S. M. Roberts
J. S. Shipman

Some Results in Two-Point Boundary Value Problems

Abstract: Two new results in two-point boundary value problems are presented. The first is a modified method of adjoints which, under certain circumstances, will solve numerically two-point boundary value problems faster than the standard method of adjoints. The second result shows that Friedrichs' solution of the operator equation $P(x) = 0$ is really the modified Newton method. Kantorovich's sufficiency conditions for the convergence of the modified Newton's method are compared with Friedrichs' sufficiency conditions; it appears that, for most applications, the former conditions allow more leeway.

1. Introduction

Many problems in optimal control theory eventually reduce to the numerical solution of two-point boundary value problems for non-linear ordinary differential equations. In view of this and in view of their inherent interest, we present here two results on the solution of two-point boundary value problems.

In order to present our findings briefly, we assume that the reader is familiar with the work of Goodman and Lance,[1] Roberts and Shipman,[2] Kantorovich,[3] and Friedrichs.[4]

Kantorovich[3] has developed the sufficient conditions for the convergence of the Newton method and the modified Newton method for the solution of the operator equation $P(x) = 0$ by the successive approximations $x_{n+1} = x_n - (P'_{x_n})^{-1} P(x_n)$. In the Newton method the operator P'_{x_n} is evaluated for each step of the process $n = 0, 1, 2, \ldots$, while in the modified Newton method the operator P'_{x_n} is evaluated once at the initial guess x_0 (and thus really should be written P'_{x_0}) and is used for all iterations. The equivalence of the Goodman and Lance method of adjoints and the Newton method for the solution of two-point boundary value problems shown by Roberts and Shipman[2] suggests a modified method in which the Jacobian matrix equivalent to P'_{x_n} is evaluated once at the initial guess x_0 and is used for all iterations.

Our first result, presented in Section 2, includes some numerical experiments with the modified method of adjoints. The rate of convergence and the computer time required for the modified method of adjoints was compared with that for the method of adjoints, and a comparison of the actual error with the error predicted by Kantorovich's theorem was also made. As expected, the modified method required more iteration steps, but less computer time per iteration. In our example the total computer time required by the modified method of adjoints was greater than the time required by the Goodman and Lance method of adjoints; this could be reversed for a different problem. Kantorovich's error estimate for the modified Newton method gave order of magitude agreement with the actual error, but not quite as good as we experienced in our earlier experiments with the Goodman and Lance method of adjoints.

Friedrichs solves the operator equation $P(x) = 0$ by forming a new operator, which turns out to be a contraction mapping under his assumptions. Accordingly we were interested to see whether this, too, led to a new method for the numerical solution of two-point boundary value problems. Our second result, Section 3, proves that Friedrichs' technique is in fact the modified Newton method. In particular we demonstrate that Friedrichs' assumptions can be subsumed under the Kantorovich sufficiency theorem for the modified Newton method; that is to say that Kantorovich's theorem implies Friedrichs'.

2. Method of adjoints and the modified method of adjoints

Let us first recall the two ways of carrying out the solution of the operator equation

$$P(x) = 0$$

by the Newton method where P is the mapping from a Banach space X to a Banach space Y. The usual Newton method is

$$x_{n+1} = x_n - (P'_{x_n})^{-1} P(x_n),$$

Table 1 Example in which method of adjoints coverages but modified method of adjoints does not.

A. Method of Adjoints Iteration No.	h_i	Kantorovich Norms β_i	η_i	K_i
0	$6.763579(10^{2})$	3.659561	1.069325	$1.728374(10^{2})$
1	$6.536509(10^{1})$	3.444090	1.663449	$1.140937(10^{1})$
2	$5.233003(10^{-1})$	$7.149846(10^{-1})$	$2.365853(10^{-1})$	3.093615
3	$8.472580(10^{-1})$	1.124802	$3.827557(10^{-2})$	$1.967965(10^{1})$
4	$3.273678(10^{-2})$	1.017268	$2.250358(10^{-3})$	$1.430042(10^{1})$
5	$4.037952(10^{-6})$	1.013542	$2.822407(10^{-6})$	$1.411560(10^{1})$
6	$4.537566(10^{-8})$	1.013534	$3.171723(10^{-9})$	$1.411526(10^{1})$

