

## Majorizing Sequences and Error Bounds for Iterative Methods

By George J. Miel

**Abstract.** Given a sequence  $\{x_n\}_{n=0}^\infty$  in a Banach space, it is well known that if there is a sequence  $\{t_n\}_{n=0}^\infty$  such that  $\|x_{n+1} - x_n\| \leq t_{n+1} - t_n$  and  $\lim t_n = t^* < \infty$ , then  $\{x_n\}_{n=0}^\infty$  converges to some  $x^*$  and the error bounds  $\|x^* - x_n\| \leq t^* - t_n$  hold. It is shown that certain stronger hypotheses imply sharper error bounds,

$$\|x^* - x_n\| \leq \frac{t^* - t_n}{(t_n - t_{n-1})^\mu} \|x_n - x_{n-1}\|^\mu \leq \frac{t^* - t_n}{(t_1 - t_0)^\mu} \|x_1 - x_0\|^\mu, \quad \mu \geq 0.$$

Representative applications to infinite series and to iterates of types  $x_n = Gx_{n-1}$  and  $x_n = H(x_n, x_{n-1})$  are given for  $\mu = 1$ . Error estimates with  $0 \leq \mu \leq 2$  are shown to be valid and optimal for Newton iterates under the hypotheses of the Kantorovich theorem. The unified convergence theory of Rheinboldt is used to derive error bounds with  $0 \leq \mu \leq 1$  for a class of Newton-type methods, and these bounds are shown to be optimal for a subclass of methods. Practical limitations of the error bounds are described.

**1. Introduction.** A majorizing sequence for a sequence  $\{x_n\}_{n=0}^\infty$  in a Banach space is a sequence of real numbers  $\{t_n\}_{n=0}^\infty$  such that  $\|x_{n+1} - x_n\| \leq t_{n+1} - t_n$ . If  $\lim t_n = t^* < \infty$  then  $\{x_n\}_{n=0}^\infty$  converges to some  $x^*$  in the space and the error bounds

$$(1.1) \quad \|x^* - x_n\| \leq t^* - t_n$$

are valid. This principle was used by Ortega [16], to present an elegant proof of the Kantorovich theorem for Newton iteration, and more generally by Rheinboldt [22], to establish a unified convergence theory for Newton-related methods. The principle was used earlier by Rall [19, Section 9], with  $x_n$  and  $t_n$  as partial sums of infinite series, to solve approximately quadratic operator equations.

This paper shows that under certain conditions, a majorizing sequence implies error bounds of type

$$(1.2) \quad \|x^* - x_n\| \leq \frac{t^* - t_n}{(t_n - t_{n-1})^\mu} \|x_n - x_{n-1}\|^\mu \leq \frac{t^* - t_n}{(t_1 - t_0)^\mu} \|x_1 - x_0\|^\mu, \quad \mu \geq 0.$$

If  $\mu = 0$  then (1.2) reduces to (1.1). For  $\mu > 0$ , the two bounds in (1.2) are sharper than the bounds in (1.1). The estimates (1.2) provide exit criteria applicable during

---

Received June 5, 1978; revised December 15, 1978 and April 30, 1979.  
 AMS (MOS) subject classifications (1970). Primary 65J05; Secondary 47H10.  
 Key words and phrases. Majorizing sequences, iterative methods, exit criteria.

automatic computation, namely, methods of stopping the calculation when an approximant  $x_n$  has a prescribed accuracy. A study of bounds of the form  $\|x^* - x_n\| \leq \alpha \|x_n - x_{n-1}\|$ ,  $\alpha$  constant, based on the behavior of successive errors  $x^* - x_n$  and  $x^* - x_{n+1}$ , is given in [11]. The approach taken here, which yields the convergence and the estimates (1.2) concurrently, is based on the behavior of successive backward differences  $x_n - x_{n-1}$  and  $x_{n+1} - x_n$ .

Conditions under which a majorizing sequence implies (1.2) are presented as two simple lemmas in the next section. No attempt is made to give a comprehensive list of consequences, but representative applications of general importance are included. An example involving the series approximation of complete elliptic integrals is given in Section 2. Section 3 presents results for iterates of types  $x_n = Gx_{n-1}$  and  $x_n = H(x_n, x_{n-1})$ ; examples involving linear iteration and a projection iterative method are given. Section 4 derives optimal error bounds for Newton's method under the hypotheses of the Kantorovich theorem; a particular case yields a simple proof of the Gragg-Tapia estimates [5]. Section 5 uses the unified convergence theory of Rheinboldt [22] to derive error bounds for a class of Newton-type methods; it is shown that these bounds are optimal for a subclass of methods, including certain processes for nondifferentiable operators. Section 6 gives a numerical example and discusses the practical difficulties inherent in the use of error bounds generated by majorizing sequences.

**2. Basic Lemmas.** In what follows,  $X$  is a Banach space and  $\nabla x_n$  denotes the backward difference  $x_n - x_{n-1}$ .

**LEMMA 2.1.** *Let  $\{x_n\}_{n=0}^\infty \subset X$ . If there is a sequence  $\{t_n\}_{n=0}^\infty$  of real numbers such that  $t_0 = 0$ ,  $t_1 \geq \|x_1 - x_0\|$ ,  $t_{n-1} < t_n$ ,  $\lim t_n = t^* < \infty$ , and*

$$\|x_{n+1} - x_n\| \leq \frac{t_{n+1} - t_n}{t_n - t_{n-1}} \|x_n - x_{n-1}\|, \quad n \geq 1,$$

*then  $\{x_n\}_{n=0}^\infty$  converges to some  $x^* \in X$  and the error bounds*

$$(2.1) \quad \|x^* - x_n\| \leq \frac{t^* - t_n}{t_n - t_{n-1}} \|x_n - x_{n-1}\| \leq \frac{t^* - t_n}{t_1} \|x_1 - x_0\|, \quad n \geq 1,$$

*are valid.*

*Proof.* If  $m \geq n$  then  $\|\nabla x_m\| \leq \nabla t_m \|\nabla x_n\| / \nabla t_n$ . Hence

$$(2.2) \quad \begin{aligned} \|x_{n+p} - x_n\| &\leq \|\nabla x_{n+p}\| + \|\nabla x_{n+p-1}\| + \cdots + \|\nabla x_{n+1}\| \\ &\leq \frac{\nabla t_{n+p} + \nabla t_{n+p-1} + \cdots + \nabla t_{n+1}}{\nabla t_n} \|\nabla x_n\| \\ &\leq \frac{t_{n+p} - t_n}{\nabla t_n} \|\nabla x_n\| \leq \frac{t_{n+p} - t_n}{\nabla t_1} \|\nabla x_1\|. \end{aligned}$$

From (2.2), we get that

$$(2.3) \quad \|x_{n+p} - x_n\| \leq t_{n+p} - t_n.$$

From (2.3), since  $\{t_n\}$  is a Cauchy sequence so is  $\{x_n\}$ . The estimates (2.1) follow from (2.2).

