

Technical Report TR-66-31
NsG-398

June 1966

On a Class of Approximate
Iterative Processes

by

James M. Ortega
Research Associate Professor
and

Werner C. Rheinboldt
Research Professor
Computer Science Center
University of Maryland
College Park, Maryland

CONTENTS

ABSTRACT

I.	Introduction	1
II.	Preliminaries and Lemmas	2
III.	Approximate Iterations	8
IV.	Implicit Iterations	15
V.	Extension to N-metric Spaces	22
VI.	References	28

ABSTRACT

31372

If F is a contractive mapping, then under certain conditions the sequence $x_{n+1} = Fx_n$ tends to the unique fixed-point of F . Because of rounding or discretization error in the numerical evaluation of F , an approximate sequence $\{y_n\}$ is in general produced in place of the exact sequence $\{x_n\}$. In this paper we combine results of Ehrmann, Ostrowski, Schmidt and Urabe, which deal with the behavior of the approximate sequence $\{y_n\}$, into a unified setting and give extensions of some of their results. The discussion is set in terms of spaces metricized by elements of a partially-ordered topological linear space.

On a Class of Approximate Iterative Processes

James M. Ortega and Werner C. Rheinboldt¹⁾

1. Introduction

The problem of approximating a solution of the fixed-point equation $x = Fx$ is closely connected with the iterative process $x_{n+1} = Fx_n$, $n = 0, 1, \dots$. Because of rounding or discretization error in the evaluation of F , an approximate sequence $\{y_n\}$ is in general produced in place of the exact sequence $\{x_n\}$ and, in a variety of settings, the effect of this error has been investigated in recent years by several authors (see, e.g., Ehrmann [4], Gardner [5], Ostrowski [10], Schmidt [14], Urabe [18], [19], Warga [22] and Zincenko [23]).

Urabe [18] studies iterations of the form $y_{n+1} = F_0 y_n$ where F_0 is "close" to F and where the deviation of F_0 from F is assumed to be caused by rounding error. Ostrowski assumes that the sequence y_n has the property that y_{n+1} differs from Fy_n by a quantity ϵ_n which tends to zero. Ehrmann considers iterations of the form $y_{n+1} = F_n y_n$ where, for example, the F_n

¹⁾ This work was supported in part by NASA grant Nsg-398 to the Computer Science Center and in part by NSF grant PIVRO6 to the Institute for Fluid Dynamics and Applied Mathematics at the University of Maryland.

may be the first n terms of a power series expansion for F . Schmidt and Warga assume that the y_n are themselves fixed-points of such operators F_n ; i.e. that $y_n = F_n y_n$. In the present work we combine these somewhat disparate results into a unified setting which exhibits the basic underlying principles. This in turn permits some extensions of the previous results.

Our discussion is set in terms of spaces which are metrized by elements of a partially-ordered topological linear space N . This is a natural extension of investigations initiated by Kantorovich (see, e.g. [6], [7]) and is essentially the same setting used by Ehrmann and Schmidt.

In Section 2 we collect some basic definitions and preliminary results. In Sections 3 and 4 we give several convergence theorems and relate these to some of the results mentioned previously. For simplicity, we assume in these sections that $N = R^m$, or in other words that the spaces are metricized by elements of the real m -dimensional Euclidean space. Finally in Section 5 we show how the results extend to more general spaces.

2. Preliminaries and Lemmas

Let N be a real linear space and C a convex cone with vertex zero; i.e. $tC \subset C$ for all $t \geq 0$, $C + C \subset C$ and $C \cap -C = 0$. Then

the relation " $x \leq y$ if and only if $y-x \in C$ " introduces a partial ordering in N which is compatible with the linear structure. If, in addition, N is a topological linear space with a locally-convex topology T under which C is closed, then we call N a partially ordered topological linear space (PTL space) with respect to C and T .

Definition 1: Let X be any nonempty set and N a PTL space.

Then a mapping ρ from $X \times X$ into N such that (i) $\rho(x,y) = 0$ if and only if $x=y$ for all $x,y \in X$ and (ii) $\rho(x,y) \leq \rho(x,z) + \rho(y,z)$ for all $x,y,z \in X$ is called an N-metric, and X is said to be an N-metric space.²⁾ If X is itself a linear space then the N-metric ρ is called invariant if $\rho(x,y) = \rho(x-y,0)$ for all $x,y \in X$.

Clearly, the two properties of an N-metric imply that $\rho(x,y) \geq 0$ and $\rho(y,x) = \rho(x,y)$. Note also that for invariant ρ we have

$$(2.1) \quad \rho(x+w, y+z) \leq \rho(x,y) + \rho(w,z) \text{ for all } x,y,w,z \in X$$

In order to deduce most of the desirable properties about

²⁾ These spaces are sometimes also called pseudometric spaces; see e.g., Collatz [3], although in his case the convergence of elements in N is not defined by a topology.

monotonic convergence in N , such as the existence of the limit of a bounded monotone sequence, it is necessary to make additional assumptions about the space N . To avoid burdening the discussion with these topological questions we restrict ourselves at first to the case $N = \mathbb{R}^m$, i.e. to \mathbb{R}^m -metric spaces, and postpone the extension of our results to general N -metric spaces until Section 5. Here \mathbb{R}^m denotes the real m -dimensional coordinate space with the usual componentwise partial-ordering. Clearly \mathbb{R}^m is a PTL space under any norm topology and if $\|\cdot\|$ denotes an arbitrary norm then the sets $\{x \in X \mid \|\rho(x, x_0)\| < r, r \text{ real}\}$ form a local neighborhood base at x_0 for a Hausdorff topology on X . Under this topology a sequence $\{x_n\} \in X$ converges to $x \in X$ if and only if $\rho(x_n, x) \rightarrow 0$ as $n \rightarrow \infty$. We call X a complete \mathbb{R}^m -metric space if every Cauchy sequence on X has a limit point in X .

\mathbb{R}^m -metric spaces play an important role in applications.

Some simple examples are the following:

1) On $X = \mathbb{R}^m$,

$$\rho(x, y) = (|x_1 - y_1|, \dots, |x_m - y_m|),$$

where $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$;

is an invariant \mathbb{R}^m -metric.