	Calculated Missing Initial Conditions $\dot{x}(0)$	$\dot{y}(0)$	$\dot{z}(0)$	Calculated Terminal Conditions $x(2)$	$y(2)$	$z(2)$
0	$-5.379999(10^{-1})$	$2.879999(10^{-1})$	$4.988299(10^{-1})$	$6.086336(10^{-1})$	$-3.733722(10^{-1})$	$-6.466988(10^{-1})$
1	-1.376313	$5.242722(10^{-1})$	$9.080641(10^{-1})$	-2.218209	$-1.837729(10^{-1})$	$-3.183034(10^{-1})$
2	$3.236898(10^{-1})$	$5.779965(10^{-1})$	1.001111	$7.165604(10^{-1})$	$9.196269(10^{-1})$	1.592828
3	$8.720452(10^{-2})$	4.489073	$7.775289(10^{-1})$	$-8.591061(10^{-2})$	$5.013211(10^{-1})$	$8.683122(10^{-1})$
4	$1.017630(10^{-1})$	$4.709831(10^{-1})$	$8.157667(10^{-1})$	$-2.289911(10^{-3})$	$5.726597(10^{-1})$	$9.918756(10^{-1})$
5	$1.016559(10^{-1})$	$4.722822(10^{-1})$	$8.180167(10^{-1})$	$-9.427341(10^{-6})$	$5.759951(10^{-1})$	$9.976526(10^{-1})$
6	$1.016587(10^{-1})$	$4.722833(10^{-1})$	$8.180185(10^{-1})$	$-7.212622(10^{-9})$	$5.759999(10^{-1})$	$9.976609(10^{-1})$

B. Modified Method of Adjoints Iteration No.	h_i	Kantorovich Norms β_i	η_i	K_i
0	$6.763579(10^{2})$	3.659561	1.069325	$1.728374(10^{2})$
1	$6.536509(10^{1})$	3.444090	1.663449	$1.140937(10^{1})$
2	$3.397647(10^{2})$	$2.991697(10^{-2})$	$8.161847(10^{-2})$	$1.391465(10^{5})$
3	$1.484853(10^{-1})$	$5.182107(10^{-1})$	3.235824	$8.855075(10^{-2})$
4	$2.688236(10^{1})$	2.106461	2.846395	4.483515
5	1.662440	$6.269471(10^{-1})$	1.432732	1.850759
6	$3.217605(10^{1})$	$2.340128(10^{-2})$	$6.246358(10^{-2})$	$2.20123354(10)^{4}$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
20	$3.275644(10^{-2})$	$5.000002(10^{-1})$	$4.869078(10^{8})$	$1.345487(10^{-5})$

	Calculated Missing Initial Conditions $\dot{x}(0)$	$\dot{y}(0)$	$\dot{z}(0)$	Calculated Terminal Conditions $x(2)$	$y(2)$	$z(2)$
0	$-5.3799999(10^{-1})$	$2.8799999(10^{-1})$	$4.9882999(10^{-1})$	$6.0863367(10^{-1})$	$-3.7337224(10^{-1})$	$-6.4669886(10^{-1})$
1	-1.3763137	$5.2427226(10^{-1})$	$9.0806412(10^{-1})$	-2.2182091	$-1.8377300(10^{-1})$	$-3.1830344(10^{-1})$
2	$-9.7733935(10^{-2})$	$8.1932398(10^{-2})$	$1.4190935(10^{-1})$	-7.0690255	-4.2945087	-7.4382167
3	2.9456488	-1.0847001	-1.8786805	6.5937129	-2.1226356	-3.6763658
4	-2.9420842	$7.2204478(10^{-1})$	1.2508203	-5.2512476	$3.9822786(10^{-1})$	$6.8986231(10^{-1})$
5	$7.6721351(10^{-1})$	$-4.8602507(10^{-1})$	$-8.4151227(10^{-1})$	1.7326841	$-8.4152145(10^{-1})$	-1.4570249
6	-1.0781199	$5.8827505(10^{-2})$	$1.0269023(10^{-1})$	$-1.3777060(10^{1})$	-3.6528749	-6.3765168
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
20	-1.2999610	$-4.8687901(10^{8})$	$2.8085323(10^{8})$	-1.5527118	$-9.7375762(10^{8})$	$5.6170623(10^{8})$

in which the inverse of the derivative operator is reevaluated at each step of the process at x_n. The modified Newton method is

$$x_{n+1} = x_n - (P'_{x_0})^{-1} P(x_n),$$

in which the inverse of the derivative operator is evaluated at the point x_0 and is used for all iterations. In an earlier paper[2] the authors have shown that the Goodman-Lance method of adjoints (MA) applied to two-point boundary value problems with non-linear ordinary differential equations is actually the Newton method, whose convergence Kantorovich has studied. This identification suggests the idea of a modified method of adjoints (MMA), which, since it is a form of the modified Newton method, requires evaluation of the inverse of the derivative operator only once.

