Pacific Journal of Mathematics

THE CLOSURE OF THE NUMERICAL RANGE CONTAINS THE SPECTRUM

EDUARDO H. ZARANTONELLO

Vol. 22, No. 3 March 1967

PACIFIC JOURNAL OF MATHEMATICS
Vol. 22, No. 3, 1967

THE CLOSURE OF THE NUMERICAL RANGE CONTAINS THE SPECTRUM

EDUARDO H. ZARANTONELLO

The purpose of this paper is to show that the well known theorem in the theory of linear operators in Hilbert space indicated in the title holds for nonlinear operators and to a certain extent for noncontinuous ones, and to provide a constructive method for solving the equations involved.

In different and more precise terms the theorem about to be generalized says:

THEOREM I. *Let T be an everywhere defined linear mapping of a complex Hilbert space $\mathscr{H}$ into itself. Then for any complex number λ at a positive distance $d(\lambda, T)$ from the numerical range of T: $\mathscr{N}(T) = \{(Tx, x), \|x\| = 1\}$, (parentheses indicating scalar product) the equation*

$$\lambda x = Tx - y \tag{1}$$

has a unique solution for every $y \in \mathscr{H}$. The operator $(T - \lambda I)^{-1}$ thus defined is bounded and $\|(T - \lambda I)^{-1}\| \leqq d^{-1}(\lambda, T)$. Moreover, for adequate choices of the averaging factor α depending on T and λ only,

$$(T - \lambda T)^{-1}y = \lim_{n\to\infty} [(1 - \alpha)I + \alpha\lambda^{-1}(T - y)]^n x_0 , \tag{2}$$

where x_0 is any point in $\mathscr{H}$ and $T - y$ the operator mapping x into $Tx - y$.

The theorem having been stated in somewhat more general terms than usual, a proof is needed.

Proof. By definition of $d(\lambda, T)$,

$$|((T - \lambda I)x, x)| = \|x\|^2 \left| \frac{(Tx, x)}{\|x\|^2} - \lambda \right| \geqq d(\lambda, T)\|x\|^2 ,$$

whence it follows by Schwarz' inequality $\|(T - \lambda I)x\| \geqq d(\lambda, T)\|x\|$, proving that $T_\lambda = T - \lambda I$ is a one-one mapping with bounded inverse and $\|T_\lambda^{-1}\| \leqq d^{-1}(\lambda, T)$. By the first inequality above any vector orthogonal to the range $\mathscr{R}(T_\lambda)$ of T_λ must vanish, meaning that $\mathscr{R}(T_\lambda)$ is dense in $\mathscr{H}$. Thus for any $y \in \mathscr{H}$ there is a sequence $\{x_n\}$ such that $T_\lambda x_n \to y$; since T_λ^{-1} is bounded $\{x_n\}$ converges to some element x. Setting $x_n^* = x_n - x$ and $y^* = y - T_\lambda x$, one obtains from the

same inequality and for any complex number σ

$$\begin{aligned} d(\lambda, T)\,\|y^*\|^2 &= \lim_{n\to\infty} d(\lambda, T)\,\|y^* + \sigma x_n^*\|^2 \\ &\leqq \lim_{n\to\infty} |\,T_\lambda(y^* + \sigma x_n^*), y^* + \sigma x_n^*)\,| \\ &= |\,(T_\lambda y^*, y^*) + \sigma\,\|y^*\|^2\,|\,, \end{aligned}$$

which by the arbitrariness of σ implies $\|y^*\| = 0$, that is, $y = T_\lambda x$. Hence $\mathscr{R}(T_\lambda) = \mathscr{H}$, as the theorem required.

Passing to the proof of the second part we remark first that by the Closed Graph Theorem T_λ is itself bounded. Taking then μ_0 as the point in $\mathscr{N}(T)$ closest to λ we observe that $\operatorname{Re}((\lambda - \mu)/(\lambda - \mu_0)) \geqq 1$ for any $\mu \in \mathscr{N}(T)$, since $\mathscr{N}(T)$ is convex ([21], p. 131). So, if $\alpha = t\lambda/(\lambda - \mu_0)$ with t real and positive,

$$\begin{aligned} \frac{\|(\alpha\lambda^{-1}T + (1-\alpha)I)x\|^2}{\|x\|^2} &= 1 + \left|\frac{\alpha}{\lambda}\right|^2 \frac{\|T_\lambda x\|^2}{\|x\|^2} - 2\operatorname{Re}\left[\frac{\alpha}{\lambda}\left(\lambda - \frac{(Tx, x)}{(x, x)}\right)\right] \\ &\leqq 1 + \frac{t^2}{d^2}\,\|T_\lambda\|^2 - 2t\,. \end{aligned}$$

For $t < 2d^2(\lambda, T)/\|T_\lambda\|^2$ the right hand member is <1, the operator

$$\alpha\lambda^{-1}T + (1-\alpha)I$$

is contractive, and the equation $x = (\alpha\lambda^{-1}T + (1-\alpha)I)x - \alpha\lambda^{-1}y -$ equivalent to (1) – can be solved by iteration, according to Banach Contraction Principle ([14], p. 43). This completes the proof.

As M. H. Stone remarked ([22], p. 149), the theorem is not valid for linear mappings only densely defined: Any maximal symmetric not essentially self-adjoint mapping has the real axis as its numerical range and only a half plane as its resolvent.

Before entering the nonlinear realm we wish to point out an interesting by-product of Theorem I, no doubt known, but perhaps not in the simple form it appears here. We have proved that if $\lambda \notin \overline{\mathscr{N}}(T)$ then T_λ, and hence T, is bounded. Therefore, if $\overline{\mathscr{N}}(T)$ does not exhaust the complex plane T is bounded. The converse of this being obvious, and since $\mathscr{N}(T)$ and $\overline{\mathscr{N}}(T)$ are simultaneously different from the whole complex plane by the convexity of $\mathscr{N}(T)$, we may conclude:

COROLLARY. *An everywhere defined linear mappings of the Hilbert space into itself is bounded if and only if its numerical range does not exhaust the complex plane.*

To make ready for the extension of Theorem I to nonlinear map-

pings we begin by assigning a meaning to the various terms appearing in its statement when applied to general mappings. The notion of numerical range generalizes naturally as follows:

DEFINITION 1. The numerical range of a mapping $T: \mathscr{H} \to \mathscr{H}$ with domain $\mathscr{D}(T)$ is the set of complex numbers

$$\mathscr{N}(T) = \left\{ \frac{(Tx_1 - Tx_2, x_1 - x_2)}{\| x_1 - x_2 \|^2}, \qquad x_1 \neq x_2, \quad x_1, x_2 \in \mathscr{D}(T) \right\}$$

In general little can be said of this set except that it is connected if T is continuous and $\mathscr{D}(T)$ connected. In the one-dimensional case, that is, when T is a mapping of complex numbers into complex numbers $\mathscr{N}(T)$ is nothing but the set of all incremental ratios $(Tx_1 - Tx_2)/(x_1 - x_2)$. For linear mappings it coincides with the ordinary notion and is always a convex set ([22], p. 130); if in addition the mapping is maximal normal the closure of $\mathscr{N}(T)$ coincides with the convex hull of the spectrum.

For any complex number λ we shall denote by $d(\lambda, T)$ and $D(\lambda, T)$ the infimum and the supremum respectively of the distances from λ to points of $\mathscr{N}(T)$. Hence

$$(3) \qquad d(\lambda, T) \leqq |\lambda - (Tx_1 - Tx_2, x_1 - x_2)/\| x_1 - x_2 \|^2 \leqq D(\lambda, T),$$

or, writing $T_\lambda = T - \lambda I$,

$$(4) \qquad \begin{aligned} d(\lambda, T) \| x_1 - x_2 \|^2 &\leqq |(T_\lambda x_1 - T_\lambda x_2, x_1 - x_2)| \\ &\leqq D(\lambda, T) \| x_1 - x_2 \|^2, \end{aligned}$$

whence by Schwartz' inequality it follows

$$(5) \qquad d(\lambda, T) \| x_1 - x_2 \| \leqq \| T_\lambda x_1 - T_\lambda x_2 \|.$$

The ordinary norm for linear operators becomes the Lipschitz norm defined as follows:

DEFINITION 2. The Lipschitz norm of an operator $T: \mathscr{H} \to \mathscr{H}$ is the number (the value $+\infty$ being allowed)

$$\| T \| = \sup \frac{\| Tx_1 - Tx_2 \|}{\| x_1 - x_2 \|},$$

where the sup is taken with regard to all pairs of distinct points in $\mathscr{D}(T)$. If $\| T \| < \infty$, T is said to be Lipschitzian.