If a sequence  $\{t_n\}$  satisfies the hypotheses of Lemma 2.1 then by (2.3) it is a majorizing sequence and the standard bounds (1.1) hold; however, the bounds in (2.1) are sharper. Not every majorizing sequence satisfies the hypotheses of the lemma. For example, the sequences

$$x_0 = t_0 = 0, \quad x_n = \frac{1 - s^n}{1 - s}, \quad t_n = x_n + \frac{1 - r^n}{1 - r}, \quad 0 < r < s < 1,$$

are such that  $\nabla x_{n+1} \leq \nabla t_{n+1}$  and  $\nabla x_{n+1} / \nabla x_n > \nabla t_{n+1} / \nabla t_n$ .

We point out that in (2.1) the error bounds

$$E_1(n) = \frac{t^* - t_n}{t_n - t_{n-1}} \|x_n - x_{n-1}\|$$

can be considerably sharper than the bounds

$$E_2(n) = \frac{t^* - t_n}{t_1} \|x_1 - x_0\|.$$

For example, Lemma 2.1 applied on the sequences

$$x_0 = t_0 = 0, \quad x_n = 1 + \frac{1}{2}r + \dots + \frac{1}{n}r^{n-1},$$

$$t_n = 1 + r + \dots + r^{n-1}, \quad r \in (0, 1),$$

yields  $E_2(n) = nE_1(n)$ . In practice, the use of  $E_1(n)$  rather than  $E_2(n)$  as an exit criterion may very well result in a substantial saving of computational work.

*Example: Infinite Series.* Consider a series  $\sum_{n=0}^{\infty} y_n$  with elements  $y_n \in X$ . If there is a series  $\sum_{n=0}^{\infty} u_n$  of positive real numbers, convergent to  $u^* < \infty$ , and such that

$$\|y_0\| \leq u_0, \quad \|y_{n+1}\| \leq \frac{u_{n+1}}{u_n} \|y_n\|, \quad n \geq 0,$$

then the series  $\sum_{n=0}^{\infty} y_n$  converges to some  $y^* \in X$  and

$$(2.4) \quad \|y^* - (y_0 + \dots + y_n)\| \leq \frac{u^* - (u_0 + \dots + u_n)}{u_n} \|y_n\|$$

$$\leq \frac{u^* - (u_0 + \dots + u_n)}{u_0} \|y_0\|, \quad n \geq 0.$$

To prove this, apply Lemma 2.1 with  $x_0 = 0$ ,  $x_n = y_0 + \dots + y_{n-1}$  and  $t_0 = 0$ ,  $t_n = u_0 + \dots + u_{n-1}$ .

If the series  $\sum_{n=0}^{\infty} y_n$  passes the ratio test,  $\|y_{n+1}\| \leq r \|y_n\|$ ,  $r < 1$ , then the geometric series  $\|y_0\| \sum_{n=0}^{\infty} r^n$  yields the error bounds

$$(2.5) \quad \|y^* - (y_0 + \dots + y_n)\| \leq \frac{r}{1 - r} \|y_n\| \leq \frac{r^{n+1}}{1 - r} \|y_0\|.$$

The sharper bounds were used by Knopp [8, Section 34] to get error estimates for the series calculation of roots and logarithms.

We now give another application. Consider the series expansion for a complete elliptic integral of the first kind [6, p. 26]:

$$y^* \equiv \frac{2}{\pi} \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} = \sum_{n=0}^{\infty} y_n,$$

$$y_0 = 1, \quad y_n = \left[ \frac{1 \cdot 3 \cdots (2n - 1)}{2 \cdot 4 \cdots (2n)} \right]^2 k^{2n}, \quad k^2 < 1.$$

The estimates (2.5) hold with  $r = k^2$ . However, sharper bounds can be obtained by using (2.4) with the binomial series,

$$u^* \equiv (1 - k^2)^{-1/2} = \sum_{n=0}^{\infty} u_n, \quad u_0 = 1, \quad u_n = \frac{1 \cdot 3 \cdots (2n - 1)}{2 \cdot 4 \cdots (2n)} k^{2n}.$$

One thus finds that

$$y^* - (y + \cdots + y_n) \leq a_n,$$

where the bounds  $a_n$  can be computed recursively with

$$a_0 = (1 - k^2)^{-1/2} - 1, \quad a_n = \frac{2n - 1}{2n} a_{n-1} - y_n.$$

One can show that analogous bounds hold for a series expansion [6] of  $1 - (2/\pi)E(k)$ , where  $E(k)$  is a complete elliptic integral of the second kind, by using the binomial series converging to  $1 - (1 - k^2)^{1/2}$ .

LEMMA 2.2. *Let  $\{x_n\}_{n=0}^{\infty} \subset X$ . If there is a sequence  $\{t_n\}_{n=0}^{\infty}$  of real numbers such that  $t_0 = 0, t_{n-1} < t_n, \lim t_n = t^* < \infty$ ,*

$$\|x_n - x_{n-1}\| \leq t_n - t_{n-1}, \quad \|x_{n+1} - x_n\| \leq \frac{t_{n+1} - t_n}{(t_n - t_{n-1})^\lambda} \|x_n - x_{n-1}\|^\lambda, \quad n \geq 1,$$

for some constant  $\lambda \geq 1$ , then  $\{x_n\}_{n=0}^{\infty}$  converges to some  $x^* \in X$  and

$$(2.6) \quad \|x^* - x_n\| \leq \frac{t^* - t_n}{(t_n - t_{n-1})^\mu} \|x_n - x_{n-1}\|^\mu \leq \frac{t^* - t_n}{t_1^\mu} \|x_1 - x_0\|^\mu, \quad n \geq 1,$$

for every  $\mu \in [0, \lambda]$ .

*Proof.* The convergence of  $\{x_n\}$  is immediate, since  $\{t_n\}$  is assumed to be a majorizing sequence. If  $m > n$  then

$$\|\nabla x_m\| \leq \nabla t_m \left[ \frac{\|\nabla x_n\|}{\nabla t_n} \right]^{\lambda m - n} \leq \nabla t_m \left[ \frac{\|\nabla x_n\|}{\nabla t_n} \right]^\lambda \leq \nabla t_m \left[ \frac{\|\nabla x_n\|}{\nabla t_n} \right]^\mu.$$

Thus for  $p \geq 1$ ,

$$\begin{aligned}
 \|x_{n+p} - x_n\| &\leq \|\nabla x_{n+p}\| + \|\nabla x_{n+p-1}\| + \cdots + \|\nabla x_{n+1}\| \\
 (2.7) \qquad &\leq (\nabla t_{n+p} + \nabla t_{n+p-1} + \cdots + \nabla t_{n+1}) \left[ \frac{\|\nabla x_n\|}{\nabla t_n} \right]^\mu \\
 &\leq \frac{t_{n+p} - t_n}{(\nabla t_n)^\mu} \|\nabla x_n\|^\mu \leq \frac{t_{n+p} - t_n}{t_1^\mu} \|\nabla x_1\|^\mu.
 \end{aligned}$$

To obtain the last inequality, we used the conditions

$$\|\nabla x_n\| \leq \nabla t_n \left[ \frac{\|\nabla x_1\|}{\nabla t_1} \right]^{\lambda^{n-1}} \leq \nabla t_n \left[ \frac{\|\nabla x_1\|}{\nabla t_1} \right] \quad \text{and} \quad \mu \geq 0.$$

Finally, the estimates (2.6) follow from (2.7) as  $p \rightarrow \infty$ .

Observe that when  $\mu = 0$ , (2.6) yields the standard bounds (1.1) and that the best bounds in (2.6) occur with  $\mu = \lambda$ .

