Existence and uniqueness regions
for solutions of nonlinear equations

A.L. Andrew

A refinement of the Newton-Kantorovich Theorem, which has many
potential applications in existence - uniqueness theory, is used
to strengthen a result of Lancaster and Rokne concerning
existence and uniqueness regions for zeros of operator
polynomials.

Most of the recent developments following from Kantorovich's analysis
of Newton's method [5] have been concerned with properties of numerical
algorithms. However Kantorovich's results have also proved useful for
establishing local existence and uniqueness results for solutions of non-
linear equations in Banach space. Special cases include differential
equations, integral equations, and finite systems of algebraic equations.

It is the purpose of this note to point out that some existence-
uniqueness results derived from Kantorovich's Theorem can be strengthened
by using Theorem 1 below. As an illustration, Theorem 1 is used to prove
two apparently new theorems concerning nonlinear operator equations which
strengthen results proved by Lancaster and Rokne using a classical "Newton-
Kantorovich" Theorem (Theorem 1 of [7]). *Henceforth the notation* "Theorem
7.n" *will denote Theorem number n of* [7]. Theorem 1 is an immediate
consequence of Theorem 1 of [2] although existence and uniqueness questions
were not emphasised in [2]. The solutions whose existence is established
in Theorems 2 and 3 may be computed by the modified Newton's method [2]
using the null operator as initial approximation. This may easily be
proved and error bounds for successive approximations established using

---

Received 21 August 1978.

results of [2] but this is not done here.

Theorem 3 is related to but not implied by a result deduced from the contraction mapping theorem by Eisenfeld [3, Theorem 2.1]. A comparison between the Kantorovich and contraction mapping approaches in a related context is made in [2]. The quantities

$$\left\|A_1^{-1}A_0\right\|^{m-1}\left\|A_1^{-1}A_m\right\|$$

occurring in Theorem 3 also arise in a related existence theorem of Isaev [4] which requires all these quantities to be less than a certain number which unfortunately is generally difficult to compute. Isaev did not give simple bounds for the solution or any uniqueness results. As explained in [3], [4], and [7], equation (2) below is important in the theory of eigen-value problems in which the eigenvalue parameter occurs nonlinearly. These problems, called nonlinear eigenparameter problems in [3], often arise in physics and engineering [1]. Solution of (2) is also required in some methods of numerical solution of nonlinear eigenparameter problems [6].

THEOREM 1. *Let* F *be a continuously Fréchet differentiable mapping of an open subset* Ω *of a Banach space* B *into* B. *Let* $\Gamma_0 = \left[F'(x_0)\right]^{-1}$, *the inverse of the Fréchet derivative of* F *at* x_0, *exist and be bounded, and let* $\Omega_0 = \left\{x \in B : \|x-x_0\| \le r\right\} \subset \Omega$. *Let*

$$\|I-\Gamma_0 F'(x)\| \le K\|x_0-x\| \quad for\ all \quad x \quad in \quad \Omega_0,$$

where I *is the identity operator. Let* $0 < h = K\|\Gamma_0 F(x_0)\| \le \frac{1}{2}$ *and* $r \ge r_-$ *where*

$$r_\pm = \left[1\pm(1-2h)^{\frac{1}{2}}\right]/K.$$

Then the equation $F(x) = 0$ *has a solution* x^* *in* $\Omega_- = \left\{x \in B : \|x-x_0\| \le r_-\right\} \subset \Omega_0$. *If also* $r < r_+$ *or* $r_- = r = r_+$ *then* x^* *is the only zero of* F *in* Ω_0.

Although the fact that $x^* \in \Omega_-$ is not explicitly stated in [2] it follows immediately on setting $r = r_-$ in Theorem 1 of [2].

