JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **103**, 162–171 (1984)

A Rapidly Convergent Iteration Method and Gâteaux Differentiable Operators*

WILLIAM O. RAY

Department of Mathematics, The University of Oklahoma, Norman, Oklahoma 73019

Submitted by V. Lakshmikantham

Let $\{F_r\}_{0<r<p}$ be a family of Banach spaces satisfying, if $0 \leqslant r_1 \leqslant r_2 \leqslant p$, (i) $F_{r_1} \supseteq F_{r_2}$; (ii) $|f|_{r_1} \leqslant |f|_{r_2}$ $(f \in F_{r_1})$; and (iii) $\varphi(r) = \ln(|f|_r)$ is a convex function. Let G_0 be a Banach space and $\mathscr{F}$ be a Gâteaux differentiable mapping, and suppose that $\mathscr{F}'(x)(F_p)$ is dense in G_0. Under appropriate assumptions, the equation $\mathscr{F}(x) = 0$ has a solution in F_r for $0 \leqslant r \leqslant p$. The results extend the Inverse Function Theorem of J. Moser to the class of Gâteaux differentiable operators.

Let X and Y be Banach spaces and $\mathscr{F}$ a (Frechet) differentiable operator from X to Y. When the operator $\mathscr{F}'(x)$ has a bounded inverse, Newton's method—the "rapidly converging" method of the title— provides an iteration scheme for solving the equation $\mathscr{F}(x) = 0$. The existence of such a solution, which essentially derives from the Contraction Mapping Principle, gives rise to the standard Inverse Function Theorem which has been of such profound importance in the study of nonlinear differential equations.

Despite the broad scope of applicability of these ideas, the requirement that $\mathscr{F}'(x)$ have a bounded inverse imposes a rather severe limitation. Indeed, it frequently happens that $\mathscr{F}$ (and hence $\mathscr{F}'(x)$) maps from a space of functions with n derivatives to a space of functions with $n - m$ derivatives, and thus, after a finite number of steps, $\mathscr{F}'(x)^{-1}$ may not even be defined. In order to circumvent such difficulties, J. Moser in 1966 [7] formulated an Inverse Function Theorem in the context of a scale of Banach spaces (for example, the Sobolev spaces $H_{2,r}$). Moser modified Newton's method, solving the equations $\mathscr{F}'(x)\,h = -\mathscr{F}(x)$ only approximately in order to preverve the smoothness of h, and he showed that Newton's method was sufficiently regular to still yield a solution to the equation $\mathscr{F}(x) = 0$.

In this paper we propose to extend Moser's now classic work in two ways. First, while Moser considered only continuously Frechet differentiable operators $\mathscr{F}$, we will allow $\mathscr{F}$ to be Gâteaux differentiable. Second, we

* Research supported in part by a grant from the University of Oklahoma Foundation.

0022-247X/84 $3.00

improve on some of Moser's technical assumptions, eliminating a quadratic estimate and obtaining solutions in a broader range of spaces.

Besides the above extensions, we feel that our methods are of independent interest. As in the earlier papers [8] and [9], we employ the following maximal principle of H. Brezis and F. E. Browder (Corollary 3 of [2]):

PROPOSITION B. *Let $(E, d, \leqslant)$ be a complete, partially ordered metric space and let $\psi: E \to [0, \infty)$ be an arbitrary function. Suppose*:

(1) *$\mathscr{S}(x) \equiv \{y \in E: y \geqslant x\}$ is closed for each $x \in E$*;

(2) *$x \leqslant y$ and $x \neq y$ imply $\psi(x) > \psi(y)$; and*

(3) *any nondecreasing sequence has compact closure. Then there is an $\bar{x} \in E$ for which $\mathscr{S}(\bar{x}) = \{\bar{x}\}$.*

Proposition B, which is independent of the axiom of choice, is general enough to embrace a range of other maximal principles, including Ekeland's famous theorem [4]. As with the Contraction Mapping Principle in the case of the standard Inverse Function Theorem, the Brezis–Browder principle provides a unified technique for proving such theorems for Gâteaux differentiable operators.