**3. Successive Approximations.** In this section, we use Lemma 2.1 to obtain the convergence and error bounds for iterates of types  $x_n = Gx_{n-1}$  and  $x_n = H(x_n, x_{n-1})$ .

**THEOREM 3.1.** *Consider a mapping  $G: D \subset X \rightarrow X$  such that*

$$(3.1) \qquad \|G(Gx) - Gx\| \leq \psi(\|Gx - x\|) \cdot \|Gx - x\|, \quad x, Gx \in D,$$

where  $\psi: [0, \infty) \rightarrow (0, \infty)$  is a nondecreasing function. Assume that for some  $x_0 \in D$  the iterates  $x_n = Gx_{n-1}$  remain in  $D$  and that the sequence  $\{t_n\}_{n=0}^\infty$  defined by

$$(3.2) \qquad t_0 = 0, \quad \|x_1 - x_0\| \leq t_1 \neq 0, \quad t_{n+1} = t_n + \nabla t_n \psi(\nabla t_n),$$

converges to  $t^* < \infty$ . Then the sequence  $\{x_n\}_{n=0}^\infty$  converges to some  $x^* \in X$  and the error bounds (2.1) are valid.

*Proof.* Consider the condition

$$\|x_{i+1} - x_i\| \leq \frac{t_{i+1} - t_i}{t_i - t_{i-1}} \|x_i - x_{i-1}\|.$$

This condition is true for  $i = 1$ . Suppose that it holds for  $i \leq n - 1$ . This hypothesis implies that

$$\|x_n - x_{n-1}\| \leq \frac{t_n - t_{n-1}}{t_1 - t_0} \|x_1 - x_0\| \leq t_n - t_{n-1}.$$

Thus (3.1), the monotonicity of  $\psi$ , and (3.2) imply that

$$\|x_{n+1} - x_n\| \leq \psi(\|\nabla x_n\|) \|x_n - x_{n-1}\| \leq \psi(\nabla t_n) \|x_n - x_{n-1}\| \leq \frac{\nabla t_{n+1}}{\nabla t_n} \|x_n - x_{n-1}\|.$$

The result follows from Lemma 2.1.

The above theorem, which is related to 12.4.3 in [17], can be extended for

more general functions  $\psi$  involving several parameters. This is done implicitly in the proofs of Theorems 4.1, 5.1, 5.2 in the sequel.

If, in Theorem 3.1,  $x^* \in D$  and  $G$  is continuous at  $x^*$  then  $x^* = Gx^*$ . If  $G$  is contractive with contraction factor  $r < 1$  then  $\psi(t) = r$  and  $t_n = t_1(1 + r + \cdots + r^{n-1})$ ; we thus get the standard estimates

$$(3.3) \quad \begin{aligned} \|x^* - x_n\| &\leq \frac{r}{1-r} \|x_n - x_{n-1}\| \\ &\leq \frac{r^n}{1-r} \|x_1 - x_0\|. \end{aligned}$$

The example below shows that it is sometimes useful to change norms in order to find a majorizing sequence.

*Example: Linear Iteration.* Given  $a \in X$  and a bounded linear operator  $A: X \rightarrow X$ , consider the equation  $x = Ax + a$ . If  $\|A^N\| \leq s \in (0, 1)$  for some positive integer  $N$ , then for any  $x_0 \in X$  the iterates  $x_n = Ax_{n-1} + a$  converge to a unique solution  $x^* = Ax^* + a$  and

$$(3.4) \quad \|x^* - x_n\| \leq \frac{r}{1-r} \left( 1 + \frac{\|A\|}{r} + \cdots + \frac{\|A^{N-1}\|}{r^{N-1}} \right) \|x_n - x_{n-1}\|, \quad r = s^{1/N}.$$

To prove this, define a new norm

$$|x| = r^{N-1} \|x\| + r^{N-2} \|Ax\| + \cdots + \|A^{N-1}x\|.$$

This norm is equivalent to  $\|\cdot\|$  since

$$(3.5) \quad r^{N-1} \|x\| \leq |x| \leq (r^{N-1} + r^{N-2} \|A\| + \cdots + \|A^{N-1}\|) \|x\|.$$

We have

$$|Ax| = r(|x| - r^{N-1} \|x\|) + \|A^N x\| \leq r|x|.$$

Thus (3.3) holds with the norm  $|\cdot|$ . Now use (3.5) to get (3.4).

We note in passing that the theory of partial orderings can also be used to derive exit criteria for linear iterations; see [12].

Let  $H: D \times D \rightarrow X$ , where  $D \subset X$ , and consider the equation

$$(3.6) \quad x = H(x, x).$$

We define successive approximations  $x_n$  to a solution of (3.6) as solutions of equations

$$(3.7) \quad x_n = H(x_n, x_{n-1}), \quad n \geq 1,$$

where  $x_0$  is some element of  $X$ .

**THEOREM 3.2.** *Assume that each equation (3.7) has a solution  $x_n$  in  $D$  and that*

$$(3.8) \quad \|H(x, y) - H(y, z)\| \leq r\|x - y\| + \psi(\|y - z\|) \cdot \|y - z\|, \quad x, y, z \in D,$$

where  $r < 1$  and  $\psi: [0, \infty) \rightarrow (0, \infty)$  is a nondecreasing function. If the sequence

$\{t_n\}_{n=0}^\infty$  defined by

$$t_0 = 0, \quad \|x_1 - x_0\| \leq t_1 \neq 0, \quad t_{n+1} = t_n + \frac{\nabla t_n}{1-r} \psi(\nabla t_n),$$

converges to  $t^* < \infty$ , then the sequence  $\{x_n\}_{n=0}^\infty$  converges to some  $x^* \in X$  and the error bounds (2.1) are valid.

*Proof.* Condition (3.8) implies that

$$\|x_{n+1} - x_n\| \leq r\|x_{n+1} - x_n\| + \psi(\|x_n - x_{n-1}\|) \cdot \|x_n - x_{n-1}\|,$$

$$\|x_{n+1} - x_n\| \leq \frac{\psi(\|x_n - x_{n-1}\|)}{1-r} \|x_n - x_{n-1}\|.$$

An induction argument similar to that for Theorem 3.1 yields the result.

If  $x^* \in D$  and  $H$  is continuous at  $(x^*, x^*)$  then  $x^*$  is a solution of (3.6).

*Example: Projection Iterative Method.* Given an operator  $T: X \rightarrow X$ , consider the equation

$$(3.9) \quad x = Tx.$$

Define

$$H(x, y) = PTx + QTy,$$

where  $P$  is a linear projection mapping  $X$  onto some subspace  $Y$ ,  $Q = I - P$ , and  $I$  is the identity operator. The successive approximations to a solution of (3.9) are solutions of the equations

$$(3.10) \quad x_n = PTx_n + QTx_{n-1}, \quad n \geq 1.$$

Each approximant  $x_n$  may be obtained by first solving an equation over the subspace  $Y$  and then translating the solution: at the  $n$ th step, when  $x_{n-1}$  is known, solve for  $y \in Y$  in the equation

$$y = PTy + PTa, \quad a = QTx_{n-1},$$

and then let  $x_n = y + a$ .