2) Let X be the space of continuous functions on $[0,1]$ and $0 = t_1 < t_2 < \dots < t_{m+1} = 1$. Then $\rho(x,y) = (r_1, \dots, r_m)$,

where

$$r_k = \sup_{t_k \leq t \leq t_{k+1}} |x(t) - y(t)|, \quad k = 1, \dots, m,$$

is an invariant R^m -metric on X .

3) On the space X of $(m-1)$ -times continuously differentiable functions on $[0,1]$, $\rho(x,y) = (r_1, \dots, r_m)$, where

$$r_{k+1} = \sup_{0 \leq t \leq 1} |x^{(k)}(t) - y^{(k)}(t)|, \quad k = 0, 1, \dots, m-1,$$

is an invariant R^m -metric.

Definition 2: A linear operator $P: R^m \rightarrow R^m$ is nonnegative if

$Pa \geq 0$ whenever $a \geq 0$. P is convergent if $\sum_{k=0}^{\infty} P^k a$ converges for all $a \in \mathbb{R}^m$.

Note that P is convergent if and only if its spectral radius is less than unity. If P is convergent then $(I-P)^{-1} = \sum_{k=0}^{\infty} P^k a, a \in \mathbb{R}^m$; this shows in particular that $(I-P)^{-1}$ is nonnegative if P is nonnegative. The converse also holds (Varga [21, p. 83]):

Lemma 1. Let $P: \mathbb{R}^m \rightarrow \mathbb{R}^m$ be a nonnegative linear operator and suppose $(I-P)^{-1}$ exists and is nonnegative. Then P is convergent.

The next lemma reduces for $m = 1$ to a special case of the well-known Toeplitz lemma [17].

Lemma 2. Let $P: \mathbb{R}^m \rightarrow \mathbb{R}^m$ be convergent and set

$$a_n = \sum_{k=0}^n P^{n-k} b_k, \quad n = 0, 1, \dots$$

Then $a_n \rightarrow 0$ if and only if $b_n \rightarrow 0$.

Proof: Since P is convergent, $\limsup \|P^n\|^{1/n} < 1$ in any norm and therefore $\sum_{k=0}^{\infty} \|P^k\| < \infty$.

From

$$\|a_n\| \leq \sum_{k=0}^{n_0} \|P^{n-k} b_k\| + \sum_{k=n_0+1}^n \|P^{n-k} b_k\|$$

it then follows that $a_n \rightarrow 0$ if $b_k \rightarrow 0$. Conversely, since

$$a_{n+1} = Pa_n + b_{n+1}, \quad a_n \rightarrow 0 \text{ implies } b_n \rightarrow 0.$$

Definition 3: Let X be an R^m -metric space and $P: R^m \rightarrow R^m$ a nonnegative convergent linear operator. Then a mapping $F: D \subset X \rightarrow X$ is called a P-contraction on D if

$$\rho(Fx, Fy) \leq P\rho(x, y), \quad x, y \in D$$

A general theorem of Schröder [16] assures the validity of the contraction principle: If $F: X \rightarrow X$ is a P-contraction on a complete R^m -metric space X , then F has a unique fixed-point $x^* = Fx^*$ in X and the sequence $x_{n+1} = Fx_n, n = 0, 1, \dots$ converges to x^* for any x_0 .

We shall be interested in approximate P-contractions that are defined only on a subdomain of X . The following lemma generalizes a result of Collatz (see [3]).

Lemma 3: Let X be an R^m -metric space and $P: R^m \rightarrow R^m$ a non-negative convergent linear operator. Suppose that $F: D \subset X \rightarrow X$ satisfies

$$\rho(Fx, Fy) \leq P\rho(x, y) + \gamma, \quad x, y \in D$$

for some fixed $\gamma \in R^m$ and that there exist $y_0, y_1 \in D$ such that

$$(2.3) \quad S \equiv \left\{ x \mid \rho(x, y_1) \leq \gamma \equiv (I-P)^{-1} [P\rho(y_1, y_0) + \gamma + \delta] \right\} \subset D$$

where $\delta \geq \rho(Fy_0, y_1)$. Then $FS \subset S$.

Proof: If $x \in S$, then

$$\begin{aligned} \rho(Fx, Y_1) &\leq \rho(Fx, FY_0) + \rho(FY_0, Y_1) \leq P\rho(x, Y_0) + \gamma + \delta \\ &\leq P\rho(x, Y_1) + [P\rho(Y_1, Y_0) + \gamma + \delta] \leq P\gamma + (I-P)\gamma = \gamma. \end{aligned}$$

3. Approximate Iterations

The following theorem, which generalizes a result of Ostrowski [10], will be basic to the subsequent discussion.

Theorem 1: Let X be a complete R^m -metric, $F:D \subset X \rightarrow X$ a P -contraction on D and $S \subset D$ a closed subset with the property that $FS \subset S$. (Hence, by the contraction principle the sequence $x_{n+1} = Fx_n$, $n = 0, 1, \dots$, starting from any $x_0 \in S$ converges to the unique fixed-point x^* of F in S .) Let y_0, y_1, \dots be an arbitrary sequence in D and set $\epsilon_n = \rho(Fy_n, y_{n+1})$, $n = 0, 1, \dots$. Then the following estimates hold:

$$(3.1) \quad \rho(y_{n+1}, x^*) \leq (I-P)^{-1} [P\rho(y_{n+1}, y_n) + \epsilon_n],$$

and

$$(3.2) \quad \rho(y_{n+1}, x^*) \leq \rho(x_{n+1}, x^*) + \sum_{k=0}^n P^{n-k} \epsilon_k + P^{n+1} \rho(x_0, y_0).$$

Moreover,

$$(3.3) \quad y_n \rightarrow x^* \text{ if and only if } \epsilon_n \rightarrow 0, n \rightarrow \infty.$$

Proof: The estimate (3.1) follows from

$$\begin{aligned} \rho(y_{n+1}, x^*) &\leq \rho(y_{n+1}, FY_n) + \rho(FY_n, FY_{n+1}) + \rho(FY_{n+1}, Fx^*) \\ &\leq \epsilon_n + P\rho(y_n, y_{n+1}) + P\rho(y_{n+1}, x^*) \end{aligned}$$

and the nonnegativity of $(I-P)^{-1}$ while (3.2) is obtained from

$$\begin{aligned} \rho(x_{n+1}, y_{n+1}) &\leq \rho(Fx_n, Fy_n) + \rho(Fy_n, y_{n+1}) \\ &\leq P\rho(x_n, y_n) + \varepsilon_n \leq \dots \leq \sum_{k=0}^n P^{n-k} \varepsilon_k + P^{n+1} \rho(x_0, y_0) \end{aligned}$$

together with $\rho(y_{n+1}, x^*) \leq \rho(y_{n+1}, x_{n+1}) + \rho(x_{n+1}, x^*)$.