We now turn to our results. Our assumptions, which are extracted from [7], are rather technical in character and so we list them separately here before stating our results. Throughout, we assume that $\{(F_r, |\cdot|_r)\}_{0 \leqslant r \leqslant p}$ is a family of Banach spaces satisfying:

(4) for $0 \leqslant r_1 \leqslant r_2 \leqslant p$, $F_{r_1} \supseteq F_{r_2}$;

(5) for all $f \in F_p$, $|f|_0 \leqslant |f|_p$; and

(6) for each $r \in (0, p)$ there is a constant C_r such that

$$|f|_r \leqslant C_r |f|_0^{(p-r)/p} |f|_p^{r/p} \qquad (f \in F_p).$$

The survey article [5] describes a wide range of spaces satisfying the above conditions. Note that (4)–(6) imply that $|\cdot|_p$ imposes a finer topology on F_r than does $|\cdot|_r$. We will also assume that $(G_0, \|\cdot\|_0)$ and $(G_\sigma, \|\cdot\|_\sigma)$ are Banach spaces with $G_0 \supseteq G_\sigma$. We will consider nonlinear operators $\mathscr{F}$ from F_0 to G_σ. In the applications considered in [7], it usually is the case that $\mathscr{F}(F_p) \subseteq F_{p-1}$, so that the range spaces of $\mathscr{F}$ have the same structure as those of the domain; however, we will require no explicit relationship to exist between $\|\cdot\|_0$ and $\|\cdot\|_\sigma$.

Within the above framework, we make the following additional hypotheses.

(H1) $\mathscr{F}$ has closed graph in $F_r \times G_0$ $(0 \leqslant r \leqslant p)$.

(H2) $\mathscr{F}$ is *Gâteaux differentiable*. By this, we mean that, for each

$f \in F_p$, there exists a (possibly nonlinear) mapping $\mathscr{F}'(f): F_p \to G_0$ satisfying, for $h \in F_p$,

$$\lim_{\xi \to 0} \| \xi^{-1}(\mathscr{F}(f+\xi h) - \mathscr{F}(f)) - \mathscr{F}'(f)\, h \|_0 = 0.$$

(Note that, while $\mathscr{F}'(f)$ need not be linear, $\mathscr{F}'(f)$ is homogeneous, i.e., $\mathscr{F}'(f)\, \alpha h = \alpha \mathscr{F}'(f)\, h$ for all α.)

(H3) There is an $M \geqslant 1$ such that, for all $K \geqslant 1$, if $|f|_p \leqslant K$ then $\|\mathscr{F}(f)\|_\sigma \leqslant MK$.

(H4) There are constants $\mu > 0$ and $C > 0$ such that, for each $\varepsilon > 0$, for each $K > 0$, for each $g \in G_0$ and each $f \in F_p$ with $\| g\|_\sigma \leqslant K$ and $|f|_p \leqslant K$, there is an $f_\varepsilon \in F_p$ satisfying:

(a) $\|\mathscr{F}'(f)(f_\varepsilon) - g\|_0 \leqslant \varepsilon^\mu K \| g\|_0$;

(b) $|f_\varepsilon|_p \leqslant \varepsilon^{-1} K \| g\|_0$; and

(c) $|f_\varepsilon|_0 \leqslant C \|\mathscr{F}'(f)(f_\varepsilon)\|_0$.

If (H4) holds, then Moser says the equation $\mathscr{F}'(f)\, h = g$ admits approximate solutions of order μ. Note that (c) does not imply that $\mathscr{F}'(f)$ is injective, since (c) applies only to f_ε in the smaller subspace F_p.

Theorem 1. *Suppose hypotheses* (H1) *through* (H4) *hold, and suppose there is an $f_0 \in F_p$ for which*

$$\|\mathscr{F}(f_0)\|_0 < \mu^2 (1+\mu)^{-(1+\mu)/\mu} M^{-1} K_0^{-1/\mu}$$

where $K_0 = \max\{M |f_0|_p, 1\}$. If

$$0 \leqslant r < p(1+\mu)^{-1} \tag{7}$$

then the equation $\mathscr{F}(f) = 0$ has a solution in F_r.