THEOREM 2. *Let* Ω *be an open subset of* L, *the noncommutative*

Banach algebra of all bounded linear operators on a Banach space B , *with the operator norm. Let* $\Omega_0 = \{ X \in L : \|X\| \le r \} \subset \Omega$. *Let* $F : \Omega \to L$ *be defined by*

$$F(X) = A_0 + A_1 X + P(X) \ ,$$

where $A_0, A_1 \in L$, A_1^{-1} *exists, and* $P : \Omega \to L$ *satisfies*

(i) $A_1^{-1} P$ *is continuously Fréchet differentiable on* Ω ,

(ii) $P(0) = 0$,

(iii) $P'(0) = 0$,

(iv) *there exists* $\dot{\rho} \ge 0$ *such that* $\left\| A_1^{-1} P'(X) \right\| \le \rho \|X\|$, *for all* $X \in \Omega_0$.

Let $h_0 = \rho \left\| A_1^{-1} A_0 \right\| \le \tfrac{1}{2}$ *and let* $r \ge t_-$ *where*

$$t_{\pm} = \left[1 \pm \left(1 - 2h_0 \right)^{\frac{1}{2}} \right] / \rho \ .$$

Then F *has a zero in* $S = \{ X \in L : \|X\| \le t_- \}$ *and if* $r < t_+$ *or* $t_- = r = t_+$ *then that zero is unique in* Ω_0 .

The proof of Theorem 2 is identical with the proof of Theorem 7.7 in [7] except that Theorem 1 is used instead of Theorem 7.1.

THEOREM 3. *Let* $A_0, \ldots, A_n \in L$, $n \ge 2$, *where* L *is as in Theorem 2, and let* A_1^{-1} *exist. Let*

$$(1) \qquad \varepsilon = 1 - \sum_{m=2}^{n} 2^{m-1} m \left\| A_1^{-1} A_0 \right\|^{m-1} \left\| A_1^{-1} A_m \right\| \ge 0 \ .$$

Then the equation

$$(2) \qquad \sum_{m=0}^{n} A_m X^m = 0$$

has a solution in

$$S = \left\{ X \in L \ : \ \|X\| \leq 2 \left\| A_1^{-1} A_0 \right\| / \left(1 + \varepsilon^{\frac{1}{2}} \right) \right\}$$

which is unique in $\left\{ X \in L \ : \ \|X\| \leq 2 \left\| A_1^{-1} A_0 \right\| \right\}$.

Proof. In Theorem 2 take $P(X) = \sum_{m=2}^{n} A_m X^m$ and $r = 2 \left\| A_1^{-1} A_0 \right\|$. Then

$$(3) \qquad P'(X)H = \sum_{m=2}^{n} \sum_{i=1}^{m} A_m X^{i-1} H X^{m-i} \ .$$

Clearly P satisfies conditions *(i)*, *(ii)*, and *(iii)* of Theorem 2 and (3) shows that with the above choice of r it also satisfies *(iv)* with

$$(4) \qquad \rho = \sum_{m=2}^{n} 2^{m-2} m \left\| A_1^{-1} A_0 \right\|^{m-2} \left\| A_1^{-1} A_m \right\| \ .$$

Hence, by (1) and the definition of h_0 in Theorem 2, $1 - 2h_0 = \varepsilon \geq 0$, and hence

$$\rho^{-1} = 2 \left\| A_1^{-1} A_0 \right\| / (1 - \varepsilon) \ .$$

The result now follows from Theorem 2.

In general the existence result in Theorem 2 is strongest when $r = t_-$ and the uniqueness result is strongest when $r = t_+$. Hence Theorem 3 may be strengthened by choosing different values of r for the existence and the uniqueness part. However, except in the special case $n = 2$, when $P''(X)$ is independent of X and this modification yields Theorem 7.2, the result obtained by such "optimum" choices of r is more complicated and harder to apply than Theorem 3.