If $\varepsilon_n \rightarrow 0$ then the convergence of y_n to x^* follows directly from (3.2) and Lemma 2. Conversely, if $y_n \rightarrow x^*$ then

$$\begin{aligned} 0 \leq \varepsilon_n = \rho(Fy_n, y_{n+1}) &\leq \rho(Fy_n, Fx^*) + \rho(x^*, y_{n+1}) \\ &\leq P\rho(y_n, x^*) + \rho(y_{n+1}, x^*) \end{aligned}$$

and $\varepsilon_n \rightarrow 0$.

Note that no assumptions were made about the sequence $\{y_n\}$ except that $\{y_n\} \subset D$. In particular, the y_n do not need to lie in S nor do the ε_n need to be small. For the estimates to be useful, we shall of course interpret the sequence $\{y_n\}$ as an approximation to the exact sequence $\{x_n\}$. In the special case $x_n = y_n$, i.e., $\varepsilon_n = 0$, the estimate (3.1) reduces to that of the contraction principle. Finally, note that (3.1) and (3.2) play different roles: (3.2) relates the exact sequence $\{x_n\}$ to the sequence $\{y_n\}$ and is useful for such theoretical purposes as proving (3.3). On the other hand, (3.1) is a computable estimate which may be used, for example, to terminate a computation.

The approximate sequence $\{y_n\}$ may be generated in a variety of ways. A rather general process, considered by Ehrmann [4], is

$$(3.4) \quad y_{n+1} = F_n y_n, \quad n = 0, 1, \dots$$

where $F_n : D \subset X \rightarrow X$. A simple consequence of Theorem 1 in this setting is:

Corollary 1.1: Let $F : X \rightarrow X$ be a P-contraction on all of X and suppose the mappings $F_n : X \rightarrow X$, $n = 0, 1, \dots$, satisfy

$$(3.5) \quad \rho(F_n x, Fx) \rightarrow 0, \text{ uniformly for } x \in X.$$

Then if $\{y_n\}$ is generated by (3.4) we have $y_n \rightarrow x^*$, where x^* is the unique fixed-point of F in X .

The proof is immediate since (3.5) implies that $\xi_n = \rho(F_n y_n, F y_n) \rightarrow 0$. Note that (3.5) cannot in general be weakened to pointwise convergence as the following simple example shows: $X = \mathbb{R}^1$, $m = 1$, $F \equiv 0$, $F_n x = \frac{1}{n+1} e^x$.

Another corollary strengthens a result of Schmidt [14, Theorem 3]. The proof follows from Lemma 3, the contraction principle and Theorem 1.

Corollary 1.2: Let $F : D \subset X \rightarrow X$ be a P-contraction on D and suppose $F_n : D \subset X \rightarrow D$, $n = 0, 1, \dots$, map D into itself. Assume that for some $y_0 \in D$, (2.3) holds with $y_1 = F_0 y_0$, $\delta \geq \rho(F y_0, F_0 y_0)$ and $\gamma = 0$. Then F has a unique fixed-point $x^* \in S$ and for

the sequence $\{y_n\}$ generated by (3.4), (3.1) and (3.3) hold.

Schmidt did not assume that the F_n map D into itself and instead required that (2.3) hold with $\gamma = 0$ and $\delta \geq 2\rho(y_{n+1}, Fy_n)$, $n = 0, 1, \dots$. However, this is an incorrect assumption since δ cannot be known until it is assured that the sequence $\{y_n\}$ exists. But this is just one of the reasons for a condition like (2.3).

With more stringent conditions on the F_n stronger results may be obtained. A natural requirement is that the F_n themselves be P-contractions, as assumed by Ehrmann, or approximate P-contractions.

Theorem 2: Let X be a complete R^m -metric space and $F: D \subset X \rightarrow X$ a P-contraction on D . Let the operators $F_n: D \subset X \rightarrow X$, $n = 0, 1, \dots$ satisfy

$$(3.6) \quad \rho(F_n x, F_n y) \leq P\rho(x, y) + 2\gamma, \quad x, y \in D$$

for some fixed $\gamma \in R^m$, and suppose there exists a $y_0 \in D$ for which

$$(3.7) \quad S \equiv \{x \mid \rho(x, F_0 y_0) \leq (I-P)^{-1} [P\rho(F_0 y_0, y_0) + 2\gamma + \delta]\} \subset D$$

where

$$(3.8) \quad \delta \geq \rho(F_n y_0, F_0 y_0), \quad n = 0, 1, \dots,$$

and

$$(3.9) \quad 2\gamma + \delta \geq \rho(Fy_0, F_0 y_0).$$

Then

- a) the sequences $y_{n+1} = F_n y_n$, $x_{n+1} = Fx_n$, $n = 0, 1, \dots$,
 $(x_0 = y_0)$ are well-defined, $x_n \rightarrow x^*$, where x^* is the unique
 fixed point of F in S , and the error estimates (3.1) and
 (3.2) hold with $\epsilon_n = \rho(F_n y_n, Fy_n)$.

Moreover, if $\gamma = 0$, each F_n has a unique fixed point z_n in S
 and the following four statements are equivalent:

- b) $y_n \rightarrow x^*$, c) $z_n \rightarrow x^*$,
 d) $\epsilon_n \rightarrow 0$, e) $\rho(F_n x^*, Fx^*) \rightarrow 0$.