If $\|\mathscr{F}(f_0)\|_0$ satisfies a slightly sharper estimate, we can obtain a solution for the remaining values of r:

Theorem 2. *Suppose hypotheses* (H1) *through* (H4) *hold and suppose there is an $f_0 \in F_p$ for which, for some $r \in (0, p]$,*

$$\|\mathscr{F}(f_0)\|_0 < \mu^2 p (1+\mu)^{-(2\mu+1)/\mu} (rM)^{-1} K_0^{-1/\mu}$$

where $K_0 = \max\{M |f_0|_p, 1\}$. If

$$p(1+\mu)^{-1} \leqslant r \leqslant p \tag{8}$$

then the equation $\mathscr{F}(f) = 0$ has a solution in F_r.

Proof of Theorem 1. Choose $\beta \in (0, 1)$ so small that

$$\|\mathscr{F}(f_0)\|_0 \leqslant \mu^2 \left(\frac{1-\beta}{1+\mu}\right)^{(1+\mu)/\mu} M^{-1} K_0^{-1/\mu}$$

$$= \mu \left(1 - \beta - \frac{1-\beta}{1+\mu}\right)\left(\frac{1-\beta}{1+\mu}\right)^{1/\mu} M^{-1} K_0^{-1/\mu}.$$

Then with $\lambda = (1-\beta)(1+\mu)^{-1}$ and $q = \beta + \lambda = (1+\beta\mu)(1+\mu)^{-1} < 1$, one sees that

$$M(K_0/\lambda)^{1/\mu} \|\mathscr{F}(f_0)\|_0 \leqslant \mu(1-q). \tag{9}$$

Next set $E = \{(f, s, K) \in F_p \times [0, \infty) \times [K_0, \infty): M|f|_p \leqslant K\}$ (note that $(f_0, 0, K_0) \in E$) and define a metric d on E by $d((f_1, s_1, K_1), (f_2, s_2, K_2)) = \max\{|s_1 - s_2|, |K_1 - K_2|, |f_1 - f_2|_p, \|\mathscr{F}(f_1) - \mathscr{F}(f_2)\|_0\}$. Since $\mathscr{F}$ has closed graph, (E, d) is a complete metric space. Define $\psi: E \to [0, \infty)$ by $\psi(f, s, K) = \|\mathscr{F}(f)\|_0$; again (H1) implies that ψ is continuous.

Next we fix r satisfying (7) and define an ordering "$\lesssim$" on E by saying $(f_1, s_1, K_1) \lesssim (f_2, s_2, K_2)$ if and only if

(10) $s_1 \leqslant s_2$;

(11) $\|\mathscr{F}(f_1) - \mathscr{F}(f_2)\|_0 \leqslant ((1+q)/(1-q))(\|\mathscr{F}(f_1)\|_0 - \|\mathscr{F}(f_2)\|_0)$;

(12) $\|\mathscr{F}(f_2)\|_0 \leqslant \|\mathscr{F}(f_1)\|_0 \exp(-(1-q)(s_2 - s_1))$;

(13) $K_2 = K_1 \exp(\mu(1-q)(s_2 - s_1))$;

(14) $|f_2 - f_1|_p \leqslant \lambda^{-1/\mu} \|\mathscr{F}(f_1)\|_0 K_1^{(\mu+1)/\mu} \int_{S_1}^{S_2} \exp(\mu(1-q)(t - s_1))\, dt$;

and

(15) $|f_2 - f_1|_r \leqslant A \|\mathscr{F}(f_1)\|_0 K_1^{(\mu+1)r/\mu p} \int_{S_1}^{S_2} \exp\{(t - s_1)(1-q) [(\mu+1) r/p - 1]\}\, dt$,

where $A > 0$ is a constant depending on r to be determined later.

It is clear that "$\lesssim$" is reflexive and anti-symmetric, and that the relations (10) through (13) are transitive as well. To see that (14) is transitive, we apply (12) and (13); suppose that $(f_1, s_1, K_1) \lesssim (f_2, s_2, K_2)$ and $(f_2, s_2, K_2) \lesssim (f_3, s_3, K_3)$. Then

$$|f_3 - f_2|_p \leqslant \int_{S_2}^{S_3} \lambda^{-1/\mu} \|\mathscr{F}(f_2)\|_0 K_2^{(\mu+1)/\mu} \exp\{\mu(1-q)(t - s_2)\}\, dt$$

$$\leqslant \int_{S_2}^{S_3} \lambda^{-1/\mu} \|\mathscr{F}(f_1)\|_0 e^{-(1-q)(s_2 - s_1)} K_1^{(\mu+1)/\mu}$$

$$e^{(\mu+1)(1-q)(s_2 - s_1)} \exp\{\mu(1-q)(t - s_2)\}\, dt$$

$$= \int_{S_2}^{S_3} \lambda^{-1/\mu} \|\mathscr{F}(f_1)\|_0 K_1^{(\mu+1)/\mu} \exp\{\mu(1-q)(t - s_1)\}\, dt.$$