MMA has the potential practical advantage of faster computation since the adjoint equations are integrated only once and since the Jacobian matrix corresponding to P'_{x_0} (based on the initial vector x_0) is inverted only once and is used for all iterations. For the same order of accuracy attained, the relative merits of MA and MMA depend on the execution time, which is the product of two factors: (1) the time per iteration and (2) the total number of iterations. MMA (as a modified Newton method) requires more iterations than MA to achieve the same accuracy. When the initial vector x_0 is such that the sufficiency conditions of Kantorovich's theorem are fulfilled, both

Table 2 Example in which both the method of adjoints and the modified method of adjoints converge. Same boundary conditions as for example in Table 1.

A. Method of Adjoints Iteration No.	h_i	β_i	*Kantorovich Norms* η_i	K_i
0	$4.2229972(10^{-1})$	1.0824180	$2.2228302(10^{-2})$	$1.7551712(10^{1})$
1	$1.9874074(10^{-2})$	1.0151739	$1.3795129(10^{-3})$	$1.4191250(10^{1})$
2	$1.8627939(10^{-5})$	1.0135362	$1.3020693(10^{-6})$	$1.4115341(10^{1})$
3	$2.2818950(10^{-8})$	1.0135348	$1.5950266(10^{-9})$	$1.4115265(10^{1})$
4	$1.8159086(10^{-11})$	1.0135348	$1.2693059(10^{-12})$	$1.4115263(10^{1})$
5	$1.5110349(10^{-14})$	1.0135348	$1.0562014(10^{-15})$	$1.4115265(10^{1})$

	Calculated Missing Initial Conditions $\dot{x}(0)$	$\dot{y}(0)$	$\dot{z}(0)$	*Calculated Terminal Conditions* $x(2)$	$y(2)$	$z(2)$
0	$8.6999999(10^{-2})$	$4.4999999(10^{-1})$	$7.9999999(10^{-1})$	$-6.5145038(10^{-2})$	$5.1882565(10^{-1})$	$9.2235671(10^{-1})$
1	$1.0182591(10^{-1})$	$4.7222478(10^{-1})$	$8.1663938(10^{-1})$	$-7.6940894(10^{-4})$	$5.7518715(10^{-1})$	$9.9469680(10^{-1})$
2	$1.0165788(10^{-1})$	$4.7228206(10^{-1})$	$8.1801915(10^{-1})$	$-2.5203506(10^{-6})$	$5.7599766(10^{-1})$	$9.9766041(10^{-1})$
3	$1.0165880(10^{-1})$	$4.7228336(10^{-1})$	$8.1801856(10^{-1})$	$-1.4735255(10^{-9})$	$5.7599999(10^{-1})$	$9.9766099(10^{-1})$
4	$1.0165880(10^{-1})$	$4.7228336(10^{-1})$	$8.1801856(10^{-1})$	$-1.8980234(10^{-12})$	$5.7599999(10^{-1})$	$9.9766099(10^{-1})$
5	$1.0165880(10^{-1})$	$4.7228336(10^{-1})$	$8.1801856(10^{-1})$	$-1.1379786(10^{-15})$	$5.7599999(10^{-1})$	$9.9766099(10^{-1})$

B. Modified Method of Adjoints Iteration No.	h_i	β_i	*Kantorovich Norms* η_i	K_i
0	$4.2229972(10^{-1})$	1.0824180	$2.2228302(10^{-2})$	$1.7551712(10^{1})$
1	$1.9874074(10^{-2})$	1.0151739	$1.3795129(10^{-3})$	$1.4191250(10^{1})$
2	$5.9124365(10^{-4})$	1.0136209	$4.1311284(10^{-5})$	$1.4119594(10^{1})$
3	$5.2390109(10^{-5})$	1.0135392	$3.6619598(10^{-6})$	$1.4115463(10^{1})$
4	$1.9951430(10^{-6})$	1.0135350	$1.3945877(10^{-7})$	$1.4115277(10^{1})$
5	$1.4947761(10^{-7})$	1.0135348	$1.0448366(10^{-8})$	$1.4115266(10^{1})$
6	$6.6473046(10^{-9})$	1.0135348	$4.6464135(10^{-10})$	$1.4115265(10^{1})$
7	$4.4217505(10^{-10})$	1.0135348	$3.0907688(10^{-11})$	$1.4115264(10^{1})$