Proof: Using (3.9), it follows from Lemma 3 that $F_n S \subset S$ and
 the contraction principle assures the existence of x^* and the
 convergence of x_n to x^* . Similarly, using (3.8), Lemma 3
 applied to each of the F_n shows that $F_n S \subset S$ and hence the
 sequence y_n is well-defined. The error estimates of Theorem 1
 now apply.

If $\gamma = 0$, then each F_n is itself a P-contraction and since
 $F_n S \subset S$ there exist unique $z_n \in S$ such that $z_n = F_n z_n$,
 $n = 0, 1, \dots$. Now

$$\begin{aligned} \rho(y_{n+1}, x^*) &\leq \rho(F_n y_n, F_n x^*) + \rho(F_n x^*, Fx^*) \leq P \rho(y_n, x^*) + \rho(F_n x^*, Fx^*) \leq \dots \\ &\leq \sum_{k=0}^n P^{n-k} \rho(F_k x^*, Fx^*) + P^{n+1} \rho(y_0, x^*) \end{aligned}$$

so that by Lemma 2, (e) implies (b). From

$$\rho(z_n, x^*) \leq \rho(F_n z_n, F_n x^*) + \rho(F_n x^*, F_n y_n) + \rho(F_n y_n, Fx^*)$$

we obtain

$$\rho(z_n, x^*) \leq (I-P)^{-1} [P\rho(x^*, y_n) + \rho(y_{n+1}, x^*)]$$

so that (b) implies (c). Next

$$\rho(F_n x^*, x^*) \leq \rho(F_n x^*, F_n z_n) + \rho(F_n z_n, Fx^*) = P\rho(x^*, z_n) + \rho(z_n, x^*)$$

and (c) implies (e). Finally, (b) and (d) are equivalent by Theorem 1 and the proof is complete.

Theorem 2 contains as corollaries several known results.

The following generalizes a theorem of Urabe [18].

Corollary 2.1: Let F be as in Theorem 2 and suppose the mappings

$F_n \equiv F_0$, $n = 0, 1, \dots$ where $F_0: D \subset X \rightarrow X$ is such that

$\rho(F_0 x, Fx) \leq \gamma$, $x \in D$. Assume that (3.7) holds with $\delta = 0$.

Then (a) of Theorem 2 is valid.

We note that here the estimates (3.1) and (3.2) imply

$$\rho(y_{n+1}, x^*) \leq (I-P)^{-1} [\rho(y_n, y_{n-1}) + \gamma],$$

and

$$\rho(y_{n+1}, x^*) \leq \rho(x_{n+1}, x^*) + (I-P)^{-1} \gamma$$

since $\gamma \geq \rho(F_0 y_n, Fy_n) = \epsilon_n$. The proof of Corollary 1 is immediate by noting that (3.6) follows from

$$\rho(F_0 x, F_0 y) \leq \rho(F_0 x, Fx) + \rho(Fx, Fy) + \rho(Fy, F_0 y)$$

while (3.8) and (3.9) are automatically satisfied.

Next we consider the main result of Ehrmann [4] which, in the case of R^m -metric spaces, is given by the following

Corollary 2.2: Let X be a complete R^m -metric space and $F_n : D \subset X \rightarrow X$ given P -contractions on D such that $F : D \subset X \rightarrow X$, (3.5) holds. Suppose there is $y_0 \in D$ such that (3.7) is true with $\gamma = 0$ and δ satisfying (3.8). Then (a) and (b) of Theorem 2 are valid.

Proof: From

$$\rho(Fx, Fy) \leq \rho(Fx, F_n x) + \rho(F_n x, F_n y) + \rho(F_n y, Fy),$$

together with (3.5), it follows that F itself is a P -contraction on D . Similarly, from

$$\rho(Fy_0, F_0 y_0) \leq \rho(Fy_0, F_n y_0) + \rho(F_n y_0, F_0 y_0)$$

the validity of (3.9) is obtained. Hence all the conditions of Theorem 2 are satisfied. In particular, (b) follows from (d) since $\epsilon_n \rightarrow 0$ by (3.5).

We note that the uniform convergence in (3.5) is required only to guarantee that $\epsilon_n \rightarrow 0$ and the equivalence of (b) and (e) in Theorem 2 yields a stronger result.

Corollary 2.3: The conclusions of Corollary 2.2 remain valid if (3.5) holds only pointwise.

4. Implicit Iterations

In Theorem 2, we considered not only the sequence $y_{n+1} = F_n y_n$ but also the sequence $z_n = F_n z_n$ of fixed-points of F_n . Implicit iterations of this kind have been considered in some detail by Warga [22] and, more recently, by Schmidt [14]. The following theorem combines and strengthens, for R^m -metric spaces, Schmidt's Theorems 1 and 2. We do not assume the existence of a fixed-point (Schmidt's Theorem 1) nor do we assume that $\rho(F_{n+1} y_n, F y_n) = 0$ (Schmidt's Theorem 2).

Theorem 3: Let X be a linear space which is a complete R^m -metric space under some invariant R^m -metric and let $F: D \subset X \rightarrow X$. Suppose that $F_n: D \subset X \rightarrow X$, $n = 0, 1, \dots$ are Q -contractions on D which possess (unique) fixed-points $y_n \in D$ and have the property that for all $n \geq 0$

$$(4.1) \quad \rho(Fx - F_n x, Fy - F_n y) \leq R\rho(x, y) \quad , \quad x, y \in D,$$

where R is a nonnegative linear operator and $P = (I - Q)^{-1}R$ is convergent. Assume further that y_0 and y_1 have the property that

$$S \equiv \{x \mid \rho(x, y_1) \leq \eta \equiv (I - P)^{-1} [P \rho(y_0, y_1) + (I - Q)^{-1} \delta_0]\} \subset D$$

where $\delta_0 \equiv \rho(F_1 y_0, F y_0)$. Then F has a unique fixed-point $x^* \in D$ and the following error estimate holds with $\delta_n = \rho(F_{n+1} y_n, F y_n)$, $n = 1, 2, \dots$:

$$(4.2) \quad \rho(x^*, y_{n+1}) \leq (I - P)^{-1} [P \rho(y_{n+1}, y_n) + (I - Q)^{-1} \delta_n].$$