Let us finally recall that a demicontinuous mapping $T: \mathscr{H} \to \mathscr{H}$ is one which is continuous from the strong topology in the domain to the weak topology in the range. Demiclosedness is the closedness notion associated with this type of continuity, and says that a mapping

T is demiclosed if the strong convergence of x_α to x together with the weak convergence of Tx_α to y imply $x \in \mathscr{D}(T)$ and $y = Tx$. In the sequel $\mathscr{B}_r(z)$ will denote the open ball of radius r about z, and $\overline{\mathscr{B}}_r(z)$ its closure.

These preliminaries settled, we may now state the main result concerning the extension of the existence part of Theorem I:

THEOREM II. *Let $T: \mathscr{H} \to \mathscr{H}$ be a demiclosed mapping defined on a closed ball $\overline{\mathscr{B}}_r(0)$, demicontinuous over any finite dimensional subspace, and vanishing at the origin. Furthermore, λ let be a number at a positive distance $d = d(\lambda, T)$ from the numerical range of T. Then the equation*

$$Tx - \lambda x = y \tag{6}$$

has a unique solution in $\overline{\mathscr{B}}_r(0)$ for any $y \in \overline{\mathscr{B}}_{rd}(0)$. The operator $(T - \lambda I)^{-1}$ thus defined is Lipschitzian and $\| (T - \lambda I)^{-1} \| \leqq d^{-1}(\lambda, T)$.

This theorem was first proved by the author [24] for mappings densely defined in a ball under the assumption of local cross-boundedness, and sequential demiclosedness, conditions which—it was shown later [14]—imply to local boundedness and $\mathscr{D}(T) = \overline{\mathscr{B}}_r(0)$. Subsequently, F. Browder [11] showed that local boundedness can be dropped if sequential demiclosedness is stepped up to demiclosedness. We do not give the proof here, the reader may find it in the mentioned articles. Instead we discuss an idea pointed out to the author by G. B. Minty according to which if T is everywhere defined the "global" numerical range, a rather unwieldy object, can be replaced by the "local closed" numerical range, a smaller and easier to handle set, defined as follows:

DEFINITION 3. The local closed numerical range of a mapping $T: \mathscr{H} \to \mathscr{H}$ is the set $\tilde{\mathscr{N}}(T) = \bigcap_{r>0} \overline{\mathscr{N}_r}(T)$, where

$$\mathscr{N}_r(T) = \left\{ \frac{(Tx_1, Tx_2, x_1 - x_2)}{\| x_1 - x_2 \|^2}, \quad x_1, x_2 \in \mathscr{D}(T), \ 0 < \| x_1 - x_2 \| \leqq r \right\}. \tag{7}$$

This is Minty's extension of Theorem II and proof.

THEOREM III. (*Minty*). *Let $T: \mathscr{H} \to \mathscr{H}$ be a demiclosed mapping everywhere defined in $\mathscr{H}$, demicontinuous on any finite dimensional subspace therein. For any complex number λ at a positive distance $\tilde{d} = \tilde{d}(\lambda, T)$ from the local closed numerical range of T, and for any $y \in \mathscr{H}$, the equation $\lambda x = Tx - y$ has a unique solution. The operator $(T - \lambda I)^{-1}$ thus defined is Lipschitzian, and $\| (T - \lambda I)^{-1} \| \leqq \tilde{d}^{-1}(\lambda, T)$.*

Proof. Clearly the $\overline{\mathscr{N}_r}(T)$ are nested closed sets in the sense that $r_1 < r_2$ implies $\overline{\mathscr{N}_{r_1}} \subset \overline{\mathscr{N}_{r_2}}$, and therefore the distance $\tilde{d}(\lambda, T)$ from λ to their intersection is equal to the limit of the distance $d_r(\lambda, T)$ to $\mathscr{N}_r(T)$ as $r \to 0$. Now assume r to be so small that $d_{2r} > 0$, and take any point $y_0 \in \mathscr{R}(T_\lambda)$ and one of its pre-images x_0 under $T_\lambda = T - \lambda I$. The mapping $T(x^0 + x) - y_0 - \lambda x_0$ restricted to $\overline{\mathscr{B}_r}(0)$ and the complex number λ satisfy all requirements of Theorem II with $d \geqq d_{2r}$, because the numerical range of the mapping is contained in $\mathscr{N}_{2r}(T)$. Hence, the range of $T(x_0 + x) - y_0 - \lambda(x + x_0)$ contains $\mathscr{B}_{rd_{2r}}(0)$, and y_0 is the center of a ball of radius rd_{2r} contained in the range of T_λ. This being so for any $y_0 \in \mathscr{R}(T_\lambda)$, $\mathscr{R}(T_\lambda)$ must be the whole space $\mathscr{H}$.

More delicate is the proof that T_λ is one-to-one; we take the finite dimensional case first. As any point in $\mathscr{D}(T)$ is the center of a ball of radius r on which T_λ is one-to-one, the pre-image by T_λ of any point $y_0 \in \mathscr{H}$ is a discrete set and therefore countable; let $x_1, x_2, \cdots$ be its elements. Hence $x_i \notin \mathscr{B}_r(x_j)$ for $i \neq j$, and the balls $\mathscr{B}_{r/2}(x_i)$, $i = 1, 2, \cdots$, are disjoint. Further, any pre-image of any point in $\mathscr{B}_{(r/2)d_{2r}}(y_0)$ has a ball of radius $r/2$ about it containing one of the x_i's; in other words it is in one of the balls $\mathscr{B}_{r/2}(x_i)$. Thus, the pre-image of $\mathscr{B}_{(r/2)d_{2r}}(y_0)$ by T_λ decomposes into a class of disjoint sets, each contained in just one ball $\mathscr{B}_{r/2}(x_i)$. These are all open sets homeomorphic to $\mathscr{B}_{(r/2)d_{2r}}(y_0)$, since T_λ restricted to any $\mathscr{B}_r(x_i)$ is a topological mapping. Under these conditions, y_0 being arbitrary and $\mathscr{H}$ simply connected, it is enough to invoke the "Monodromy Principle" ([12], p. 146) to conclude that T_λ is one-to-one. To extend this to infinite dimensions, simply take any two points x_1 and x_2 in $\mathscr{D}(T)$, construct the space $\mathscr{H}_0$ generated by x_1, x_2, and apply the above result to the restriction of the mapping $E_0 T$ to $\mathscr{H}_0$ (here E_0 is the orthogonal projection on $\mathscr{H}_0$). If $x_1 \neq x_2$, then $E_0 T_\lambda x_1 \neq E_0 T_\lambda x_2$, and in consequence $T_\lambda x_1 \neq T_\lambda x_2$.