If  $PT$  and  $QT$  are contractive operators on  $X$  such that

$$\|PTx - PTy\| \leq r\|x - y\|, \quad \|QTx - QTy\| \leq s\|x - y\|, \quad r + s < 1,$$

then for any  $n$  and  $x_0 \in X$  the equations (3.10) are uniquely solvable for  $x_n$ , the sequence  $\{x_n\}_{n=0}^\infty$  converges to a unique solution  $x^*$  of (3.9), and the estimates

$$\|x^* - x_n\| \leq \frac{s}{1-r-s} \|x_n - x_{n-1}\| \leq \frac{s^n}{(1-r)^{n-1}(1-r-s)} \|x_1 - x_0\|$$

are valid.

The proof is straightforward. The unique solvability of (3.10) follows from the hypothesis that  $PT$  is contractive. Apply Theorem 3.2 with  $\psi(t) = s$ , thus obtaining

the majorizing sequence

$$t_n = t_1 \left[ 1 + \left( \frac{s}{1-r} \right) + \cdots + \left( \frac{s}{1-r} \right)^{n-1} \right].$$

Finally, a standard argument shows that  $\lim x_n = x^*$  is a unique solution of (3.9).

The above result is contained in Theorem 9.1 of Kurpel' [9]. Similar results for other projection iterative methods hold. Theorem 3.2 with  $\psi(t) = s < 1 - r$  is a special case of Theorems 8.1 and 14.1 in [9].

**4. Newton Iteration.** This section uses Lemma 2.2 to derive optimal error bounds for Newton's method under the hypotheses of the Kantorovich theorem [7], [16]. In what follows,  $X$  and  $Y$  are Banach spaces,  $D$  is an open convex subset of  $X$ ,  $F: D \rightarrow Y$  is Fréchet differentiable on  $D$  and such that

$$\|F'(x) - F'(y)\| \leq K\|x - y\|, \quad x, y \in D.$$

The open ball  $\{x: \|x - x_0\| < r\}$  and its closure are denoted by  $S(x_0, r)$  and  $\bar{S}(x_0, r)$  respectively.

**THEOREM 4.1.** *For some  $x_0 \in D$ , assume that  $F(x_0)^{-1}$  is defined on all of  $Y$ , and that*

$$\|F(x_0)^{-1}Fx_0\| \leq a, \quad \|F(x_0)^{-1}\| \leq b, \quad \bar{S}(x_0, t^*) \subset D,$$

where  $t^* = 2a(1 - \sqrt{1-h})/h$ ,  $h = 2Kab \leq 1$ . Then the iterates  $x_{n+1} = x_n - F'(x_n)^{-1}Fx_n$  exist, remain in  $S(x_0, t^*)$ , and converge to the only root  $x^*$  of  $F$  in  $D \cap S(x_0, t^{**})$ , where  $t^{**} = 2a(1 + \sqrt{1-h})/h$ . Moreover, the constants defined recursively by

$$(4.1) \quad A_1 = \frac{1}{a}B_1, \quad A_{n+1} = A_n(2 - \Delta A_n), \quad \Delta = t^{**} - t^*,$$

$$(4.2) \quad B_1 = \frac{t^*}{a} - 1 = \frac{t^*}{t^{**}}, \quad B_{n+1} = B_n^2,$$

$$(4.3) \quad C_1 = B_1, \quad C_{n+1} = \frac{C_n^2}{2C_n + \Delta/a},$$

are such that the error bounds,

$$(4.4) \quad \|x^* - x_n\| \leq A_n \|x_n - x_{n-1}\|^2 \leq B_n \|x_n - x_{n-1}\| \leq C_n \|x_1 - x_0\|,$$

are valid and best possible.

*Proof.* First, we outline a standard proof [16] of the convergence. If  $x \in S(x_0, t^*)$  then  $Gx = x - F'(x)^{-1}Fx$  exists. If also  $Gx \in S(x_0, t^*)$  then

$$(4.5) \quad \|G(Gx) - Gx\| \leq \frac{bK\|Gx - x\|/2}{1 - bK\|Gx - x_0\|} \|Gx - x\|.$$

Consider the scalar function

$$f(t) = \frac{K}{2} t^2 - \frac{1}{b} t + \frac{a}{b} = \frac{K}{2} (t^* - t)(t^{**} - t)$$

and its Newton iterates

$$(4.6) \quad t_0 = 0, \quad t_n = t_{n-1} - f'(t_{n-1})^{-1} f(t_{n-1}).$$

The sequence  $\{t_n\}$  satisfies the conditions  $\|x_1 - x_0\| \leq t_1 = a$ ,  $t_{n-1} < t_n$ ,  $\lim t_n = t^*$ , and

$$(4.7) \quad \frac{bK/2}{1 - bKt_n} = \frac{t_{n+1} - t_n}{(t_n - t_{n-1})^2}.$$

An induction argument shows that  $\{x_n\}$  exists and that  $\{t_n\}$  is a majorizing sequence. Hence

$$(4.8) \quad \|x_{n+p} - x_n\| \leq t_{n+p} - t_n$$

and  $\{x_n\}$  converges to some  $x^*$ . That  $Fx^* = 0$  follows from the continuity at  $x^*$  of  $F$  and  $F'$ . The uniqueness statement follows by consideration of the simplified Newton method.

At this point, use (4.5) with  $x = x_{n-1}$ , (4.8), and (4.7) to get

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \frac{bK/2}{1 - bK\|x_n - x_0\|} \|x_n - x_{n-1}\|^2 \leq \frac{bK/2}{1 - bKt_n} \|x_n - x_{n-1}\|^2 \\ &\leq \frac{t_{n+1} - t_n}{(t_n - t_{n-1})^2} \|x_n - x_{n-1}\|^2. \end{aligned}$$

Thus the sequence  $\{t_n\}$  satisfies the hypotheses of Lemma 2.2 with  $\lambda = 2$ . Taking  $\mu = 2$  and  $\mu = 1$ , the bounds (4.4) are valid with

$$(4.9) \quad A_n = \frac{t^* - t_n}{(t_n - t_{n-1})^2}, \quad B_n = \frac{t^* - t_n}{t_n - t_{n-1}}, \quad C_n = \frac{t^* - t_n}{t_1}.$$

The second inequality in (4.4) follows from the inequality  $\|\nabla x_n\|/\nabla t_n \leq 1$ .

We now show that the constants defined in (4.9) satisfy the recurrence relations (4.1), (4.2), (4.3). Let  $e_n = t^* - t_n$  and  $E_n = t^{**} - t_n$ ; then (4.6) yields

$$(4.10) \quad \nabla t_{n+1} = \frac{e_n E_n}{e_n + E_n}, \quad e_{n+1} = \frac{e_n^2}{e_n + E_n}, \quad E_{n+1} = \frac{E_n^2}{e_n + E_n}.$$

Since  $A_{n+1} = e_{n+1}/(\nabla t_{n+1})^2 = E_{n+1}^{-1}$  and  $E_{n+1} = E_n^2/(2E_n - \Delta)$ , (4.1) follows. Since  $B_{n+1} = e_{n+1}/\nabla t_{n+1} = e_n/E_n$  and  $e_{n+1}/E_{n+1} = (e_n/E_n)^2$ , we get (4.2). Relation (4.3) follows from  $C_n = e_n/a$  and  $e_{n+1} = e_n^2/(2e_n + \Delta)$ . The bounds (4.4) with constants (4.9) are optimal since they are clearly attained with  $F = f$ ,  $x_0 = t_0$ . This completes the proof.