Moreover, the following four statements are equivalent:

- (a) $\epsilon_n \equiv \rho(y_{n+1}, Fy_n) \rightarrow 0$ (b) $\delta_n \rightarrow 0$
 (c) $y_n \rightarrow x^*$ (d) $\rho(F_n x^*, Fx^*) \rightarrow 0$

Proof: By (2.1) and (4.1) we have for all $x, y \in D$

$$\rho(Fx, Fy) \leq \rho(Fx - F_n x, Fy - F_n y) + \rho(F_n x, F_n y) \leq (Q+R) \rho(x, y).$$

With $T = Q + R$, $I-T = (I-Q)(I-P)$ and since P and Q are convergent and nonnegative,

$$(4.3) \quad (I-T)^{-1} = (I-P)^{-1} (I-Q)^{-1}$$

exists and is nonnegative. Therefore Lemma 1 shows that T is convergent and hence F is a T -contraction on D .

Now let $S' = \{x \mid \rho(x, Fy_1) \leq (I-T)^{-1} T \rho(Fy_1, y_1)\}$. Then for $x \in S'$,

$$\begin{aligned} \rho(x, y_1) &\leq \rho(x, Fy_1) + \rho(Fy_1, y_1) \\ &\leq (I-T)^{-1} T \rho(Fy_1, y_1) + \rho(Fy_1, y_1) = (I-T)^{-1} \rho(Fy_1, y_1) \\ &\leq (I-T)^{-1} [\rho(Fy_1 - F_1 y_1, Fy_0 - F_1 y_0) + \rho(Fy_0, F_1 y_0)] \\ &\leq (I-T)^{-1} [R\rho(y_1, y_0) + \delta_0] = (I-P)^{-1} [P\rho(y_1, y_0) + (I-Q)^{-1} \delta_0] = \eta \end{aligned}$$

That is, $S' \subset S$. But then, by Lemma 3, $FS' \subset S'$ and, by the contraction principle, F has a fixed point $x^* \in S'$.

The equivalence of (a) and (c) follows directly from Theorem 1. Using this equivalence and

$$\delta_n \equiv \rho(F_{n+1}Y_n, F_{n+1}Y_{n+1}) + \rho(y_{n+1}, Fy_n) \equiv Q\rho(y_n, Y_{n+1}) + \epsilon_n$$

we see that (c) implies (b). Now

$$(4.4) \quad \rho(y_{n+1}, x^*) \equiv \rho(F_{n+1}Y_{n+1}, F_{n+1}x^*) + \rho(Fy_n - F_{n+1}Y_n, Fx^* - F_{n+1}x^*) \\ + \rho(Fy_n, F_{n+1}Y_n) \equiv Q\rho(y_{n+1}, x^*) + R\rho(y_n, x^*) + \delta_n$$

or

$$\rho(y_{n+1}, x^*) \equiv P\rho(y_n, x^*) + (I-Q)^{-1}\delta_n \\ \leq \dots \leq \sum_{k=0}^n P^{n-k} (I-Q)^{-1}\delta_k + P^{n+1} \rho(y_0, x^*)$$

which by Lemma 2 shows that (b) implies (c). Finally the equivalence of (c) and (d) is proved as in Theorem 2. The error estimate follows directly from (4.4) together with (4.3) since

$$\rho(x^*, Y_{n+1}) \equiv Q\rho(y_{n+1}, x^*) + R\rho(y_n, x^*) + \delta_n \\ \equiv (Q+R) \rho(y_{n+1}, x^*) + R\rho(y_{n+1}, Y_n) + \delta_n$$

or

$$\rho(y_{n+1}, x^*) \equiv (I-T)^{-1} [R\rho(y_{n+1}, Y_n) + \delta_n].$$

This completes the proof.

The condition (4.1) on the difference operator $F_n - F$ can be replaced by a stronger condition on F itself. Assume that instead of (4.1) we require that F is an R_0 -contraction and $P = (I-Q)^{-1}(R_0+Q)$ is convergent. Then

$$\rho(Fx - F_n x, Fy - F_n y) \equiv \rho(Fx, Fy) + \rho(F_n x, F_n y) \equiv (R_0 + Q) \rho(x, y)$$

so that with $R = R_0 + Q$, (4.1) is satisfied. This represents an extension of the result of Warga [22].

In a somewhat different way we can again remove the condition (4.1) and also the requirement of an invariant metric; in addition, the following theorem appears to be more suitable for applications.

Theorem 4: Let X be a complete R^m -metric space and $F:D \subset X \rightarrow X$, $F_n : D \times D \subset X \times X \rightarrow X$ ($n = 1, 2, \dots$) given operators such that

$$(4.5) \quad \rho(F_n(x, z), F_n(y, z)) \leq Q\rho(x, y), \quad x, y, z \in D,$$

$$(4.6) \quad \rho(F_n(z, x), F_n(z, y)) \leq R\rho(x, y), \quad x, y, z \in D,$$

where Q is nonnegative and convergent, R is nonnegative and $P = (I-Q)^{-1}R$ is convergent. Let $\delta_n(x) = \rho(F_n(x, x), Fx)$, $x \in D$, and assume that

$$(4.7) \quad \lim_{n \rightarrow \infty} \delta_n(x) = 0, \text{ pointwise for } x \in D.$$

Suppose further that there exists a $y_0 \in D$ such that the equation $y = F_1(y, y_0)$ has a solution $y_1 \in D$ and

$$S \equiv \{x \mid \rho(x, y_1) \leq (I-P)^{-1} [P\rho(y_1, y_0) + (I-Q)^{-1}\delta]\} \subset D$$

where $\delta \geq \delta_n(y_1)$ ($n = 1, 2, \dots$). Then the equations

$y = F_n(y, y_{n-1})$ ($n = 1, 2, \dots$) have unique solutions y_n in S

and $y_n \rightarrow x^*$ where $x^* \in S$ is the unique fixed-point of F in D .