From the above arguments one sees that $\| T_\lambda x_1 - T_\lambda x_2 \| < r d_{2r}$ implies $\| x_1 - x_2 \| < r$, and this is turn, $\| x_1 - x_2 \| \leqq d_{2r}^{-1} \| T_\lambda x_2 - T_\lambda x_1 \|$. This proves the local Lipschitz character of T_λ^{-1}, of which—as we shall see—the global Lipschitz character is an easy consequence. For any couple of points x' and x'' in $\mathscr{D}(T)$ and any positive integer n, let $x_0, \cdots, x_n$ be the points in $\mathscr{D}(T)$ such that

$$T_\lambda x_k = T_\lambda x' + \frac{k}{n}(T_\lambda x'' - T_\lambda x'), \qquad k = 0, \cdots, n.$$

If n is sufficiently large,

$$\| T_\lambda x_k - T_\lambda x_{k-1} \| = \frac{1}{n} \| Tx'' - Tx' \|$$

is smaller than rd_{2r} and

$$\| x_k - x_{k-1} \| \leqq d_{2r} \| T_\lambda x_k - T_\lambda x_{k-1} \| = \frac{d_{2r}}{n} \| T_\lambda x'' - T_\lambda x' \| ,$$
$$k = 0, 1, \cdots, n .$$

Adding these inequalities together one obtains,

$$\| x'' - x' \| \leqq \sum_1^n \| x_k - x_{k-1} \| \leqq d_{2r} \| T_\lambda x'' - T_\lambda x' \| ,$$

whence passing to the limit $r \to 0$, the sought for Lipschitz condition follows.

We now turn to the discussion of the second part of Theorem I, namely to the construction of the solution of equation (1) as a limit of a sequence of succesive averages

$$x_k = (1 - \alpha_k)x_{k-1} + \alpha_k \lambda^{-1}(Tx_{k-1} - y) , \tag{8}$$

between a vector and its transform by $\lambda^{-1}(T - y)$, built from an original approximation x_0 by means of a suitable choice of the averaging constants α_k. For the purpose of generalization the fact that all averaging constants can be taken as equal should be looked upon as merely accidental to the linear case. The extension to nonlinear equations is here again possible but, as it was to be expected, not under conditions as general as those for the first part. Furthermore, to prevent the iterates from running out of the domain of definition when the mappings are not everywhere defined, we shall find it necessary to modify scheme (8) with an extra factor in front of y. This, however, is a point of secondary importance. Our first step is to show that the idea is feasible, that is, that there exists at least one sequence $\{\alpha_k\}$ leading to the solution:

THEOREM III. *Let $T: \mathscr{H} \to \mathscr{H}$ be a mapping vanishing at the origin, defined on $\overline{\mathscr{B}}_r(0)$, continuous and bounded there. Further, let λ be a nonvanishing complex number at a positive distance $d = d(\lambda, T)$ from the numerical range of T, and x the solution of the equation $\lambda x = Tx - y$ for some $y \in \overline{\mathscr{B}}_{rd}(0)$. Then for any sequence of complex numbers $\{\gamma_k\}_1^\infty$ such that*

$$|\gamma_k| \leqq 1 , \qquad \sum_1^\infty (1 - |\gamma_k|^2) = \infty , \tag{9}$$

and any vector x_0 satisfying $\| x \| + \| x - x_0 \| \leqq r$, the vector sequence $\{x_k\}_0^\infty$ having x_0 as its first term and the successive ones determined by

$$x_k = (1 - \alpha_k)x_{k-1} + \alpha_k \lambda^{-1}(Tx_{k-1} - y) \tag{10}$$

where the α_k's are calculated from

$$\|T_\lambda x - T_\lambda x_{k-1}\|^2 \alpha_k = -\lambda(\overline{T_\lambda x - T_\lambda x_{k-1}, x - x_{k-1}})(1 + \gamma_k), \quad (T_\lambda = T - \lambda I) \tag{11}$$

can be continued indefinitely and converges to x.

Proof. Notice that α_k is uniquely determined by (10) except when $\|T_\lambda x - T_\lambda x_{k-1}\| = 0$, in which case any complex number satisfies (11). But then $x_{k-1} = x$ by (6) and x_k coincides with x regardless of the value of α_k, and so do all successive terms. As the theorem's conclusions are obviously valid for the resulting sequence, we may discard its occurrence and assume that $\|T_\lambda x - T_\lambda x_{k-1}\|$ vanishes for no k.

The sequence $\{x_k\}$ may terminate only by running out of the domain of definition of T. To see that this does not happen it is enough to show that the relation $\|x\| + \|x - x_{k-1}\| \leqq r$ (valid for $k - 1 = 0$) propagates from one term to the next. Subtracting (10) from

$$x = (1 - \alpha_k)x + \alpha_k\lambda^{-1}(Tx - y)$$

one obtains

$$\begin{aligned} x - x_k &= (1 - \alpha_k)(x - x_{k-1}) + \alpha_k\lambda^{-1}(Tx - Tx_{k-1}) \\ &= (x - x_{k-1}) + \alpha_k\lambda^{-1}(T_\lambda x - T_\lambda x_{k-1}) , \end{aligned}$$

and hence

$$\begin{aligned} \|x - x_k\|^2 = \|x - x_{k-1}\|^2 &+ 2\,\mathrm{Re}\,[\alpha_k\lambda^{-1}(T_\lambda x - T_\lambda x_{k-1}, x - x_{k-1})] \\ &+ |\alpha_k\lambda^{-1}|^2 \|T_\lambda x - T_\lambda x_{k-1}\|^2 . \end{aligned}$$

Replacing α_k by its value given by (11), and transferring terms,

$$\begin{aligned} &\|x - x_{k-1}\|^2 - \|x - x_k\|^2 \\ &= \frac{|(T_\lambda x - T_\lambda x_{k-1}, x - x_{k-1})|^2}{\|T_\lambda x - T_\lambda x_{k-1}\|^2}[2\,\mathrm{Re}\,(1 + \gamma_k) - |1 + \gamma_k|^2] \\ &= \frac{|(T_\lambda x - T_\lambda x_{k-1}, x - x_{k-1})|^2}{\|T_\lambda x - T_\lambda x_{k-1}\|^2}(1 - |\gamma_k|^2) . \end{aligned} \tag{12}$$

Hence

$$\|x - x_k\| \leqq \|x - x_{k-1}\| ,$$

and

$$\|x\| + \|x - x_k\| \leqq r$$

if

$$\|x\| + \|x - x_{k-1}\| \leqq r.$$

The existence of $\{x_k\}_0^\infty$ is thus fully guaranteed.

Making use of the estimates

$$|(T_\lambda x - T_\lambda x_{k-1}, x - x_{k-1})|^2 \geqq d^2 \|x - x_{k-1}\|^2,$$
$$\|T_\lambda x - T_\lambda x_{k-1}\|^2 \leqq B^2,$$

one obtains from (12),

$$\|x - x_{k-1}\|^2 - \|x - x_k\|^2 \geqq \frac{d^2 \|x - x_{k-1}\|^4}{B^2}(1 - |\gamma_k|^2).$$

Dividing this inequality by $\|x - x_{k-1}\|^2 \|x - x_k\|^2$ and observing that $(\|x - x_{k-1}\|^2)/(\|x - x_k\|^2) \geqq 1$, one gets

$$\frac{1}{\|x - x_k\|^2} - \frac{1}{\|x - x_{k-1}\|^2} \geqq \frac{\|x - x_{k-1}\|^2}{\|x - x_k\|^2} \frac{d^2(1 - |\gamma_k|^2)}{B^2}$$
$$\geqq \frac{d^2(1 - |\gamma_k|^2)}{B^2}$$

which added together from $k = 1$ to $k = n$ yield

$$\frac{1}{\|x - x_n\|^2} \geqq \frac{1}{\|x - x_0\|^2} + \frac{d^2}{B^2} \sum_1^n (1 - |\gamma_k|^2).$$

Clearly $\|x - x_n\| \to 0$ by the divergence of $\sum_1^\infty (1 - |\gamma_k|^2)$.