	Calculated Missing Initial Conditions $\dot{x}(0)$	$\dot{y}(0)$	$\dot{z}(0)$	*Calculated Terminal Conditions* $x(2)$	$y(2)$	$z(2)$
0	$8.6999999(10^{-2})$	$4.4999999(10^{-1})$	$7.9999999(10^{-1})$	$-6.5145038(10^{-2})$	$5.1882565(10^{-1})$	$9.2235671(10^{-1})$
1	$1.0182591(10^{-1})$	$4.7222478(10^{-1})$	$8.1663938(10^{-1})$	$-7.6940894(10^{-4})$	$5.7518715(10^{-1})$	$9.9469680(10^{-1})$
2	$1.0164567(10^{-1})$	$4.7224205(10^{-1})$	$8.1798921(10^{-1})$	$-7.9355623(10^{-5})$	$5.7590793(10^{-1})$	$9.9755301(10^{-1})$
3	$1.0165895(10^{-1})$	$4.7228369(10^{-1})$	$8.1801490(10^{-1})$	$-2.5247112(10^{-6})$	$5.7599841(10^{-1})$	$9.9765308(10^{-1})$
4	$1.0165877(10^{-1})$	$4.7228322(10^{-1})$	$8.1801847(10^{-1})$	$-2.0419541(10^{-7})$	$5.7599971(10^{-1})$	$9.9766070(10^{-1})$
5	$1.0165880(10^{-1})$	$4.7228336(10^{-1})$	$8.1801854(10^{-1})$	$-7.1454555(10^{-9})$	$5.7599999(10^{-1})$	$9.9766097(10^{-1})$
6	$1.0165880(10^{-1})$	$4.7228336(10^{-1})$	$8.1801856(10^{-1})$	$-5.5531839(10^{-10})$	$5.7599999(10^{-1})$	$9.9766099(10^{-1})$
7	$1.0165880(10^{-1})$	$4.7228336(10^{-1})$	$8.1801856(10^{-1})$	$-1.9003882(10^{-11})$	$5.7599999(10^{-1})$	$9.9766099(10^{-1})$

methods will converge, so that total computing time is the measure of superiority of one method over the other. If, as frequently occurs in practice, x_0 is such that Kantorovich's conditions are not fulfilled, then it is felt that it is better to use MA rather than MMA. There are, however, circumstances which do call for the modified method. For example, if the matrix P'_{z_n} is near-singular or large, it may very well pay to employ MMA. For large problems one may want to use mixed strategy in which MA is used at first and then once elements of the Jacobian matrix change only moderately from iteration to iteration, the program switches over to MMA.

To illustrate more or less typical behavior that might be encountered on comparing MA and MMA, let us consider the two-body equations

$$\ddot{x} = -\frac{kx(t)}{r^3}, \qquad \ddot{y} = -\frac{ky(t)}{r^3}, \qquad \ddot{z} = -\frac{kz(t)}{r^3},$$

where $r = (x^2 + y^2 + z^2)^{1/2}$ and $k = 1.0$ for canonical units.

The data of the problem are:

$$t_0 = 0 \qquad t_f = 2$$
$$x(0) = 1.076000 \qquad x(2) = 0.000000$$
$$y(0) = 0.000000 \qquad y(2) = 0.576000$$
$$z(0) = 0.000000 \qquad z(2) = 0.997661.$$

Table 1A illustrates a case where MA converges even though the initial trial vector and the next few iterates did not satisfy the sufficiency conditions of the Kantorovich theorem. Table 1B illustrates for the same initial conditions as Table 1A that MMA does not converge even after 20 iterations. We therefore have an example where MA works but MMA does not.

In the example given in Table 2A and 2B (MA and MMA,), both methods converge. MA converges in about

Table 3 Comparison of theoretical and computed error bounds on the initial velocity vector.* The bounds are given by the norm of the initial velocity vector, which is defined as Max $\{|\dot{x}_c - \dot{x}_l|, |\dot{y}_c - \dot{y}_l|, |\dot{z}_c - \dot{z}_l|\}$, where subscript c means calculated initial value and subscript l means true velocity. The "true" values were obtained from the tenth iterate of the method of adjoints.