Proof: For $x, y \in D$ and $n \geq 1$ we have

$$\begin{aligned} \rho(Fx, Fy) &\cong \rho(Fx, F_n(x, x)) + \rho(F_n(x, x), F_n(x, y)) + \rho(F_n(x, y), F_n(y, y)) \\ &\quad + \rho(F_n(y, y), Fy) \cong (Q+R) \rho(x, y) + \delta_n(x) + \delta_n(y) \end{aligned}$$

and, by (4.7), $\rho(Fx, Fy) \cong T\rho(x, y)$ where $T = Q+R$. Hence it follows as in Theorem 3 that F is a T -contraction on D .

Let $S' = \{x \mid \rho(x, Fy_1) \cong (I-T)^{-1}T\rho(Fy_1, y_1)\}$. Then $S' \subset S$ since for $x \in S'$ an argument similar to that in Theorem 3 shows

$$\begin{aligned} \rho(x, y_1) &\cong (I-T)^{-1}[\rho(Fy_1, F_1(y_1, y_1)) + \rho(F_1(y_1, y_1), F_1(y_1, y_0))] \\ &\cong (I-T)^{-1}[R\rho(y_1, y_0) + \delta_1(y_1)] \cong (I-P)^{-1}[P\rho(y_1, y_0) + (I-Q)^{-1}\delta]. \end{aligned}$$

Thus, by Lemma 3, $FS' \subset S'$ and F has a fixed-point $x^* \in S'$.

Next let $S'' = \{x \mid \rho(x, y_1) \cong (I-Q)^{-1}\rho(F_n(y_1, \bar{y}), y_1)\}$ for some fixed n and arbitrary fixed $\bar{y} \in S$. Then with

$$\eta = (I-P)^{-1}[P\rho(y_1, y_0) + (I-Q)^{-1}\delta],$$

which implies

$$P\eta + P\rho(y_1, y_0) = \eta - (I-Q)^{-1}\delta \cong \eta,$$

we have

$$\begin{aligned} (I-Q)^{-1}\rho(F_n(y_1, \bar{y}), y_1) &= (I-Q)^{-1}\rho(F_n(y_1, \bar{y}), F_n(y_1, y_0)) \\ &\cong (I-Q)^{-1}R\rho(\bar{y}, y_0) = P[\rho(\bar{y}, y_1) + \rho(y_1, y_0)] \\ &\cong P\eta + P\rho(y_1, y_0) \cong \eta, \end{aligned}$$

so that $S'' \subset S$. Hence, by Lemma 3, $F_n(\cdot, \bar{y})$ maps S' into itself

and therefore the y_n exist and are unique in D .

Finally,

$$\begin{aligned} \rho(y_n, x^*) &\leq \rho(F_n(y_n, y_{n-1}), F_n(x^*, y_{n-1})) + \rho(F_n(x^*, y_{n-1}), F_n(x^*, x^*)) \\ &\quad + \rho(F_n(x^*, x^*), Fx^*) \\ &\leq Q\rho(y_n, x^*) + R\rho(y_{n-1}, x^*) + \delta_n(x^*) \end{aligned}$$

or

$$\rho(y_n, x^*) \leq P\rho(y_{n-1}, x^*) + (I-Q)^{-1}\delta_n(x^*)$$

which by Lemma 2 and (4.7) assures that $y_n \rightarrow x^*$. This completes the proof.

We note that if the operators F_n do not depend on the first variable, i.e., if $F_n(x, y) \equiv F_n y$, then we are considering the explicit process $y_{n+1} = F_n y_n$. Here $Q \equiv 0$, $P \equiv R$ and Theorem 4 reduces essentially to Corollary 3 of Theorem 2 although the two constants δ are defined in slightly different ways.

Similarly, if the F_n do not depend on the second variable, Theorem 4 gives a convergence result for the implicit process $y_n = F_n y_n$, $n = 0, 1, \dots$.

As an example of the application of Theorem 4 let $X = R^m$ and let

$$Fx \equiv (f_i(x_1, \dots, x_m) \mid i = 1, \dots, m) = 0$$

be a nonlinear system of equations. Set

$$F_n(x, y) \equiv F_0(x, y) \equiv (f_i(x_1, \dots, x_i, y_{i+1}, \dots, y_m) \mid i = 1, \dots, m)$$

Then $F_0(x, x) = F(x)$ and we have the implicit Gauss-Seidel process studied by Bers [1] and Schechter [13]:

$$x_i^{(k+1)} = f_i(x_1^{(k+1)}, \dots, x_i^{(k+1)}, x_{i+1}^{(k)}, \dots, x_m^{(k)}), \quad (i = 1, \dots, m)$$

Conditions (4.5) and (4.6):

$$|f_i(x_1, \dots, x_i, z_{i+1}, \dots, z_m) - f_i(y_1, \dots, y_i, z_{i+1}, \dots, z_m)| \leq \sum_{j=1}^i q_{ij} |x_j - y_j|$$

and

$$|f_i(z_1, \dots, z_i, x_{i+1}, \dots, x_m) - f_i(z_1, \dots, z_i, y_{i+1}, \dots, y_m)| \leq \sum_{j=i+1}^m r_{ij} |x_j - y_j|$$

are then satisfied if for the triangular matrices $Q = (q_{ij})$ and $R = (r_{ij})$ we have $q_{ii} < 1$, $i = 1, \dots, m$ and $(I-Q)^{-1}R$ is convergent. This latter condition is satisfied if $Q+R$ is convergent.

Other examples are provided by linear decompositions of F . Let X be a linear R^m -metric space and let F, G_n, H_n be nonlinear operators on X such that $Fx = F_n x + H_n x$, $n = 0, 1, \dots$. Then if we define $F_n(x, y) \equiv G_n x + H_n y$ we obtain the process

$$y_{n+1} = G_n y_{n+1} + H_n y_n, \quad n = 0, 1, \dots$$

Theorem 4 then requires that all G_n are Q -contractions and $\rho(H_n x, H_n y) \leq R\rho(x, y)$, $n = 0, 1, \dots$ with a nonnegative linear R such that $(I-Q)^{-1}R$ is convergent.