It is possible to derive concurrently the convergence of  $\{x_n\}$  and the two right-most error bounds in (4.4) by an induction argument and the use of Lemma 2.1.

Note, however, that the bounds  $A_n \|x_n - x_{n-1}\|^2$  are sharper than the other two bounds in (4.4). The recurrence relations (4.1), (4.2), (4.3) provide a very convenient way of calculating the error estimates during on-line computation. The following corollary gives explicit forms for the constants.

**COROLLARY 4.2.** *Under the hypotheses of the theorem, the constants in (4.4) are given by*

$$A_n = \frac{2^{n-1}}{a}, \quad B_n = 1, \quad C_n = 2^{-n+1},$$

if  $h = 1$  and by

$$A_n = \frac{h}{4a\sqrt{1-h}} (1 - \theta^{2n}), \quad B_n = \theta^{2^{n-1}}, \quad C_n = \frac{4\sqrt{1-h}}{h} \cdot \frac{\theta^{2n}}{1 - \theta^{2n}}, \quad \theta = \frac{t^*}{t^{**}},$$

if  $h < 1$ .

*Proof.* As before, let  $e_n = t^* - t_n$ . From (4.10),

$$(4.11) \quad e_0 = t^*, \quad e_{n+1} = \frac{e_n^2}{2e_n + \Delta}, \quad \Delta = t^{**} - t^*.$$

If  $h = 1$ , then  $\Delta = 0$  and

$$(4.12) \quad e_n = 2^{-n+1}a.$$

If  $h < 1$ , let  $e_n = \Delta(s_n - 1)/2$  to get  $s_{n+1} = (s_n^2 + 1)/2s_n$ . As in [5],

$$(4.13) \quad \frac{s_{n+1} - 1}{s_{n+1} + 1} = \left( \frac{s_n - 1}{s_n + 1} \right)^2, \quad \frac{s_n - 1}{s_n + 1} = \left( \frac{s_0 - 1}{s_0 + 1} \right)^{2^n} = \theta^{2^n},$$

and consequently,

$$(4.14) \quad e_n = \frac{4\sqrt{1-h}}{h} \cdot \frac{\theta^{2n}}{1 - \theta^{2n}} a.$$

The explicit form for  $C_n$  follows since  $C_n = e_n/a$ . Use  $A_{n+1} = E_{n+1}^{-1} = (e_{n+1} + \Delta)^{-1}$  and (4.12), (4.14) to obtain the explicit form for  $A_n$ . Finally, solve directly the difference equation (4.2) to get  $B_n = \theta^{2^{n-1}}$ .

An interesting background of the Kantorovich theorem is given in [17, p. 428]. The convergence argument outlined in the proof of Theorem 4.1 is due to Ortega [16]. Under hypotheses different from the usual ones, Ostrowski [18] proved the optimality of bounds  $\|x^* - x_n\| \leq D_n \|x_1 - x_0\|$ . Gragg and Tapia [5] derived the explicit forms of the optimal constants  $B_n$  and  $C_n$  stated in Corollary 4.2. These explicit expressions follow from their explicit solution of the nonlinear difference equation (4.6). We used their argument in (4.13) in order to solve (4.11). Their derivation of  $B_n$  relies on the original recurrence relations of Kantorovich [7]. A simpler argument, which works directly with the majorizing sequence, is given in [13]. The above proof based on Lemma 2.2 is even simpler and it also yields the sharper bounds  $A_n \|x_n - x_{n-1}\|^2$  and the

inequalities in (4.4) between the bounds themselves. These inequalities, as well as the optimal error estimates presented below, seem to be new.

COROLLARY 4.3. *Assume that the conditions of Theorem 4.1 hold. Let*

$$S_n(\mu) = (2^{-n+1}a)^{1-\mu}, \quad T_n(\mu) = 2^{-n+1}a^{1-\mu},$$

if  $h = 1$  and

$$S_n(\mu) = \left[ \frac{4a\sqrt{1-h}}{h(1-\theta^{2^n})} \right]^{1-\mu} \theta^{(2-\mu)2^{n-1}}, \quad T_n(\mu) = \frac{4\sqrt{1-h}}{h} \cdot \frac{\theta^{2^n}}{1-\theta^{2^n}} \cdot a^{1-\mu},$$

if  $h < 1$ . Then the error bounds

$$(4.15) \quad \|x^* - x_n\| \leq S_n(\mu) \|x_n - x_{n-1}\|^\mu \leq T_n(\mu) \|x_1 - x_0\|^\mu$$

are valid and best possible for each  $\mu \in [0, 2]$ .

*Proof.* From Lemma 2.2,  $S_n(\mu) = (t^* - t_n)/(t_n - t_{n-1})^\mu = e_n/(e_{n-1} - e_n)^\mu$  and  $T_n(\mu) = (t^* - t_n)/t_1^\mu = e_n/a^\mu$ . Use (4.12) and (4.14) to obtain the desired expressions.

Observe that if  $\mu = 0$  then (4.15) yields the standard bounds in (1.1) with the closed forms (4.12) and (4.14) for  $e_n = t^* - t_n$ , that the sharpest bounds in (4.15) occur with  $\mu = 2$ , and that the constants in Corollary 4.2 are given by  $A_n = S_n(2)$ ,  $B_n = S_n(1)$ , and  $C_n = T_n(1)$ .

**5. Newton-Type Iterations.** The results of Section 2 are used to derive error bounds for a class of Newton-related methods studied by Rheinboldt [22, Theorem 4.3]. In what follows, the operator  $F$  is as in Section 4 and  $L(X, Y)$  denotes the Banach space of bounded linear operators with domain  $X$  and range in  $Y$ .

THEOREM 5.1. *Assume that  $\mathcal{D}: D \rightarrow L(X, Y)$  and  $x_0 \in D$  are such that  $\|\mathcal{D}(x) - \mathcal{D}(x_0)\| \leq L\|x - x_0\|$ ,  $\|F'(x) - \mathcal{D}(x)\| \leq M\|x - x_0\| + c$ ,  $x \in D$  and  $M, c \geq 0$ ,  $\mathcal{D}(x_0)^{-1} \in L(Y, X)$  exists,  $\|\mathcal{D}(x_0)^{-1}\| \leq b$ ,  $\|\mathcal{D}(x_0)^{-1}Fx_0\| \leq a$ , and  $\bar{S}(x_0, t^*) \subset D$  where*

$$t^* = \frac{2a}{(1-bc)h} (1 - \sqrt{1-h}), \quad h = \frac{2\sigma Kab}{(1-bc)^2} \leq 1,$$

$$\sigma = \max\left(1, \frac{L+M}{K}\right), \quad bc < 1.$$

Define

$$t^{**} = \frac{2a}{(1-bc)h} (1 + \sqrt{1-h}), \quad \hat{t} = \frac{2\sigma a}{(1-bc)h} \left(1 + \sqrt{1 + \frac{h}{\sigma}}\right), \quad \Delta = t^{**} - t^*.$$

Then the iterates  $x_{n+1} = x_n - \mathcal{D}(x_n)^{-1}Fx_n$  exist, remain in  $S(x_0, t^*)$ , and converge to the only root  $x^*$  of  $F$  in  $D \cap S(x_0, \hat{t})$ . Moreover, the following error bounds are valid:

$$(5.1) \quad \|x^* - x_n\| \leq B_n \|x_n - x_{n-1}\| \leq C_n \|x_1 - x_0\|,$$

where the constants are defined recursively by

$$(5.2) \quad B_1 = \frac{t^*}{a} - 1, \quad B_{n+1} = \frac{(LQ - PR)B_n^2 + Q(L - P)B_n - PQ}{(LR + PQ - 2PR)B_n + P(Q - R)},$$

$$(5.3) \quad C_1 = B_1, \quad C_{n+1} = \frac{C_n(aPC_n + Q)}{aLC_n + R},$$

with

$$P = L - \frac{\sigma K}{2}, \quad Q = R - \frac{\sigma K \Delta}{2}, \quad R = \frac{1}{b} - Lt^*.$$

If  $K = L + M$  then these error bounds are best possible.