Method of Adjoints Iteration No.	*Norm of initial velocity vector*	
	Theoretical	*Computed*
0	$4.5(10^{-2})$	$2.2(10^{-2})$
1	$1.6(10^{-2})$	$1.4(10^{-3})$
2	$3.8(10^{-3})$	$1.3(10^{-6})$
3	$4.6(10^{-4})$	$1.6(10^{-9})$
4	$1.3(10^{-5})$	$6.0(10^{-13})$
5	$2.2(10^{-8})$	$1.0(10^{-16})$
Modified Method of Adjoints Iteration No.	*Norm of initial velocity vector*	
	Theoretical	*Computed*
0	$2.9(10^{-2})$	$2.2(10^{-2})$
1	$1.3(10^{-2})$	$1.4(10^{-3})$
2	$5.9(10^{-3})$	$4.0(10^{-5})$
3	$2.7(10^{-3})$	$3.7(10^{-7})$
4	$1.2(10^{-3})$	$1.4(10^{-7})$
5	$5.4(10^{-4})$	$1.0(10^{-8})$

* This table was computed from data in Table 2 by using an adjusted value for h_0. The h_0 in Table 2 was calculated from $K_0 = 1.7551712(10^1)$, which is too high compared to other point estimates for K given in the table. Since a better estimate of the true K is given by $1.4115(10^1)$, a corrected value for h_0 is given by $4.2229972(10^{-1})$ $(1.4115/1.7551712) = 3.4578(10^{-1})$. The value used for η_0 is $2.2228303(10^{-2})$.

one-half as many iterations as MMA, to the same order of accuracy. Based on timing both methods for a total of 11 iterations we find that MA takes 6.2 sec/iteration and MMA takes 4.9 sec/iteration. For this example we may conclude that the terminal conditions were satisfied to single precision accuracy in 3 iterations and 18.6 sec by MA and in 6 iterations and 29.4 sec by MMA. In this case MA procedure is superior to MMA.

We have compared the error bounds calculated from Kantorovich's formulas (Table 3) for MA and MMA with the errors actually observed. As expected, the estimated errors are too large. However Kantorovich's estimates are fairly sharp, especially if we update them by using later iterates as the initial vector x_0 with associated parameters h_0, η_0.

An interesting sidelight to the Newton method and the modified Newton method is the evaluation of the two error bounds of Kantorovich. The error bound for the Newton method is given by

$$\frac{||x^* - x_n||}{\eta_0} \leq \frac{1}{2^n h_0} (2h_0)^{2^n}$$

and the error bound estimate for the modified Newton method is given by

$$\frac{||x^* - x_n||}{\eta_0} \leq \frac{1}{h_0} (1 - \sqrt{1 - 2h_0})^{n+1},$$

where x_n is the n^{th} iterate, x^* is the value approached, and h_0, and η_0 are as defined in Ref. 3. While these bounds show that the Newton method converges faster than the modified Newton method in general, it is interesting that for $n = 0$ and $n = 1$ the second formula gives smaller estimates over a wide range of h_0 $(10^{-7} < h_0 < 0.4)$ than the first, although x_0 and x_1 are identical for the two methods.

3. Friedrich's solution of P(x) = 0

The preceding section indicates that, generally speaking, any method of approximating the solution to the abstract operator equation $P(x) = 0$ (where P maps the Banach space X into Banach space Y) may be expected to lead to a concrete method for solving the two-point boundary value problem once the proper identifications have been made.

Friedrichs[4] is among the authors whose treatment of the solution of the operator equation $P(x)$ appear to be different from the Newton-Kantorovich method, and we were therefore interested to see if his approach leads to another method of solving two-point boundary value problems. He assumes that $P(x)$ has a "first variation," that is, there is defined for each $x \in X$ a linear operator $P_1(x, h)$ from X to Y such that if

$$R(x, h) = \begin{cases} \dfrac{P(x + h) - P(x) - P_1(x, h)}{||h||} & \text{for } h \in X \neq 0 \\ 0 & \text{for } h = 0, \end{cases} \quad (1)$$

then $||R(x, h)|| \to 0$ as $||h|| \to 0$. Now if x_0 is an approximate solution of $P(x) = 0$, put $x = x_0 + \Delta x$. By (1)