5. Extension to N-metric spaces

In order to extend the results of the previous sections to more general N-metric spaces it is necessary to place additional conditions on the PTL space N. For details about the following definitions we refer to the literature on partially-ordered topological linear spaces e.g. Birkhoff [2], Nachbin [8], Namioka [9] and Schaefer [12]. (See also Vandergraft [20] and Rheinboldt and Vandergraft [11].)

Definition 4: Let N be a PTL space with positive cone C and topology T. Then

- a) N is called solid if C has an interior point;
- b) N is called regular if any order-bounded monotone increasing sequence has a limit i.e. if whenever $a_n \leq a_{n+1} \leq a, n = 0, 1, \dots, \{a_n\}$ converges to some element in N;
- c) N is called normal if for any neighborhood base U of the origin, there exists a constant $\alpha > 0$ such that for any $a \geq 0$ in N, and any neighborhood $U \in U$ we have $\{b \mid 0 \leq b \leq a\} \subset \alpha U$.

As examples, we note that R^m is regular, solid and normal. $C[0,1]$ and $L^p[0,1]$ ($1 \leq p \leq \infty$) are normal under their usual orderings and topologies. Furthermore, $C[0,1]$ and $L^\infty[0,1]$ are

solid but not regular while $L^p[0,1]$ for $1 \leq p \leq \infty$ are regular but not solid.

For PTL spaces in which T is a norm topology, it can be shown that the space is normal if and only if there exists a constant $\alpha > 0$ such that $0 \leq a \leq b$ implies $\|a\| \leq \alpha \|b\|$. We can then introduce an equivalent monotone norm $\|\cdot\|'$, i.e. $0 \leq a \leq b$ implies $\|a\|' \leq \|b\|'$.

In general, a sequence $\{b_n\}$ in a partially ordered linear space N is called relatively uniformly convergent to zero if there exists a real sequence $t_1 \geq t_2 \geq \dots \geq 0$ with $t_i \rightarrow 0$ and an element $b \geq 0$ in N such that

$$(5.1) \quad -t_n b \leq b_n \leq t_n b, \quad n = 0, 1, \dots$$

For normal PTL spaces it may be shown that

$$(5.2) \quad 0 \leq a_n \leq b_n, \quad n = 0, 1, \dots \text{ where } b_n \rightarrow 0, \text{ implies } a_n \rightarrow 0$$

and thus in a normal PTL space relative uniform convergence implies topological convergence. For solid PTL spaces the converse holds. Therefore, in a normal and solid PTL space topological convergence is equivalent with relative uniform convergence.

All of the sequential convergence conditions used by Collatz [3] and Schröder [16] are satisfied in a normal PTL space N . The conditions of Ehrmann [4] require that N be solid

as well, while Schmidt [14] assumed that convergence is equivalent to relative uniform convergence, a condition which is satisfied if N is both solid and normal.

Let X be an N -metric space where N is normal. If U is a neighborhood base of the origin in N , then for each $x_0 \in X$ the sets $\{x \in X \mid \rho(x, x_0) \in U \in U\}$ form a local neighborhood base of x_0 for a uniform topology on X . Under this topology, convergence of a sequence $\{x_n\} \subset X$ to $x \in X$ again means that $\rho(x_n, x) \rightarrow 0$; we call X complete if any Cauchy sequence has a limit in X .

As before, a continuous linear operator $P: N \rightarrow N$ is called nonnegative if $Pa \geq 0$ whenever $a \geq 0$ and convergent if $\sum_{k=0}^{\infty} P^k a$ exists for all $a \in N$. It is well-known that for convergent P , $(I-P)^{-1}$ exists and $(I-P)^{-1}a = \sum_{k=0}^{\infty} P^k a$ for all $a \in N$. Clearly, $(I-P)^{-1}$ is nonnegative if P is nonnegative.

On an N -metric space X a mapping $F: D \subset X \rightarrow X$ will again be called a P -contraction on D if there exists a continuous nonnegative convergent linear operator $P: N \rightarrow N$ such that $\rho(Fx, Fy) \leq P\rho(x, y)$, $x, y \in D$. If N is a normal PTL space the contraction principle again follows from the general result of Schröder [16].

It is easily checked that on any N -metric space X for which N is normal, Lemma 3, Theorem 1 with the exception of (3.3), and part (a) of Theorem 2 remain valid and the proofs hold word

for word. In order to conclude (3.3) and hence all of the remaining results, we need Lemma 2. In addition, Theorems 3 and 4 have used Lemma 1 while Theorems 2, 3 and 4 have used the continuity of $(I-P)^{-1}$.

For the continuity of $(I-P)^{-1}$ two general results can be cited: If N is a solid and normal PTL space then Namioka [9] has shown that every nonnegative linear operator is continuous. On the other hand, if N is semi-metrizable and topologically complete then the Banach-Steinhaus Theorem assures that $(I-P)^{-1}$, as the limit of continuous linear operators, is again continuous.

Next we consider the generalization of Lemma 2. If

$$(5.1) \quad a_n = \sum_{k=0}^n P^{n-k} b_k,$$

then whenever $P:N \rightarrow N$ is continuous it follows as before that $a_n \rightarrow 0$ implies $b_n \rightarrow 0$. For the converse one general result is the following: If N is a normed linear space and $P:N \rightarrow N$ a linear operator with spectral radius $\sigma(P) \equiv \limsup \| P^n \|^{1/n} < 1$, then $\sum_0^\infty \| P^n \| < \infty$ and the previous proof of Lemma 2 holds. But if we only know that P is convergent then we may have $\sigma(P) = 1$ and the proof breaks down. A different approach is the following:

Lemma 4: Let N be a normal PTL-space and $P:N \rightarrow N$ a nonnegative convergent continuous linear operator. Then if $b_k \rightarrow 0$ relatively uniformly, $a_n \rightarrow 0$ in the topology of N .