*Proof.* First, we outline and rearrange slightly the proof of the convergence in [17, Section 12.6]. If  $x \in S(x_0, t^*)$  then  $Gx = x - \mathcal{D}(x)^{-1}Fx$  exists. If also  $Gx \in S(x_0, t^*)$  then

$$(5.4) \quad \|G^2x - Gx\| \leq \psi(\|Gx - x\|, \|Gx - x_0\|, \|x - x_0\|) \cdot \|Gx - x\|,$$

where

$$\psi(u, v, w) = \frac{b}{1 - bLv} \left( \frac{\sigma K}{2} u + c + (\sigma K - L)w \right).$$

Consider the scalar functions

$$(5.5) \quad f(t) = \frac{\sigma K}{2} t^2 - \frac{(1 - bc)t}{b} + \frac{a}{b} = \frac{\sigma K}{2} (t^* - t)(t^{**} - t), \quad d(t) = Lt - \frac{1}{b},$$

and the iterates

$$(5.6) \quad t_0 = 0, \quad t_{n+1} = t_n - d(t_n)^{-1}f(t_n).$$

The sequence  $\{t_n\}$  satisfies the conditions  $\|x_1 - x_0\| \leq t_1 = a$ ,  $t_{n-1} < t_n$ ,  $\lim t_n = t^*$ , and

$$(5.7) \quad \psi(\nabla t_n, t_n, t_{n-1}) = \nabla t_{n+1} / \nabla t_n.$$

An induction argument shows that  $\{x_n\}$  exists and that  $\{t_n\}$  is a majorizing sequence. Thus

$$(5.8) \quad \|x_{n+p} - x_n\| \leq t_{n+p} - t_n$$

and  $\lim x_n = x^*$  exists. A standard argument shows that  $x^*$  is a unique root of  $F$  in the indicated domain.

Now, use (5.4) with  $x = x_{n-1}$ , (5.8) and the monotonicity of  $\psi$  in each variable, and then (5.7) to get

$$\|x_{n+1} - x_n\| \leq \frac{\nabla t_{n+1}}{\nabla t_n} \|x_n - x_{n-1}\|.$$

Hence, the sequence  $\{t_n\}$  satisfies the hypotheses of Lemma 2.1 (or Lemma 2.2 with

$\lambda = 1$ ), and so, the estimates (5.1) are valid with

$$(5.9) \quad B_n = \frac{t^* - t_n}{t_n - t_{n-1}}, \quad C_n = \frac{t^* - t_n}{t_1}.$$

If  $K = L + M$  then these estimates are optimal since they are attained with  $F = f$ ,  $\mathcal{D} = d$ ,  $x_0 = t_0$ .

We now show that the constants (5.9) satisfy the recurrence relations (5.2) and (5.3). Letting  $e_n = t^* - t_n$  in (5.5) gives

$$f(t_n) = \frac{\alpha K}{2} e_n(e_n + \Delta) \quad \text{and} \quad d(t_n) = -(Le_n + R);$$

(5.6) then yields

$$(5.10) \quad e_{n+1} = e_n(Pe_n + Q)/(Le_n + R).$$

Relation (5.3) follows since  $C_n = e_n/a$ . Use (5.10) to get

$$B_n = \frac{e_n}{e_{n-1} - e_n} = \frac{Pe_{n-1} + Q}{(L - P)e_{n-1} + (R - Q)},$$

and then

$$(5.11) \quad e_{n-1} = \frac{(Q - R)B_n + Q}{(L - P)B_n - P}, \quad \frac{Pe_{n-1} + Q}{Le_{n-1} + R} = \frac{B_n}{B_n + 1}.$$

From (5.10) and (5.11), we obtain

$$(5.12) \quad e_n = \frac{(Q - R)B_n + Q}{(L - P)B_n - P} \cdot \frac{B_n}{B_n + 1}.$$

Finally, put (5.12) in

$$B_{n+1} = \frac{Pe_n + Q}{(L - P)e_n + (R - Q)},$$

to get the desired relation (5.2). This completes the proof.

The use of Lemma 2.2 in the above proof yields the estimates

$$(5.13) \quad \|x^* - x_n\| \leq \frac{t^* - t_n}{(t_n - t_{n-1})^\mu} \|x_n - x_{n-1}\|^\mu \leq \frac{t^* - t_n}{t_1^\mu} \|x_1 - x_0\|^\mu, \quad 0 \leq \mu \leq 1,$$

where the majorizing sequence is defined by (5.6). The sharpest case occurs with  $\mu = 1$  and the constants defined by (5.2) and (5.3) correspond to this case. The sharper bound in (5.13) with  $\mu = 1$  was originally derived [14] by extending the argument given in [13] for Newton's method. The above proof of the more general estimates (5.13) is considerably simpler. The recurrence relations (5.2) and (5.3), which provide convenient means of calculating the error estimates during automatic computation, also appear to be new.

Important special cases of Theorem 5.1 include the Newton, simplified Newton, parallel chord, Newton-Jacobi, and Newton-SOR methods [17]. It turns out that the

error bounds are best possible for the first three methods. Practical aspects of the implementation of the bounds in computer programs are discussed in [14, Section 6]. For the Newton case, when  $\mathcal{D}(x) = F'(x)$ , we have  $L = K$ ,  $M = c = 0$ ,  $\sigma = 1$ ,  $P = K/2$ ,  $R = \Delta P$ ,  $Q = 0$  and the recurrence relations (5.2) and (5.3) reduce to (4.2) and (4.3) respectively. Hence Theorem 5.1 contains Theorem 4.1 with the two rightmost bounds in (4.4).