$$P(x) = P(x_0) + P_1(x_0, \Delta x) + ||\Delta x|| R(x_0, \Delta x), \quad (2)$$

so the equation $P(x) = 0$ takes the form

$$P_1(x_0, \Delta x) = -P(x_0) - ||\Delta x|| R(x_0, \Delta x). \quad (3)$$

If the linear operator $P_1(x_0, \Delta x)$ has a bounded inverse $\Gamma_0 y \equiv \Gamma(x_0, y)$ (which is necessarily linear), that is, $\Gamma[x_0, P_1(x_0, \Delta x)] = \Delta x$ and $||\Gamma_0 y|| \leq ||\Gamma_0|| \, ||y||$, where the real number $||\Gamma_0|| > 0$ is a bound for Γ_0, then Γ_0 can be applied to Eq. (3) to give

$$\Delta x = -\Gamma_0 P(x_0) - ||\Delta x|| \, \Gamma_0 R(x_0, \Delta x). \quad (4)$$

But Eq. (4) is of the form $z = H(z)$, and for such equations it is known that the sequence $\{z_n\}$, where $z_{n+1} \equiv H(z_n)$, converges to z provided H is a contraction operator. Friedrichs then establishes conditions which guarantee that $H(z) = -\Gamma P(x_0) - ||z|| \, \Gamma R(x_0, z)$ be a contraction

operator with contraction constant θ, $0 < \theta < 1$. Under these conditions, the sequence $\{\Delta x_n\}$, where

$$\Delta x_{n+1} = -\Gamma_0 P(x_0) - ||\Delta x_n||\, \Gamma_0 R(x_0, \Delta x_n) \qquad (5)$$

converges to the solution Δx of Eq. (4), which in turn furnishes the solution $x = x_0 + \Delta x$ of $P(x) = 0$. Ficken[5] in his investigation of the continuation method for solving operator equations $P(x, \mu) = 0$ takes Friedrichs' approach as a point of departure.

Friedrichs' results may be summed up in the following theorem, which is a restatement of his Theorem 10.1.

Theorem

Let x_0 be an approximate solution of $P(x) = 0$, and suppose the following assumptions hold:

I. $P(x)$ has a first variation $P_1(x, \delta x)$ throughout a certain sphere $||x - x_0|| \leq A$, and for each $\epsilon < 0$ there exist numbers $\alpha(\epsilon, x_0) < 0$ and $\beta(\epsilon, x_0) < 0$ such that:

a. $||x - x_0|| \leq \alpha(\epsilon, x_0)$ and $||\delta x|| \leq \delta(\epsilon, x_0)$

imply

$||R(x, \delta x)|| \leq \epsilon$;

b. $||x - x_0|| \leq \beta(\epsilon, x_0)$

implies

$||P_1(x, \delta x) - P_1(x_0, \delta x)|| \leq \epsilon\, ||\delta x||$.

II. Further, $P_1(x_0, y)$ has a bounded inverse $\Gamma(x_0, y) = \Gamma_0$ with bound $||\Gamma_0|| = \mu(x_0) > 0$.

III. Finally, x_0 is such that

$$||P(x_0)|| \leq \frac{1 - \mu\epsilon}{\mu} \min\left(\alpha, \frac{\beta}{2}\right).$$

Choose θ such that $0 < \theta < 1$, and define $\epsilon = \theta/2\,\mu(x_0)$. If assumptions I and III hold for x_0 and this value of ϵ, then the equation $P(x) = 0$ has a unique solution in the sphere

$$||x - x_0|| \leq \min\,[\alpha(\epsilon, x_0), \tfrac{1}{2}\beta(\epsilon, x_0)],$$

to which the sequence $\{x_n\}$, where $x_n = x_0 + \Delta x_n$ and Δx_n is given by Eq. (5), converges.

However, although it is not explicitly stated in Ref. 4, it turns out that Friedrichs' method for solving $P(x)$ is simply the modified Newton method, as can be easily shown. If we recognize that in Eq. (1) the first variation $P_1(x_0, h)$ is the Frechet differential, we may write it in the more usual notation $P'(x_0)\, h$. Then, using (1)

$$||\Delta x_n||\, R(x_0, \Delta x_n) = P(x_0 + \Delta x_n) - P(x_0) - P'(x_0)\, \Delta x_n, \qquad (6)$$

which when substituted into (5) gives

$$\Delta x_{n+1} = \Delta x_n - \Gamma_0 P(x_0 + \Delta x_n), \qquad (7)$$

since $\Gamma_0\, P'(x_0)\, \Delta x_n = \Delta x_n$ by the definition of the operator Γ_0. Then adding x_0 to both sides of (7) and setting $x_0 + \Delta x_n = x_n$ we get the sequence $\{x_n\}$ where,

$$x_{n+1} = x_n - \Gamma_0 P(x_n). \qquad (8)$$

Equation (8) is the modified Newton method for solving the operator equation $P(x) = 0$.