Proof: For some fixed m , let $c_n = P^{n-m} \sum_{k=0}^m P^{m-k} b_k$ and $d_n = \sum_{k=m+1}^n P^{n-k} b_k$, $n \geq m$; then $a_n = c_n + d_n$. Since $b_k \rightarrow 0$ relatively uniformly, there exists $b \geq 0$ in N and a real sequence $t_k \geq t_{k+1} \geq \dots \geq 0$, $t_k \rightarrow 0$ such that (5.1) holds. Hence, since P is nonnegative and convergent, $-t_{m+1}(I-P)^{-1}b \leq d_n \leq t_{m+1}(I-P)^{-1}b$, $n \geq m$, and given any neighborhood U of the origin, the normality of N shows that we can choose m such that $d_n \in \frac{1}{2}U$, $n \geq m$. But since P is convergent $c_n \rightarrow 0$ for any fixed m and hence $a_n \in U$, $n \geq m$, i.e. $a_n \rightarrow 0$.

Since in normal, solid PTL spaces, sequential topological convergence is equivalent with relative uniform convergence it follows from Lemma 4 that in such spaces Lemma 2 holds. Hence in N -metric spaces for which N is normal and solid, Theorems 1 and 2 remain valid, and the proofs are the same.

Theorems 3 and 4 also required Lemma 1 in order to conclude that $Q + R$ was a convergent operator. One generalization of Lemma 1 is the following:

Lemma 4: Let N be a regular PTL-space and assume that the positive cone C is reproducing, i.e., $C - C = N$. Let $P: N \rightarrow N$ be a nonnegative linear operator such that $(I-P)^{-1}$ exists and is nonnegative. Then P is convergent.

Proof: For any $b \geq 0$,

$$c_n = \sum_{k=0}^n P^k b \leq (I-P)^{-1} b.$$

Hence the monotonically increasing sequence $\{c_n\}$ is order-bounded and has a limit. For arbitrary b there exist $b_1, b_2 \in C$ such that $b = b_1 - b_2$. But then $\sum P^k b_1$ and $\sum P^k b_2$ converge and hence also $\sum P^k b$.

By itself this result is satisfactory but taken together with the earlier assumptions of normality and solidness, the class of spaces for which Theorems 3 and 4 apply is considerably restricted. Of course, it is possible to add the assumption, or to ascertain otherwise, that both the operators $Q + R$ and $P = (I - Q)^{-1}R$ are convergent; then Theorems 3 and 4 will again be valid if N is only solid and normal.

In summary, we have given some sufficient conditions that Lemma 1 and 2 and hence all our results of Sections 3 and 4 remain valid, although these conditions are rather stringent. An open question remains as to what other conditions are possible and, indeed, what are necessary conditions.

6. References

- [1] L. Bers, On mildly nonlinear partial difference equations of elliptic type, J. Res. Nat. Bur. Standards, 51, 229-236 (1953).
- [2] G. Birkhoff, Lattice Theory, Am. Math. Soc. Colloquium Publ. V. XXV, Am. Math. Soc., 1948.
- [3] L. Collatz, Funktionalanalysis und Numerische Mathematik, Springer Verlag, Germany, 1964.
- [4] H. Ehrmann, Iterationsverfahren mit veränderlichen Operatoren, Arch. Rat. Mech. Anal. 4, 45-64 (1959).
- [5] G. Gardner, Numerical Errors in Iterative Processes, Techn. Report, Division of Applied Mathematics, Brown University, August 1965.
- [6] L. Kantorovich, The Method of Successive Approximations for Functional Equations, Acta. Math. 71, 63-97 (1939).
- [7] L. Kantorovich, B. Vulich, A. Pinsker, Functional Analysis in Partially Ordered Spaces, Gostechisdat, Moscow, 1950.
- [8] L. Nachbin, Topology and Order, Van Nostrand Math. Studies No. 4, Van Nostrand, Princeton, 1965.
- [9] I. Namioka, Partially Ordered Linear Topological Spaces, Mem. Am. Math. Soc. No. 24, Am. Math. Soc., 1957.
- [10] A. Ostrowski, The rounding-off stability of iterations, Basel Math. Notes, BMN-12, August 20, 1964.

- [11] W. Rheinboldt, J. Vandergraft, Partially-Ordered Topological Linear Spaces in Numerical Analysis (in preparation).
- [12] H. Schaefer, Topological Vector Spaces, MacMillan, New York, 1966.
- [13] S. Schechter, Iteration Methods for Nonlinear Problems, Trans. Amer. Math. Soc. 104, 179-189 (1962).
- [14] J. Schmidt, Konvergenzuntersuchungen und Fehlerabschätzungen für ein verallgemeinertes Iterationsverfahren, Arch. Rat. Mech. Anal. 6, 261-276 (1960).
- [15] J. Schmidt, H. Schönheinz, Fehlerschranken zum Differenzenverfahren unter ausschliesslicher Benutzung verfügbarer Grössen, Arch. Rat. Mech. Anal. 10, 311-322 (1962).
- [16] J. Schröder, Das Iterationsverfahren bei verallgemeinertem Abstandsbegriff, Math. Z. 66, 111-116 (1956).
- [17] O. Toeplitz, Über lineare Mittelbildungen, Prace matem. - fizyczne 22, 113-119 (1911).
- [18] M. Urabe, Convergence of Numerical Iteration in Solution of Equations, Journal of Sci., Hiroshima Univ., A 19, 479-489 (1956).
- [19] M. Urabe, Error Estimation in Numerical Solution of Equations by Iteration Process, J. Sci., Hiroshima Univ. Ser. A-I 26, 77-91 (1962).

- [20] J. Vandergraft, On Newton's Method for Convex Operators on Partially Ordered Topological Linear Spaces, Univ. of Maryland, Comp. Sci. Center, Techn. Report TR-66-29, May 1966.
- [21] R. Varga, Matrix Iterative Analysis, Prentice Hall, Englewood Cliffs, N. J., 1962.
- [22] J. Wurga, On a Class of Iterative Procedures for Solving Normal Systems of Ordinary Differential Equations, J. Math. Phys. 31, 223-243 (1952).
- [23] O. I. Zincenko, A Class of Approximate Methods for Solving Operator Equations with Non-Differentiable Operators, Dopovidi Akad. Nauk Ukrain RSR, No. 7, 852-856 (1963).

Computer Science Center
University of Maryland
College Park, Maryland, USA