We have thus demonstrated the possibility of passing by repeated averaging from an approximating x_0 with $\|x\| + \|x_0 - x\| < r$ to the solution x itself. If no such approximation is known a priori, the averaging scheme may be modified so as to yield the solution starting from any approximation whatever. We shall not enter into the details here, but with the reassurance derived from Theorem 4 shall proceed to the more substantial task of establishing a recursive procedure for the determination of the averaging factors. This we are able to do only under the assumption that the mapping satisfy a Hölder condition of exponent $>1/2$. Yet, in order to bring the ideas involved into focus we shall confine our attention to the simpler case of cross-Lipschitzian mappings. These are mappings for which the quantities

$$\|T\|^\perp = \sup_{\substack{x_1 \neq x_2 \\ x_1, x_2 \in \mathfrak{D}(T)}} \left\| Tx_1 - Tx_2 - \left(Tx_1 - Tx_2 \frac{x_1 - x_2}{\|x_1 - x_2\|}\right) \frac{(x_1 - x_2)}{\|x_1 - x_2\|} \right\| \Big/ \|x_1 - x_2\|$$

called cross-Lipschitz norms, are finite. It has been shown [14] that this notion of continuity, obtained by replacing the ordinary increment $Tx_1 - Tx_2$ by its component orthogonal to $x_1 - x_2$, is strictly weaker than Lipschitz continuity and stronger than Hölder continuity of any order $\nu < 1$. The interested reader will find the modifications necessary to extend the theory below to ν-Hölder continuous mappings ($\nu > 1/2$) in the

Appendix at the end of [25]. The next two lemmas form the core of the averaging theory.

LEMMA I. *Let* $T: \mathscr{H} \to \mathscr{H}$ *be a continuous, cross-Lipschitzian mapping defined on* $\overline{\mathscr{B}_r}(0)$ *vanishing at the origin,* λ *a nonvanishing complex number at a positive distance* $d = d(\lambda, T)$ *from the numerical range of* T, y *a point in* $\overline{\mathscr{B}_{rd}}(0)$. *Then for any* $x_0 \in \overline{\mathscr{B}_r}(0)$ *not a solution of* (1) *and any complex number* α *such that*

$$x_\alpha = (1 - \alpha)x_0 + \alpha\lambda^{-1}(Tx_0 - y)$$

belongs to $\overline{\mathscr{B}_r}(0)$,

$$\frac{\| T_\lambda x_0 - y \|^2 - \| T_\lambda x_\alpha - y \|^2}{\| T_\lambda x_0 - y \|^2} \geqq R(1 - |\gamma|^2), \quad (T_\lambda = T - \lambda I), \tag{13}$$

where

$$R = \frac{d^2(\lambda, T)}{d^2(\lambda, T) + (\| T \|^\perp)^2}, \tag{14}$$

$$\gamma = \frac{\alpha}{R\lambda} \frac{(T_\lambda x_0 - T_\lambda x_\alpha, x_0 - x_\alpha)}{\| x_0 - x_\alpha \|^2} + 1. \tag{15}$$

Proof. We have

$$\begin{aligned} \| T_\lambda x_\alpha - y \|^2 &= \| (T_\lambda x_0 - y) + (T_\lambda x_\alpha - T_\lambda x_0) \|^2 \\ &= \| T_\lambda x_0 - y \|^2 + \| T_\lambda x_\alpha - T_\lambda x_0 \|^2 \\ &\quad + 2 \operatorname{Re} (T_\lambda x_\alpha - T_\lambda x_0, T_\lambda x_0 - y), \end{aligned}$$

and

$$\begin{aligned} &\frac{\| T_\lambda x_0 - y \|^2 - \| T_\lambda x_\alpha - y \|^2}{\| T_\lambda x_0 - y \|^2} \\ &= -2 \operatorname{Re} \frac{(T_\lambda x_\alpha - T_\lambda x_0, T_\lambda x_0 - y)}{\| T_\lambda x_0 - y \|^2} - \frac{\| T_\lambda x_\alpha - T_\lambda x_0 \|^2}{\| T_\lambda x_0 - y \|^2}. \end{aligned} \tag{16}$$

Since

$$x_\alpha = x_0 + \frac{\alpha}{\lambda}(T_\lambda x_0 - y), \tag{17}$$

the first term on the right of (16) is readily seen to be $-2 \operatorname{Re} (R(\gamma - 1))$. As to the second, decomposing the increment $T_\lambda x_\alpha - T_\lambda x_0$ into its longitudinal and transversal components and observing that the latter is equal to the transversal component of $Tx_\alpha - Tx_0$, whose modulus does not exceed $\| T \|^\perp \| x_\alpha - x_0 \|$, one obtains

$$
\begin{aligned}
\frac{\| T_\lambda x_\alpha - T_\lambda x_0 \|^2}{\| T_\lambda x_0 - y \|^2} &\leqq \frac{1}{\| T_\lambda x_0 - y \|^2}\left[\left| \frac{(T_\lambda x_\alpha - T_\lambda x_0,\, x_\alpha - x_0)}{\| x_\alpha - x_0 \|} \right|^2 \right. \\
&\qquad \left. + (\| T \|^\perp)^2 \| x_\alpha - x_0 \|^2 \right] \\
&\leqq \left| \frac{\alpha}{\lambda} \frac{(T_\lambda x_\alpha - T_\lambda x_0,\, x_\alpha - x_0)}{\| x_\alpha - x_\alpha \|^2} \right|^2 + \left| \frac{\alpha}{\lambda} \right|^2 (\| T \|^\perp)^2 \\
&= | R(\gamma - 1) |^2 + \left| \frac{\alpha}{\lambda} \right|^2 (\| T \|^\perp)^2 .
\end{aligned} \tag{18}
$$

Notice now that by virtue of (5), $| R(\gamma - 1)(\lambda/\alpha) | \geqq d$, and so that $| \alpha/\lambda | \leqq (R/d) | r - 1 |$ and insert this in (18),

$$
\frac{\| T_\lambda x_\alpha - T_\lambda x_0 \|^2}{\| T_\lambda x_0 - y \|^2} \leqq R^2 \left(1 + \frac{(\| T \|^\perp)^2}{d^2}\right) | \gamma - 1 |^2 = R | r - 1 |^2 .
$$

Therefore, returning to (16),

$$
\begin{aligned}
\frac{\| T_\lambda x_0 - y \|^2 - \| T_\lambda x_\alpha - y \|^2}{\| T_\lambda x_0 - y \|^2} &\geqq -2 \operatorname{Re} (R(\gamma - 1)) - R | \gamma - 1 |^2 \\
&= R(1 - | \gamma |^2) ,
\end{aligned}
$$

as we set out to prove.

Formula (13) is the key to the averaging theory; it points to the fundamental fact that, by averaging, the error committed in solving equation (1) can be made strictly smaller whenever there is an α whose corresponding γ lies in the interior of the unit disc about the origin. This calls for a study of the mapping $\alpha \rightarrow \gamma$, which is precisely the content of the next lemma. To state the lemma we shall need the concept of "admissible approximation".

DEFINITION 4. Let T: $\mathscr{H} \rightarrow \mathscr{H}$ be a mapping defined on $\overline{\mathscr{B}_r}(0)$ vanishing at the origin, λ a complex number at a positive distance $d = d(\lambda, T)$ from the numerical range of T, and y a vector in $\overline{\mathscr{B}_{rd}}(0)$. A vector $x_0 \subset \overline{\mathscr{B}_r}(0)$ is said to be an admissible approximation to the solution of equation (1), if

$$
\| y \| + \| T_\lambda x_0 - y \| \leqq rd , \qquad (T_\lambda = T - \lambda I) . \tag{19}
$$

It is important to notice that the test for admissibility does not require the knowledge of the solution. One sees at once that $x_0 = 0$ is admissible for those equations with $\| y \| \leqq rd/2$, and that any x_0 is admissible only if T is defined everywhere. For y's with $\| y \| = rd$, x_0 is admissible if and only if it is a solution.