Zinčenko [25] has shown that the differentiability condition on  $F$  assumed in Theorem 5.1 can be replaced by corresponding conditions on  $\mathcal{D}$ . His result was proved with majorizing sequences by Rheinboldt [22, Theorem 4.5]. Our last theorem incorporates optimal error bounds in this result. Suppose now that  $F: D \rightarrow Y$  is continuous and let  $Z: D \rightarrow Y$  be Fréchet differentiable and such that

$$\begin{aligned} \|Z'(x) - Z'(y)\| &\leq K\|x - y\|, & \|H(x) - H(y)\| &\leq c\|x - y\|, \\ H &= F - Z, & x, y &\in D. \end{aligned}$$

**THEOREM 5.2.** *Assume that  $x_0 \in D$  is such that*

$$Z'(x)^{-1} \in L(Y, X) \text{ exists, } \|Z'(x_0)^{-1}Fx_0\| \leq a, \quad \|Z'(x_0)^{-1}\| \leq b, \quad bc < 1.$$

Define  $h$ ,  $t^*$ ,  $t^{**}$ ,  $\Delta$ ,  $P = K/2$ ,  $Q = c$ ,  $R = \Delta P + Q$  as in Theorem 5.1 with  $\sigma = 1$  and  $L = K$ . If  $\bar{S}(x_0, t^*) \subset D$ , then the iterates  $x_{n+1} = x_n - Z(x_n)^{-1}Fx_n$  remain in  $S(x_0, t^*)$  and converge to the only root  $x^*$  of  $F$  in  $D \cap S(x_0, t^{**})$ . Moreover, the error bounds (5.1), with  $B_n$  and  $C_n$  defined by (5.2) and (5.3), are valid and best possible.

*Proof.* From [22, p. 57], if  $x$  and  $Gx \in S(x_0, t^*)$ , then

$$\|G^2x - Gx\| \leq \psi(\|Gx - x\|, \|Gx - x_0\|) \cdot \|Gx - x\|,$$

where

$$\psi(u, v) = \frac{b}{1 - bKv} \left( \frac{1}{2}Ku + c \right).$$

The sequence defined by

$$t_0 = 0, \quad t_{n+1} = t_n - z'(t_n)^{-1}f(t_n),$$

$$f(t) = \frac{K}{2}t^2 - \frac{(1 - bc)t}{b} + \frac{a}{b} = \frac{K}{2}(t^* - t)(t^{**} - t), \quad z(t) = \frac{K}{2}t^2 - \frac{1}{b}t + \frac{a}{b},$$

is a majorizing sequence which satisfies the hypotheses of Lemma 2.1. The recurrence relations for  $B_n$  and  $C_n$  follow as before.

**6. Limitations.** This section illustrates and discusses shortcomings inherent in the use of the error bounds given by the semilocal convergence theorems resulting from Lemmas 2.1 and 2.2.

Recall that a parallel chord method has iterates of the form

$$(6.1) \quad x_{n+1} = x_n - A^{-1}Fx_n,$$

where  $A$  is a fixed operator. Using Theorem 5.1 with  $\mathcal{D}(x) = A$  and  $\|F'(x_0) - A\| \leq c$ , we get  $\|F'(x) - A\| \leq K\|x - x_0\| + c$ , and so,  $L = 0$  and  $M = K$ . Hence, provided that the pertinent hypotheses are satisfied, the iterates (6.1) converge to a root  $x^*$  and the error bounds  $\|x^* - x_n\| \leq B_n \|x_n - x_{n-1}\|$  are valid and optimal with

$$(6.2) \quad B_0 = \frac{t^*}{a} - 1, \quad B_{n+1} = \frac{B_n^2 + (1 - \xi)(B_n + 1)}{(1 + \xi)B_n + \xi},$$

where  $\xi = (1 - bc)\sqrt{1 - h}$ .

We now use the simplified Newton method in order to find a solution to

$$(6.3) \quad Fx = (2r^3 - s^2 - 1, rs^3 - s - 4) = 0, \quad x = (r, s).$$

The  $l_\infty$ -norm is used. A preliminary graphical analysis shows that the two curves intersect inside the square  $S = S(x_0, 1)$ ,  $x_0 = (1, 2)$ .

The operator  $F''(x)$  is represented by the pair of matrices

$$A_1(x) = \begin{bmatrix} 12r & 0 \\ 0 & -2 \end{bmatrix}, \quad A_2(x) = \begin{bmatrix} 0 & 3s^2 \\ 3s^2 & 6rs \end{bmatrix}.$$

Hence

$$\|F''(x)\| \leq \max(\|A_1(x)\|, \|A_2(x)\|) = \max(2, 12|r|, 3s^2 + 6|rs|).$$

For  $x \in \bar{S}$ , the maximum is reached at the corner point  $(2, 3)$ . Thus we take  $K = 63$ .

The iteration has the form (6.1) with

$$x_0 = (1, 2), \quad A = F'(x_0) = \begin{bmatrix} 6 & -4 \\ 8 & 11 \end{bmatrix}.$$

The results of the computation, done in double precision on a CDC 6600, are shown below.

$m$	$n$	$r$	$s$	$B_n \ x_n - x_{n-1}\ $	$\ x^* - x_n\ $
1		1.255 102 041	1.632 653 061		
3	0	1.235 228 392	1.660 232 811		
4	1	1.234 074 961	1.661 758 476	$.83 \times 10^{-2}$	$.23 \times 10^{-3}$
8	5	1.234 274 104	1.661 526 930	$.18 \times 10^{-4}$	$.46 \times 10^{-6}$
13	10	1.234 274 484	1.661 526 467	$.79 \times 10^{-8}$	$.19 \times 10^{-9}$
18	15	1.234 274 484	1.661 526 467	$.34 \times 10^{-11}$	$.75 \times 10^{-13}$

At steps  $m = 1, 2, 3, 4$  the program calculated  $h = 2Kab/(1 - bc)^2$ , where  $a = \|x_m - x_{m-1}\|$ ,  $b = \|A^{-1}\|$ , and  $c = \|F'(x_{m-1}) - A\|$ . The conditions  $bc < 1$  and  $h \leq 1$  were satisfied at step  $m = 4$ . At this point, the program renumbered  $x_3$  as  $x_0$  and  $x_4$  as  $x_1$ , and it kept fixed the values  $a, c$ , and  $h$ . The recurrence relation (6.2) was then used to compute the error bounds  $B_n \|x_n - x_{n-1}\|$ . The exact values of  $\|x^* - x_n\|$  are indicated in the last column.

The above numerical example illustrates three shortcomings associated with the

error bounds generated by majorizing sequences. Although the iterates of the example converge, the stringent hypotheses of the theorem do not take hold initially. In general, the iterates must be in the vicinity of a root before a semilocal theorem will guarantee convergence and provide error bounds. The second disadvantage is the amount of work involved in computing the required constants, especially the Lipschitz constant  $K$ . For most problems, the calculation of  $K$  is extremely difficult, and usually results in conservative values. For finite systems of equations, when  $F''(x)$  exists, an automatic estimation of  $K$  is in principle possible with the use of interval arithmetic [20, p. 163], but such procedures are considered too prohibitive by most people involved in programming efforts. The third shortcoming is that the values of  $B_n \|x_n - x_{n-1}\|$ , although sharper than those of  $C_n \|x_1 - x_0\|$ , are usually very pessimistic error bounds and get worse as  $n$  increases.

For Newton iterates, Lancaster [10] and Rockne [23] presented bounds which include the effects of rounding errors; an example involving the solution of (6.3) by Newton's method is given in [10]. Their analysis is generalized in [15] for the class of methods engendered by Rheinboldt's theorem. In view of the severe shortcomings of the theoretical bounds just described, and the huge effort needed to account for roundoff, such analyses are of limited practical interest.