Theorem 3 strengthens Theorem 7.8, the only result in [7] dealing with (2) with $n > 2$, in three ways:

(i)   Theorem 7.8 considers only the case $n = 3$;

(ii)   the existence region S established for the solution by
      Theorem 3 is smaller than that established by Theorem 7.8;

(iii)   in the case $n = 3$ the inequality in (1) becomes

$$4 \left\| A_1^{-1} A_0 \right\| \left(3 \left\| A_1^{-1} A_3 \right\| \left\| A_1^{-1} A_0 \right\| + \left\| A_1^{-1} A_2 \right\| \right) \leq 1$$

whereas Theorem 7.8 requires the more restrictive condition

$$4\left\|A_1^{-1}A_0\right\|\left(6\left\|A_1^{-1}A_3\right\|\left(\left\|A_1^{-1}A_0\right\|+\left\|A_1^{-1}A_2\right\|\right)\right) < 1 \ .$$

This last improvement was made possible by the use of Theorem 1. The weakening of the Lipschitz conditions in Theorems 7.1 and 7.7 to those in Theorems 1 and 2 respectively, made possible the halving of the coefficient of $\left\|A_1^{-1}A_3\right\|$. The replacement of the strict inequality of [7] by weak inequality was possible since Theorem 1, following Kantorovich [5], considers separately the case $r_- = r_+$ (when $h = \frac{1}{2}$) whereas Theorem 7.1 does not.

Using Theorem 7.1 instead of Theorem 1, a marginally more complicated version of the above proof of Theorem 3 yields the similar but weaker result with ε replaced by

$$\varepsilon^* = 1 - \sum_{m=2}^{n} 2^{m-1}m(m-1)\left\|A_1^{-1}A_0\right\|^{m-1}\left\|A_1^{-1}A_m\right\| \ .$$

Comparing ε^* with ε shows that for all $n > 2$ $\left(\text{assuming } A_n \neq 0 \right)$, the use of Theorem 1 instead of Theorem 7.1 establishes local existence and uniqueness of a solution of (2) under more general conditions and even when $\varepsilon^* \geq 0$ it establishes a smaller existence region for the solution. The gain from using Theorem 1 increases as n increases.

Further information concerning (2), including cases when (1) is not satisfied, may often be obtained from Theorem 3 by means of the change of variable $Y = X - C$, where C is a fixed element of L chosen so that the new coefficients satisfy (1).

References

[1] A.L. Andrew, "Eigenvalue problems with nonlinear dependence on the eigenvalue parameter. A bibliography" (Technical Report, Department of Mathematics, La Trobe University, Bundoora, Victoria, 1974).

[2] A.L. Andrew, "Error bounds for the modified Newton's method", *Bull. Austral. Math. Soc.* 14 (1976), 427-433.

[3] J. Eisenfeld, "Operator equations and nonlinear eigenparameter
        problems", *J. Functional Analysis* 12 (1973), 475-490.

[4] G.A. Isaev, "The linear factorization of polynomial operator pencils",
        Math. Notes 13 (1973), 333-338.

[5] L.V. Kantorovich and G.P. Akilov, *Functional analysis in normed spaces*
        (translated by D.E. Brown. International Series of Monographs in
        Pure and Applied Mathematics, 46. Pergamon Press, Oxford,
        London, Edinburgh, New York, Paris, Frankfurt, 1964).

[6] P. Lancaster, "A review of numerical methods for eigenvalue problems
        nonlinear in the parameter", *Numerik und Anwendungen von
        Eigenwertaufgaben und Verzweigungsproblemen*, 43-67
        (Vortragsauszüge der Tagung über Numerik und Anwendungen von
        Eigenwertaufgaben und Verzweigungsproblemen, Oberwolfach,
        November 1976. International Series of Numerical Mathematics,
        38. Birkhäuser Verlag, Basel, Stuttgart, 1977).

[7] Peter Lancaster and Jon G. Rokne, "Solutions of nonlinear operator
        equations", *SIAM J. Math. Anal.* 8 (1977), 448-457.

Department of Mathematics,
La Trobe University,
Bundoora,
Victoria.