LEMMA II. *Let T: $\mathscr{H} \rightarrow \mathscr{H}$ be a continuous, cross-Lipschitzian mapping defined on $\overline{\mathscr{B}_r}(0)$ vanishing at the origin, λ a nonvanishing*

complex number at a positive distance $d = d(\lambda, T)$ *from the numerical range of* T, y *a vector in* $\overline{\mathscr{B}}_{r_d}(0)$, *and* $x_0 \in \mathscr{B}_r(0)$ *an admissible approximation, not the solution of* (1). *Then if*

$$x_\alpha = (1 - \alpha)x_0 + \alpha\lambda^{-1}(Tx_0 - y)\,,$$

the mapping of the complex plane into itself

$$\alpha \longrightarrow \gamma = \frac{\alpha}{R\lambda}\left(\frac{T_\lambda x_0 - T_\lambda x_\alpha,\ x_0 - x_\alpha}{\|x_0 - x_\alpha\|^2}\right) + 1 \tag{20}$$

is a homeomorphism having as domain the circular disc

$$\left|\alpha + \lambda\frac{(T_\lambda x_0 - y, x_0)}{\|T_\lambda x_0 - y\|^2}\right|^2 \leqq |\lambda|^2\frac{r^2 - \|x_0\|^2}{\|T_\lambda x_0 - y\|^2} + \left|\lambda\frac{(T_\lambda x_0 - y, x_0)}{\|T_\lambda x_0 - y\|^2}\right|^2, \tag{21}$$

and as range a closed domain containing the unit disc about the origin. Its inverse is Lipschitzian with norm not exceeding $R|\lambda|/d$.

Proof. The mapping (20) is defined for those α's for which $\|x_\alpha\| \leqq r$. By (17) such a set coincides with the circular disc (21). On use of (17) again one may write, setting $y_0 = T_\lambda x_0$,

$$\gamma = \frac{1}{R}\frac{(y_0 - T_\lambda(x_0 + \alpha\lambda^{-1}y_0),\ y_0 - y)}{\|y_0 - y\|^2} + 1\,. \tag{22}$$

It is clear that γ is a continuous function of α. Moreover, from (22) one may also derive

$$\begin{aligned}\frac{\gamma_1 - \gamma_2}{\alpha_1 - \alpha_2} &= -\frac{1}{R}\frac{(T_\lambda x_{\alpha_1} - T_\lambda x_{\alpha_2},\ y_0 - y)}{(\alpha_1 - \alpha_2)\|y_0 - y\|^2} \\ &= -\frac{1}{R\lambda}\frac{(T_\lambda x_{\alpha_1} - T_\lambda x_{\alpha_2},\ x_{\alpha_1} - x_{\alpha_2})}{\|x_{\alpha_1} - x_{\alpha_2}\|^2}\end{aligned}$$

and by (6)

$$\left|\frac{\gamma_1 - \gamma_2}{\alpha_1 - \alpha_2}\right| \geqq \frac{d}{R|\lambda|}\,. \tag{23}$$

Hence the mapping is one-to-one and has a Lipschitzian inverse with the required norm. As it was continuous it is a homeomorphism, and by Brouwer's Domain Invariance Theorem ([1], p. 156) its range is a closed domain, whose boundary is in a one-to-one correspondence with the boundary of disc (21), that is, with the set of α's for which $\|x_\alpha\| = r$. Since $\|x_0\| < r$, $\alpha = 0$ is an interior point of (21) and therefore its image $\gamma = 1$ an interior point of the range, which by this fact has a nonempty intersection with the disc $|\gamma| < 1$. Moreover, no point of the disc is on the range boundary, because if $|\gamma| < 1$

then by (13) and the admissibility of x_0,

$$\begin{aligned} \| x_\alpha \| &\leqq d^{-1} \| T_\lambda x_\alpha \| \leqq d^{-1}(\| T_\lambda x_\alpha - y \| + \| y \|) \\ &< d^{-1}(\| T_\lambda x_0 - y \| + \| y \|) \leqq r , \end{aligned}$$

showing that α is not in the boundary of (21) and hence that neither is γ in the range boundary. Therefore the disc $|\gamma| < 1$ being connected must be entirely contained in the range of the mapping (20). The lemma has thus been proved.

This lemma assures us that any γ in the unit disc is the image of an α, far in excess of what is really needed for the effective calculation of the solution of equation (1). In fact it is enough to be able to determine just *one* α so that γ lies in the interior of the unit disc uniformly, say, within a disc of smaller radius ρ. If $\delta(\varepsilon)$ is the modulus of continuity of γ as a function of α, any $\delta(\rho)$-net covering the disc (21) contains one such α at least, and so if $\delta(\varepsilon)$ is known the inequality $|\gamma(\alpha)| \leqq \rho$ can be solved effectively.

THEOREM V. *Let $T: \mathscr{H} \to \mathscr{H}$ be a continuous, cross-Lipschitzian mapping defined on $\overline{\mathscr{B}}_r(0)$ vanishing at the origin, λ a nonvanishing complex number at a positive distance $d = d(\lambda, T)$ from the numerical range of T, and y any point in $\overline{\mathscr{B}}_{rd}(0)$. Then for any admissible approximation x_0 and any sequence $\{\gamma_k\}_1^\infty$ of complex numbers satisfying*

$$|\gamma_k| \leqq 1 , \qquad \sum_1^\infty (1 - |\gamma_k|^2) = \infty \tag{24}$$

the sequence of successive averages

$$x_k = (1 - \alpha_k)x_{k-1} + \alpha_k \lambda^{-1}(Tx_{k-1} - y) \tag{25}$$

starting with x_0 and constructed with averaging factors α_k satisfying the equations

$$\| x_k - x_{k-1} \|^2 (\gamma_k - 1) = \frac{\alpha_k}{R\lambda} (T_\lambda x_k - T_\lambda x_{k-1}, x_k - x_{k-1}) , \tag{26}$$

can be continued indefinitely and converges to the solution of (1). *Moreover, if x is the solution,*

$$\| x_n - x \| \leqq \frac{\| T_\lambda x_0 - y \|}{d(\lambda, T)} \prod_{k=1}^{n} [1 - R(1 - |\gamma_k|^2)]^{1/2} . \tag{27}$$

Proof. Should two consecutive terms coincide, say x_{k-1} and x_k, then it would follow from (25) that x_{k-1} is a solution of the equation and all successive terms are equal to that solution regardless of the

values of the α_k's, which, consistently with this, are no longer determined by (26). As the Lemma clearly holds in this case, we may exclude it from our considerations.

In other cases, inequality (13) indicates that admissibility is passed from one term to the next. Hence the x_k's never come out of $\overline{\mathscr{B}_r}(0)$, are defined and admissible for all values of k, and

$$\frac{\| T_\lambda x_{k-1} - y \|^2 - \| T_\lambda x_k - y \|^2}{\| T_\lambda x_{k-1} - y \|^2} \geqq R(1 - |\gamma_k|^2), \qquad k = 1, 2, \cdots,$$

that is,

$$\| T_\lambda x_k - y \| \leqq \| T_\lambda x_{k-1} - y \| [1 - R(1 - |\gamma_k|^2)]^{1/2}, \qquad k = 1, 2, \cdots.$$

Multiplying these inequalities together, one obtains

$$\| T_\lambda x_k - y \| \leqq \| T_\lambda x_0 - y \| \prod_{k=1}^{n} [1 - R(1 - |\gamma_k|^2)]^{1/2}, \tag{28}$$

which, since the divergence of $\sum_1^\infty (1 - |\gamma_k|^2)$ implies the divergence to zero of the product on the right, proves that $T_\lambda x_n \to y$. But T_λ^{-1} being Lipschitzian the x_n converge to an x, and $T_\lambda x = y$. Finally, (27) follows from (28) upon recalling that

$$\| x_n - x \| \leqq d^{-1}(\lambda, T) \| T_\lambda x_n - T_\lambda x \|.$$

Hence, a recursive averaging procedure to construct the solution of equation (1) from any admissible approximation has been established. The slightly more complicated scheme below applies to cases where no admissible approximation is known beforehand. The leading idea is the following: if x_δ is admissible for $\lambda x = Tx - \delta y$ then by averaging an $x_{\delta'}$ can be found which is admissible for $\lambda x = Tx - \delta' y$ with some δ' larger than δ, and so, starting with $x_{\delta_0} = 0$—which is admissible for $\lambda x = Tx - \delta_0 y$, $0 \leqq \delta_0 \leqq 1/2$—it is possible to construct by repeated averaging an admissible approximation for $\lambda x = Tx - \delta_1 y$, $\delta_1 = 1$, and then the solution of (1).