Rheinboldt's hypotheses in Theorem 5.1 include the condition  $\|\mathcal{D}(x) - \mathcal{D}(x_0)\| \leq L \|x - x_0\|$ , which turns out to be restrictive, as there is a type of efficient iteration which violates the condition. Dennis [1, Theorem 3.2] used a majorizing sequence to extend Rheinboldt's theorem for methods based on certain derivative approximations, which have in some sense bounded deterioration. Contrary to the gloomy note in [17, NR 12.6-4], this result includes generalized secant techniques, as well as algorithms due to Broyden for systems of equations, and to Davidon for the minimization of functionals. Such algorithms initiated the considerable research that led to the so-called quasi-Newton methods recently surveyed in [2], [4].

To be competitive an algorithm should be superlinearly convergent. A characterization of this important property is given by Dennis and Moré [3]. A trait of superlinearly convergent sequences is that

$$(6.4) \quad \lim_{n \rightarrow \infty} \frac{\|x_n - x_{n-1}\|}{\|x^* - x_{n-1}\|} = 1,$$

provided that  $x_n \neq x^*$ . The asymptotic relation (6.4) provides some justification for stopping criteria involving  $\|x_n - x_{n-1}\|$ . Gragg and Tapia [5] point out that, under the hypotheses of the Kantorovich theorem, Newton iterates satisfy

$$(6.5) \quad 2(\sqrt{2} - 1)\|x_{n+1} - x_n\| \leq \|x^* - x_n\| \leq \|x_n - x_{n-1}\|.$$

We note that (6.5) refines the bounds of Dennis [1, p. 457], and that the upper bound in (6.5) follows immediately from (4.2). Rall [21] has shown that if  $x^*$  is a simple root, then there exists an open ball centered at  $x^*$  such that any  $x_0$  in that ball satisfies the Kantorovich conditions. Thus, any Newton sequence converging to a simple  $x^*$  will satisfy (6.5) after a finite number of terms.

In addition to the local nature, the difficult evaluation of the pertinent constants, and the pessimistic values of the associated error bounds, majorizing sequences have the disadvantage of usually not providing the traditional order of convergence. This is perhaps not surprising since orders of  $Q$ -convergence are stated in local theorems which assume the existence of a solution  $x^*$ , whereas majorizing sequences are used in a semilocal setting, and are hampered by the task of establishing the existence of  $x^*$ . For Newton iterates, the argument outlined in the proof of Theorem 4.1 gives only quadratic  $R$ -convergence, not the stronger quadratic  $Q$ -convergence. It is interesting to note that the latter does follow from the original Kantorovich recurrence relations proof; the distinction is subtle and depends on the difference between right and left inverses. In the general case, if a majorizing sequence  $\{t_n\}$  is  $Q$ -convergent to  $t^*$  with order  $p$ , then the related sequence  $\{x_n\}$  is  $R$ -convergent to  $x^*$  with same order  $p$ . If

$$(6.6) \quad \overline{\lim} \frac{t^* - t_n}{\|x^* - x_n\|} < \infty,$$

then  $R$ - and  $Q$ -convergence for  $\{x_n\}$  are the same. If condition (6.6) does not hold then the error bounds  $t^* - t_n$  become infinitely bad.

**Acknowledgement.** The author is grateful to the referee for raising the points, indicated in the last section, on the shortcomings of majorizing sequences.

Department of Mathematical Sciences  
University of Nevada, Las Vegas  
Las Vegas, Nevada 89154

1. J. E. DENNIS, "Toward a unified convergence theory for Newton-like methods," *Nonlinear Functional Analysis and Applications* (L. B. Rall, Ed.), Academic Press, New York, 1971.
2. J. E. DENNIS, "A brief introduction to quasi-Newton methods," *Numerical Analysis* (G. H. Golub and J. Olinger, Eds.), Proc. Sympos. Appl. Math., vol. 22, Amer. Math. Soc., Providence, R. I., 1978.
3. J. E. DENNIS & J. J. MORE', "A characterization of superlinear convergence and its application to quasi-Newton methods," *Math. Comp.*, v. 28, 1974, pp. 549-560.
4. J. E. DENNIS & J. J. MORE', "Quasi-Newton methods, motivation and theory," *SIAM Rev.*, v. 19, 1977, pp. 46-89.
5. W. B. GRAGG & R. A. TAPIA, "Optimal error bounds for the Newton-Kantorovich theorem," *SIAM J. Numer. Anal.*, v. 11, 1974, pp. 10-13.
6. H. HANCOCK, *Elliptic Integrals*, Dover, New York, 1958.
7. L. V. KANTOROVICH, "Functional analysis and applied mathematics," *Uspehi Mat. Nauk*, v. 3, 1948, pp. 89-185 (Russian); English transl., Rep. 1509, National Bureau of Standards, Washington, D. C., 1952.
8. K. KNOPP, *Theory and Application of Infinite Series*, Blackie & Son Ltd., London and Glasgow, 1928.
9. N. S. KURPEL', *Projection-Iterative Methods for Solution of Operator Equations*, Transl. Math. Monographs, vol. 46, Amer. Math. Soc., Providence, R. I., 1976.
10. P. LANCASTER, "Error analysis for the Newton-Raphson method," *Numer. Math.*, v. 9, 1966, pp. 55-68.
11. G. J. MIEL, "On a posteriori error estimates," *Math. Comp.*, v. 31, 1977, pp. 204-213.
12. G. J. MIEL, "Cones and error bounds for linear iterations," *Aequationes Math.* (To appear.)
13. G. J. MIEL, "The Kantorovich theorem with optimal error bounds," *Amer. Math. Monthly*, v. 86, 1979, pp. 212-215.
14. G. J. MIEL, *Exit Criteria for Newton-Type Iterations*, Research Paper No. 363, Dept. of Math. and Stat., Univ. of Calgary, 1977.

15. G. J. MIEL, "Unified error analysis for Newton-type methods," *Numer. Math.* (To appear.)
16. J. M. ORTEGA, "The Newton-Kantorovich theorem," *Amer. Math. Monthly*, v. 75, 1968, pp. 658-660.
17. J. M. ORTEGA & W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
18. A. M. OSTROWSKI, "La méthode de Newton dans les espaces de Banach," *C. R. Acad. Sci. Paris Sér. A.*, v. 272, 1971, pp. 1251-1253.
19. L. B. RALL, "Quadratic equations in Banach spaces," *Rend. Circ. Mat. Palermo*, v. 10, 1961, pp. 314-332.
20. L. B. RALL, *Computational Solution of Nonlinear Operator Equations*, Wiley, New York, 1969.
21. L. B. RALL, "A note on the convergence of Newton's method," *SIAM J. Numer. Anal.*, v. 11, 1974, pp. 34-36.
22. W. C. RHEINBOLDT, "A unified convergence theory for a class of iterative processes," *SIAM J. Numer. Anal.*, v. 5, 1968, pp. 42-63.
23. J. ROCKNE, "Newton's method under mild differentiability conditions with error analysis," *Numer. Math.*, v. 18, 1972, pp. 401-412.
24. A. H. SHERMAN, "On Newton-iterative methods for the solution of systems of nonlinear equations," *SIAM J. Numer. Anal.*, v. 15, 1978, pp. 755-771.
25. A. I. ZINČENKO, "A class of approximate methods for solving operator equations with nondifferentiable operators," *Dopovīdi Akad. Nauk Ukraїn. RSR*, 1963, pp. 852-855. (Ukrainian)