A direct comparison between Kantorovich's theorem (Theorem 6 (1.XVIII) of Ref. 3) and Friedrichs' theorem on the convergence of $\{x_n\}$ to the solution x does not seem possible, since Kantorovich requires that the second derivative $P''(x)$ be uniformly bounded on a certain sphere Ω_0 containing the initial approximation x_0, while Friedrichs requires that the first derivative $P'(x)$ satisfy a Lipschitz condition in a (possibly different) sphere around x_0. However, since a bound on the derivative of a function implies a Lipschitz condition on the function itself, the bounded derivative requirement is often made in theory, and is sometimes the only practical way to guarantee that a function satisfy a Lipschitz condition. Therefore let us assume that the hypotheses of Kantorovich's theorem are satisfied and see what this implies with regard to Friedrichs' theorem. We have then that an x_0 has been found such that

$$(1') \qquad ||\Gamma_0|| \leq B',$$

$$(2') \qquad ||P(x_0)|| \leq \eta',$$

$$(3') \qquad ||P''(x)|| \leq K'$$

in a sphere $S(r_0, x_0)$ of radius $r_0 = 2B'\,\eta'$ around x_0, and $h = K'\,B'^2\,\eta \leq \frac{1}{2}$. Under these conditions the modified Newton method will converge to the solution x of $P(x) = 0$ which, moreover, will be in $S(r_0, x_0)$.

Turning to Friedrichs' theorem, assumption II is satisfied with $\mu(x_0) \equiv B'$. Assume a θ (the contraction constant) has been picked with $0 < \theta < 1$, which then defines $\epsilon = (\theta/2B')$. Assumption I is satisfied with $A = 2B'\eta'$, $\alpha(x_0, \epsilon) = A = 2B'\eta'$, and $\beta(x_0, \epsilon) = \epsilon/K'$. Here the bound on $P''(x)$ has been used to verify the Lipschitz condition on the first derivative, Ib. Finally assumption III requires that $\eta' \leq (2 - \theta/2B')\, \min\,(2B'\,\eta', \epsilon/2K')$. If $2B'\,\eta'$ is the minimum, the condition is essentially vacuous, reducing to $1 \leq 2 - \theta$ or $\theta \leq 1$. So we may assume that $\epsilon/2K'$ is the minimum, which leads to the requirement (since $\epsilon = \theta/2B'$):

$$\eta' \leq \frac{(2 - \theta)\theta}{8B'^2K'} \quad \text{or}$$

$$h = \eta' B'^2 K' \leq \frac{(2 - \theta)\theta}{8}.$$

This in turn implies that h be less than $1/8$ instead of $\frac{1}{2}$, as Kantorovich requires. Moreover if it can be assumed that K' and B' are the same in both theorems so that only the

value of η' is open, then the implication is that Friedrichs' theorem requires a better initial approximation x_0 in order to guarantee convergence.

Our conclusion is thus that Friedrichs' method is the modified Newton's method and that for those applications where a bounded second derivative may be assumed, the theorem of Kantorovich is stronger and more precise than the corresponding theorem of Friedrichs.

References

1. T. R. Goodman and G. N. Lance, "The Numerical Solution of Two-Point Boundary Value Problems," *Mathematical Tables and Other Aids to Computation* **10**, 82–86 (1956).
2. S. M. Roberts and J. S. Shipman, "The Kantorovich Theorem and Two-Point Boundary Value Problems," *IBM Journal* **10**, 402–406 (1966).
3. L. V. Kantorovich and G. P. Akilov, *Functional Analysis in Normed Spaces*, The Macmillan Company, New York, 1964.
4. K. O. Friedrichs, *Functional Analysis and Applications*, New York University, Institute of Mathematical Sciences, 1953.
5. F. A. Ficken, "The Continuation Method for Functional Equations," *Comm. Pure and Applied Math.* **4**, 435–456 (1951).
6. R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, (1960).

Received October 20, 1966