THEOREM IV. *Let T: $\mathscr{H} \to \mathscr{H}$ be a continuous, cross-Lipschitzian mapping defined on $\overline{\mathscr{B}_r}(0)$ vanishing at the origin, λ a nonvanishing complex number at a positive distance $d = d(\lambda, T)$ from the numerical range of T, and y any point in $\overline{\mathscr{B}_{rd}}(0)$. Then for any sequence $\{\gamma_k\}_1^\infty$ of complex numbers satisfying*

$$|\gamma_k| \leqq 1, \qquad \sum_1^\infty (1 - |\gamma_k|^2) = \infty \tag{29}$$

the sequence of modified successive averages

$$x_k = (1 - \alpha_k) x_{k-1} + \alpha_k \lambda^{-1} (T x_{k-1} - \delta_{k-1} y) \tag{30}$$

having $x_0 = 0$ *as its first term and successive ones contructed with correcting factors*

$$(31)\quad \delta_k = 1 - (1 - \delta_0)\prod_{h=1}^{k}\frac{1}{2}\{1 + [1 - R(1 - |\gamma_h|^2)]^{1/2}\}, \qquad 0 < \delta_0 \leqq \frac{1}{2},$$

and averaging factors α_k *determined as solutions of*

$$(32)\quad \|x_k - x_{k-1}\|^2(\gamma_k - 1) = \frac{\alpha_k}{R\lambda}(T_\lambda x_k - T_\lambda x_{k-1}, x_k - x_{k-1}).$$

can be continued indefinitely, and converges to the solution of (1). *Further if* x *is the solution,*

$$(33)\quad \|x_n - x\| \leqq \frac{2(1 - \delta_0)\|y\|}{d(\lambda, T)}\prod_{k=1}^{n}\frac{1}{2}\{1 + [1 - R(1 - |\gamma_k|^2)]^{1/2}\}.$$

Proof. The factors in the product defining δ_k being all positive numbers not exceeding one, the δ_k's form a nondecreasing sequence in the interval $(0, 1)$. Further, since

$$\frac{R}{4}(1 - |\gamma_k|^2) \leqq 1 - \frac{1}{2}\{1 + [1 - R(1 - |\gamma_k|^2)]^{1/2}\} \leqq \frac{R}{2}(1 - |\gamma_k|^2)$$

the divergence of $\sum_1^\infty (1 - |\gamma_k|^2)$ implies the divergence to zero of the product on the right member of (31). Therefore $\delta_k \nearrow 1$.

We prove that x_k is defined for all values of the index k by showing that if $x_{k=1}$ satisfies $\|\delta_{k-1}y\| + \|T_\lambda x_{k-1} - \delta_{k-1}y\| \leqq \|y\|$ then x_k is defined and $\|\delta_k y\| + \|T_\lambda x_k - \delta_k y\| \leqq \|y\|$. Under the above assumptions x_{k-1} is admissible for $\lambda x = T_\lambda x - \delta_{k-1}y$. If it is a solution of this equation then $x_k = x_{k-1}$, regardless of the value of α_k, which is now undetermined; if it is not a solution, Lemma 2 guarantees the existence of α_k satisfying (32), and x_k is again defined. Then in either case,

$$\begin{aligned}\|\delta_k y\| + \|T_\lambda x_k - \delta_k y\| &= \|\delta_k y\| + \|(T_\lambda x_k - \delta_{k-1}y) - (\delta_k - \delta_{k-1})y\| \\ &\leqq (2\delta_k - \delta_{k-1})\|y\| + \|T_\lambda x_k - \delta_{k-1}y\|.\end{aligned}$$

By Lemma 1, (13)

$$\|T_\lambda x_k - \delta_{k-1}y\| \leqq [1 - R(1 - |\gamma_k|^2)]^{1/2}\|T_\lambda x_{k-1} - \delta_{k-1}y\|,$$

a relation which holds even if x_{k-1} is the solution of $\lambda x = Tx - \delta_{k-1}y$. Taking account of this in the previous inequality and recalling that $\|T_\lambda x_{k-1} - \delta_{k-1}y\| \leqq \|y\| - \|\delta_{k-1}y\|$ one obtains

$$\|\delta_k y\| + \|T_\lambda x_k - \delta_k y\| \leqq \{(2\delta_k - \delta_{k-1}) + [1 - R(1 - |\gamma_k|^2)]^{1/2}(1 - \delta_{k-1})\}\|y\|,$$

which is the desired conclusion, since by the definition of the δ_k's the

expression in braces on the right is equal to 1. Having proved the existence of the sequence $\{x_k\}_0^\infty$ and the relations

$$\| \delta_k y \| + \| T_\lambda x_k - \delta_k y \| \leqq \| y \|, \qquad k = 0, 1, \cdots, \tag{34}$$

the rest follows easily. We have

$$\begin{aligned} \| T_\lambda x_k - y \| &= \| T_\lambda x_k - \delta_k y + (\delta_k - 1) y \| \\ &\leqq \| T_\lambda x_k - \delta_k y \| + (1 - \delta_k) \| y \| \\ &\leqq 2(1 - \delta_k) \| y \|, \end{aligned} \tag{35}$$

and as $\delta_k \nearrow 1$, $T_\lambda x_k \to y$, and by the usual argument, $x_k \to x$, x satisfying $T_\lambda x = y$. Finally, coupling (35) with $\| x_k - x \| \leqq d^{-1} \| T x_k - y \|$, one obtains (33).

The computational strategy provided by Theorems V and VI is very broad and flexible. In practice one is not given the γ_k's from which one determines the α_k's; on the contrary, one tries to make judicious choices of the α_k's in order to place the γ_k's within the unit circle, and this only a finite number of times until the desired approximation is attained; naturally, the divergence of $\sum (1 - |\gamma^2|^2)$ never comes up in actual calculation. In some cases—presumably when the computations are done by hand—it may be preferable to reduce the number of steps by making them individually more effective, that is, by bringing the γ_k's as close as possible to zero, while in others, it may be more expedient to make rough estimates and thus gain in simplicity what is lost in accuracy. In a looser manner still, the α_k's make be picked at random—à la Monte-Carlo—keeping only those that improve the approximation; the method seems to be probabilistically convergent.

Averaging computational techniques have been used in the past—not always with success. (For a reference to extensive calculations see [3], Ch. X, and [2]). Aside from cases where no averaging could possibly converge, the failures may be attributed in part to the insistence in using constant averaging factors, after the linear model. In view of this situation it is not inappropriate to stop to investigate the circumstances under which the usual procedure works. We shall broaden the scope of our investigation so as to also be able to give a partial answer to questions of when the averaging factors can all be taken of same modulus, or of same argument. In terms of the family of homeomorphisms $\alpha \leftrightarrow \gamma$—which depend on λ, y, and x_0—these problems amount to a search for conditions under which the images of $|\gamma| < 1$ by the functions $|\alpha(\gamma)|$, or by the functions $\arg(\alpha(\gamma))$ have, for fixed λ and y and varying x_0, an open interval in common. Our findings in this direction are contained in the following lemma,

which can be considered as a continuation of Lemma II:

LEMMA III. *Under the hypotheses of Lemma IV, the following holds*:

a. *If the numerical range of T is bounded then for any t on the interval $(0, 2\,|\,\lambda\,|\,R/D(\lambda, T))$ there are α's in the disc* (21) *such that*

$$|\,\alpha\,| = t\,, \qquad |\,\gamma\,| \leqq 1 - \min\left(\frac{dt}{|\,\lambda\,|\,R}, 2 - \frac{Dt}{|\,\lambda\,|\,R}\right); \tag{36}$$

b. *if the numerical range of T is viewed from λ under an angle of amplitude $\Omega < \pi$, and if ω_0 is the bisector direction, then for any θ in the interval*

$$\left(\arg\lambda + \pi - \omega_0 - \frac{\pi - \Omega}{2}, \arg\lambda + \pi - \omega_0 + \frac{\pi - \Omega}{2}\right)$$

there are α's in the disc (21) *such that*

$$\arg\alpha = \theta\,, \qquad |\,\gamma\,| \leqq \sin\left(|\,\theta - (\pi - \omega_0) - \arg\lambda\,| + \frac{\Omega}{2}\right); \tag{37}$$

c. *if the numerical range is both bounded and viewed from λ under an angle of amplitude $\Omega < \pi$, the area between the open discs*

$$\left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 - \Omega/2)}\right| < \frac{|\,\lambda\,|\,R}{D}\,, \qquad \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 + \Omega/2)}\right| < \frac{|\,\lambda\,|\,R}{D} \tag{38}$$

is contained in the disc (21), *and for any α therein*

$$|\,\gamma\,|^2 \leqq 1 - \frac{dD}{R^2\,|\,\lambda\,|^2} \min\left\{\frac{|\,\lambda\,|^2 R^2}{D^2} - \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 - \Omega/2)}\right|^2, \frac{|\,\lambda\,|^2 R^2}{D^2} - \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 + \Omega/2)}\right|^2\right\}. \tag{39}$$

Proof. a. If the range is bounded $D(\lambda, T) < \infty$, and since $R\lambda(\gamma - 1)/\alpha$ is the difference between λ and a point in $\mathscr{N}(T)$,

$$\frac{d(\lambda, T)}{R\,|\,\lambda\,|} \leqq \frac{|\,\gamma - 1\,|}{|\,\alpha\,|} \leqq \frac{D(\lambda, T)}{R\,|\,\lambda\,|}\,. \tag{40}$$

It follows that if α maps into $\gamma = \rho$, with ρ real and $0 < \rho < 1$, then $|\,\alpha\,| \leqq (1 - \rho)R\,|\,\lambda\,|/d$, while if it maps into $\gamma = -\rho$, $|\,\alpha\,| \geqq (1 + \rho)R\,|\,\lambda\,|/D$. Therefore by continuity, if $(1 - \rho)R\,|\,\lambda\,|/d \leqq (1 + \rho)R\,|\,\lambda\,|/D$, that is if $(D - d)(D + d) \leqq \rho < 1$, the pre-image of $|\,\gamma\,| \leqq \rho$ contains α's whose moduli fill the closed interval $I_\rho = ((1 - \rho)R\,|\,\lambda\,|/d, (1 + \rho)R\,|\,\lambda\,|/D)$, which as ρ approaches 1 grows to the limiting open interval $(0, 2R\,|\,\lambda\,|/D)$. The proof is concluded by remarking that for any t in this interval

the smallest ρ such that $t \in I_\rho$ is precisely given by the right hand member of inequality (36).

b. By definition of Ω and ω_0

$$\omega_0 - \frac{\Omega}{2} \leqq \arg\left| \frac{(Tx_1 - Tx_2, x_1 - x_2)}{\| x_1 - x_2 \|^2} - \lambda \right| \leqq \omega_0 + \frac{\Omega}{2}$$

and so, recalling the meaning of $R\lambda(\gamma - 1)/\alpha$,

$$\omega_0 - \arg\lambda - \frac{\Omega}{2} \leqq \arg\left(\frac{\gamma - 1}{\alpha}\right) \leqq \omega_0 - \arg\lambda + \frac{\Omega}{2}\,. \tag{41}$$

As γ varies over the disc $|\gamma| \leqq \rho$, $\arg(\gamma - 1)$ varies over the closed interval $(\pi - \arcsin\rho, \pi + \arcsin\rho)$. Therefore by (41) the pre-image of the disc $|\gamma| \leqq \rho$ contains an α with

$$\arg\alpha \geqq \pi - \omega_0 + \arg\lambda + \arcsin\rho - \frac{\Omega}{2}$$

and an α with

$$\arg\alpha \leqq \pi - \omega_0 + \arg\lambda - \arcsin\rho + \frac{\Omega}{2}\,,$$

and in consequence, by continuity, contains α's whose arguments cover the closed interval

$$J_\rho = \left(\pi - \omega_0 + \arg\lambda - \arcsin\rho + \frac{\Omega}{2},\ \pi - \omega_0 + \arg\lambda + \arcsin\rho - \frac{\Omega}{2}\right),$$

provided $\arcsin\rho \geqq \Omega/2$. As $\rho \nearrow 1$ this interval becomes

$$\left(\pi - \omega_0 + \arg\lambda - \frac{\pi - \Omega}{2},\ \pi - \omega_0 + \arg\lambda + \frac{\pi - \Omega}{2}\right),$$

and for any θ in it the smallest ρ such that $\theta \in J_\rho$ is

$$\rho \sin\left(|\theta - (\pi - \omega_0) - \arg\lambda| + \frac{\Omega}{2}\right),$$

in agreement with (37).

c. In this case both (40) and (41) hold, and if α belongs to the two discs (38),

$$\begin{aligned} 0 < \frac{|\lambda|^2 R^2}{D^2} - \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 \pm \Omega/2)}\right|^2 &= \frac{2R}{D} Re(\alpha\bar{\lambda} e^{i(\omega_0 \pm \Omega/2)}) - |\alpha|^2 \\ &= |\alpha| \left| \frac{2R}{D} |\lambda| \cos\left(\arg\alpha - \left(\arg\lambda - \omega_0 \pm \frac{\Omega}{2}\right)\right) - |\alpha| \right| \end{aligned} \tag{42}$$

whence one deduces first $|\arg\alpha - (\arg\lambda - \omega_0 \pm (\Omega/2))| < \pi/2$, and then by (41),

$$-\frac{\pi}{2} < \arg\alpha - \arg\lambda + \omega_0 - \frac{\Omega}{2} \leqq \arg(\gamma - 1)$$
$$\leqq \arg\alpha - \arg\lambda + \omega_0 + \frac{\Omega}{2} < \frac{\pi}{2},$$

which, as $\cos x$ is a concave function in the interval $(-\pi/2, \pi/2)$, yields

$$0 < \min\left\{\cos\left(\arg\alpha - \arg\lambda + \omega_0 - \frac{\Omega}{2}\right), \cos\left(\arg\alpha - \arg\lambda + \omega_0 + \frac{\Omega}{2}\right)\right\}$$
$$\leqq \cos(\arg(\gamma - 1)).$$

Therefore taking the minimum in (42),

$$0 < \min\left\{\left[\frac{|\lambda|^2 R^2}{D^2} - \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 + \Omega/2)}\right|^2\right],\right.$$
$$\left.\left[\frac{|\lambda|^2 R^2}{D^2} - \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 - \Omega/2)}\right|^2\right]\right\}$$
$$\leqq |\alpha|\left[\frac{2R|\lambda|}{D}\cos(\arg(\gamma - 1)) - |\alpha|\right].$$

The right hand member of this inequality does no decrease if $|\alpha|$ outside the square brackets is replaced by its upper bound $(R|\lambda|/d)|\gamma - 1|$, and the $|\alpha|$ inside by its lower bound $(R|\lambda|/D)|\gamma - 1|$, both obtained from (40). Thus

$$0 < \min\left\{\left[\frac{|\lambda|^2 R^2}{D^2} - \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 + \Omega/2)}\right|^2\right],\right.$$
$$\left.\left[\frac{|\lambda|^2 R^2}{D^2} - \left|\alpha - \frac{\lambda R}{D} e^{-i(\omega_0 - \Omega/2)}\right|^2\right]\right\}$$
$$\leqq \frac{R|\lambda|}{d}|\gamma - 1|\left[\frac{2R|\lambda|}{D}\cos(\arg(\gamma - 1)) - \frac{R|\lambda|}{D}|\gamma - 1|\right] \tag{43}$$
$$= \frac{R^2|\lambda|^2}{dD}(1 - |\gamma|^2).$$

This is (39). It remains to check that the whole area between the two discs (38) is contained in the disc (21). In the first place it is clear that there are points simultaneously interior to the three discs, because the first two—which have a nonempty intersection—pass through the origin, whereas the third contains in its interior. Secondly, by virture of (43) and Lemma 1, no point interior to the area in equation belongs to the boundary of disc (21). Therefore, the area being connected is entirely contained in the disc (21).