Combining this with (14) and the triangle inequality gives that (14) is transitive. In exactly the same fashion one shows that (15) is transitive as well, and thus "$\lesssim$" is a partial order.

Now set $\tilde{E} = \mathscr{S}(f_0, 0, K_0) \equiv \{(f, s, K) \in E: (f, s, K) \gtrsim (f_0, 0, K_0)\}$. In view of (4) through (6), convergence in $|\cdot|_p$ implies convergence in $|\cdot|_r$, and thus it follows readily that $\mathscr{S}(f, s, K)$ is a closed subset of $\tilde{E}$ for each $(f, s, K) \in \tilde{E}$. Also, in view of (13) and (14), if $(f_1, s_1, K_1) \lesssim (f_2, s_2, K_2)$ and $(f_1, s_1, K_1) \neq (f_2, s_2, K_2)$, then $s_1 \neq s_2$, and thus, by (12), $\psi(f_1, s_1, K_1) > \psi(f_2, s_2, K_2)$. We will show that $\mathscr{F}(f) = 0$ has a solution for some $f \in F_r$ which is a limit (in F_r) of elements in $\tilde{E}$; suppose for contradiction, that $\mathscr{F}(f) \neq 0$ for all such f.

In order to apply the Brezis–Browder principle, all that remains is to show that, if $\{(f_n, s_n, K_n)\}$ is a nondecreasing sequence in $\tilde{E}$, then $\{(f_n, s_n, K_n)\}$ has compact closure. By (12) and (13)

$$\|\mathscr{F}(f_n)\|_0 \leqslant \|\mathscr{F}(f_0)\|_0 \, e^{-(1-q)s_n}$$

and

$$K_n = K_0 e^{\mu(1-q)s_n}.$$

Applying these relations to (15) gives, for $n < m$,

$$\begin{aligned} |f_n - f_m|_r \leqslant & A \, \|\mathscr{F}(f_0)\|_0 \, e^{-(1-q)s_n} K_0^{(\mu+1)r/\mu p} \\ & \exp(r(\mu+1)(1-q)\, s_n\, p^{-1}) \\ & \int_{s_n}^{s_m} \exp\left\{(t - s_n)(1-q)\left[\frac{(\mu+1)\,r}{p} - 1\right]\right\} dt \\ = & A \, \|\mathscr{F}(f_0)\|_0 \, K_0^{(\mu+1)r/\mu p} \\ & \int_{s_n}^{s_m} \exp\left\{t(1-q)\left[\frac{(\mu+1)\,r}{p} - 1\right]\right\} dt. \end{aligned}$$

By (10), $\{s_n\}$ is a nondecreasing sequence, and, by (7), $(\mu+1)\,rp^{-1} < 1$. Consequently, $\{f_n\}$ is a Cauchy sequence in F_r in both the cases $s_n \to \infty$ and $s_n \to s < \infty$ as $n \to \infty$; denote by f_∞ the limit of $\{f_n\}$ in F_r. By (12), $\{\|\mathscr{F}(f_n)\|_0\}$ is a nonincreasing sequence in $[0, \infty)$, and thus convergent; (11) then implies that $\{\mathscr{F}(f_n)\}$ is a Cauchy sequence in G_0. In view of (H1) and (12),

$$\|\mathscr{F}(f_\infty)\|_0 = \lim \|\mathscr{F}(f_n)\|_0 \leqslant \lim \|\mathscr{F}(f_0)\|_0 \, e^{-(1-q)s_n}.$$

Now if $\{s_n\}$ is unbounded, $\mathscr{F}(f_\infty) = 0$, contrary to our earlier assumption; hence $\{s_n\} \uparrow s_\infty < \infty$ as $n \to \infty$. This implies that $\{K_n\}$ converges to $K_\infty = K_0 e^{\mu(1-q)s_\infty}$ and that $\{f_n\}$ is a Cauchy sequence in F_p, i.e., $\{f_n\}$

converges in F_p to f_∞. We have shown that $\{(f_n, s_n, K_n)\}$ is a convergent sequence in $(\tilde{E}, d)$, more than was required.