Upon observing that the ranges of values assigned by a., b., and c. to $|\alpha|$, $\arg\alpha$, and α respectively, as well as the bounds for γ do not depend on y nor on the admissible approximation x_0, one may draw

from Lemma 3 the following interesting conclusion.

THEOREM VII. *If, under the hypotheses of either Theorem* V *or Theorem* VI, *the numerical range of* T *is either bounded, or viewed from* λ *under an angle of amplitude less than* π, *or both, then it is possible to construct geometrically convergent averaging schemes with averaging factors of constant modulus in the first case, of constant argument in the second, and with constant averaging factors in the third.*

It is to be noticed that under the cross-Lipfchitz condition, the requirement that $\mathcal{N}(T)$ be bounded amounts to asking that T be Lipschitzian. In fact, if $D(0, T)$ is the supremum of the distances to the origin of points in $\mathcal{N}(T)$, then $\| T \|^2 \leqq D^2(0, T) + (\| T \|^{\perp})^2$. On the other hand, the condition that $\mathcal{N}(T)$ be viewed from λ under an angle of amplitude less than π can be expressed by requiring that the convex hull of $\mathcal{N}(T)$ contain no straight line and that λ be at a positive distance from it.

Point c. of Lemma 3 raises the following natural question: Is the contractiveness of the operator $(1 - \alpha)I + \alpha\lambda^{-1}(T - y)$ the reason why, for α's satisfying (48), the solution of (1) can be obtained by iteration? In other words, is $(1 - \alpha)I + \alpha\lambda^{-1}(T - y)$ contractive for such α's? The answer is yes, and can be proved through calculations very similar to those leading to c.; the contraction factor turns out to be not larger than

$$1 - \frac{dD}{R|\lambda|^2} \min \left\{ \frac{|\lambda|^2 R^2}{D^2} - \left| \alpha - \frac{\lambda R}{D} e^{-i(\omega_0 + \Omega/2)} \right|^2, \right.$$
$$\left. \frac{|\lambda|^2 R^2}{D^2} - \left| \alpha - \frac{\lambda R}{D} e^{-i(\omega_0 - \Omega/2)} \right|^2 \right\}.$$

A related question is that of the asymptotic behavior of the sequence of averaging factors. When may the sequence terminate with some sort of constancy? It is evident that the answer must involve only the local behavior of the operator, for, if a sequence of averages converges, it finally dwells within an arbitrarily small neighborhood of the solution and thus involves only the values of the operator in such a neighborhood. Therefore, if any of the conditions of Theorem VII is valid locally the corresponding assertion holds asymptotically. This applies for instance to locally Lipschitzian mappings, in which case the averaging factors can be taken as having a constant modulus from an index on, or to continuously differentiable ones, which having local numerical ranges arbitrarily close to the numerical ranges of linear mappings make asymptotically constant averaging schemes possible.

Let us finally remark that the value $\lambda = 0$, excluded from our considerations, can be brought within the range of the averaging theory by replacing equation $0 = Tx - y$ by its equivalent $x = (T + I)x - y$.

The theory offered here is a development beyond the closely related theory of monotone operators started by the author [23] and continued by C. L. Dolph [13], G. J. Minty [17]-[21], I. I. Kolodner [16] and F. E. Browder [4]-[11], which can be summarized in the statement "the closure of the convex hull of the numerical range contains the spectrum".

BIBLIOGRAPHY

1. P. A. Alexandroff, *Combinatorial Topology*, Translation, Graylock Press, Rochester, N. Y., 1956.
2. G. Birkhoff, H. H. Goldstine and E. H. Zarantonello, *Calculations of plane cavity flows past curved obstacles*, Rend. Sem. Mat. Univ. Polytech di Torino **13** (1954), 205-244.
3. G. Birkhoff and E. H. Zarantonello, *Jets, Wakes, and Cavities*, Academic Press, New York, 1957.
4. F. E. Browder, *Solvability of nonlinear functional equations*, Duke Math. J. **30** (1963), 557-566.
5. ———, *Variational boundary value problems for quasilinear elliptic equations of arbitrary order*, Proc. Nat. Acad. Sci. **50** (1963), 31-37.
6. ———, *Variational boundary value problems for quasilinear elliptic equations, II*, Proc. Nat. Acad. Sci. **50** (1963), 592-598.
7. ———, *Variational boundary value problems for quasilinear elliptic equations, III*, Proc. Nat. Acad. Sci. **50** (1963), 794-798.
8. ———, *Nonlinear elliptic boundary value problems*, Bull. Amer. Math. Soc. **69** (1963), 862-874.
9. ———, *Nonlinear parabolic boundary value problems of arbitrary order*, Bull. Amer. Math. Soc. **69** (1963), 858-561.
10. ———, *Nonlinear elliptic problems, II*, Bull. Amer. Math. Soc. **70** (1964), 299-302.
11. ———, *Remarks on nonlinear functional equations*, Proc. Nat. Acad. Sci. **51** (1964), 985-989.
12. C. Chevalley, *Theory of Lie Groups*, Princeton University Press, 1946.
13. C. L. Dolph and G. J. Minty, *On non-linear equations of the Hammerstein type*, Proceeding Seminar Non-Linear Integral Equations, Univ. of Wis. Press, 1963.
14. F. Grünbaum and E. H. Zarantonello, *Cross-continuity vs. continuity*, Studia Mathematica **27** (1966), 273-288.
15. A. W. Kolmogorov and S. V. Fomin, *Elements of the Theory of Functions and Functional Analysis*, Translation, Graylock Press, Rochester, N. Y., 1957.
16. I. I. Kolodner, *Equations of Hammerstein Type in Hilbert Spaces*, Tech. Rep No. 47, Dept. of Math., Univ. of N. Mexico, November, 1963.
17. G. J. Minty, *On the maximal domain of a 'monotone' function*, Michigan Math. J. **8** (1961), 135-137.
18. ———, *Monotone (non-linear) operators in Hilbert space*, Duke Math. J. **29** (1962), 341-346.
19. ———, *On a 'monotonicity' method for the solution of non-linear equations in Banach spaces*, Proc. Nat. Acad. Sci. **50** (1963), 1038-1041.

20. ———, *Two theorems on nonlinear functional equations in Hilbert space*, Bull. Amer. Math. Soc. **69** (1963), 691-692.
21. ———, *On the montonicity of the gradient of a convex function*, Pacific J. Math. **14** (1964), 243-247.
22. M. H. Stone, *Linear transformations in Hilbert Space*, Amer. Math. Soc. Coll. Pub. XV, New York, 1932.
23. E. H. Zarantonello, *Solving functional equations by contractive averaging*, Math. Res. Center, Tech, Rep. No. 160, June 1960.
24. ———, *The closure of the numerical range contains the Spectrum*, Bull. Amer. Math. Soc. **70** (1964), 781-787.
25. ———, *The closure of the numerical range contains the spectrum*, Univ. of Kansas, Tech. Rept. **7** (new series), July 1964.

Received January 20, 1965. Research sponsored in part by National Science Foundation Grant GP-439. A summary of results appeared as a Research Announcement in the Bulletin of the American Mathematical Society [**24**]. See also [**25**].

UNIVERSIDAD NACIONAL DE CÓRDOBA, ARGENTINA
UNIVERSITY OF KANSAS

PACIFIC JOURNAL OF MATHEMATICS

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced). The first paragraph or two must be capable of being used separately as a synopsis of the entire paper. It should not contain references to the bibliography. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, Richard Arens at the University of California, Los Angeles, California 90024.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is published monthly. Effective with Volume 16 the price per volume (3 numbers) is $8.00; single issues, $3.00. Special price for current issues to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: $4.00 per volume; single issues $1.50. Back numbers are available.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 103 Highland Boulevard, Berkeley 8, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), 7-17, Fujimi 2-chome, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.

Pacific Journal of Mathematics
Vol. 22, No. 3 March, 1967