Consequently, there is an $(\bar{f}, \bar{s}, \bar{K}) \in E$ for which $\mathscr{S}(\bar{f}, \bar{s}, \bar{K}) = \{(\bar{f}, \bar{s}, \bar{K})\}$; we will now use (H4) to derive a contradiction to this conclusion. In (H4), take $K = \bar{K}$, $\varepsilon = (\lambda/\bar{K})^{1/\mu}$, $g = -\mathscr{F}(\bar{f})$ and $f = \bar{f}$. (Since $\bar{K} \geqslant 1$ and $|\bar{f}|_p \leqslant \bar{K}M^{-1}$, it follows from (H3) that $\|-\mathscr{F}(\bar{f})\|_\sigma \leqslant \bar{K}$.) Corresponding to these choices, select $f_\varepsilon \in F_p$ so that (a), (b) and (c) of (H4) hold. Next, choose $\xi > 0$ so small that

$$\|\mathscr{F}(\bar{f} - \xi f_\varepsilon) - \mathscr{F}(\bar{f}) - \xi \mathscr{F}'(\bar{f})(f_\varepsilon)\|_0 \leqslant \xi\beta \|\mathscr{F}(\bar{f})\|_0 \tag{16}$$

where $\beta > 0$ is the number chosen in the first line of the proof. Set $(f^+, s^+, K^+) = (\bar{f} + \xi f_\varepsilon, \bar{s} + \xi, \bar{K}e^{\mu(1-q)\xi})$; we will show that $(f^+, s^+, K^+) \gtrsim (\bar{f}, \bar{s}, \bar{K})$. Since clearly $(f^+, s^+, K^+) \neq (\bar{f}, \bar{s}, \bar{K})$, this is the desired contradiction.

In view of (16) and our choice of ε,

$$\begin{aligned} \|\mathscr{F}(f^+) - (1-\xi)\mathscr{F}(\bar{f})\|_0 &\leqslant \xi\beta \|\mathscr{F}(\bar{f})\|_0 + \xi\varepsilon^\mu \bar{K} \|\mathscr{F}(\bar{f})\|_0 \\ &= \xi q \|\mathscr{F}(\bar{f})\|_0. \end{aligned}$$

This implies via the triangle inequality that

$$\|\mathscr{F}(f^+) - \mathscr{F}(\bar{f})\|_0 \leqslant (1+q)\,\xi \|\mathscr{F}(\bar{f})\|_0$$

and

$$\|\mathscr{F}(f^+)\|_0 \leqslant (1 - (1-q)\,\xi) \|\mathscr{F}(\bar{f})\|_0.$$

The latter inequality implies $\|\mathscr{F}(f^+)\|_0 \leqslant \|\mathscr{F}(\bar{f})\|_0\, e^{-(1-q)(s^+ - \bar{s})}$ while combining the two inequalities gives

$$\|\mathscr{F}(f^+) - \mathscr{F}(\bar{f})\|_0 \leqslant \frac{1+q}{1-q} (\|\mathscr{F}(\bar{f})\|_0 - \|\mathscr{F}(f^+)\|_0).$$

Consequently (10), (11) and (12) hold.

For (13), we need to verify that $M|f^+|_p \leqslant K^+ = \bar{K}e^{\mu(1-q)\xi}$. Applying (b) of (H4) and our choice of ε gives

$$\begin{aligned} M|f^+|_p &\leqslant M(|\bar{f}|_p + \xi |f_\varepsilon|_p) \\ &\leqslant \bar{K} + \xi M \bar{K}^{(\mu+1)/\mu} \lambda^{-1/\mu} \|\mathscr{F}(\bar{f})\|_0 \\ &= \bar{K}(1 + \xi M \bar{K}^{1/\mu} \lambda^{-1/\mu} \|\mathscr{F}(\bar{f})\|_0). \end{aligned}$$

Thus $M|f^+|_p \leqslant K^+$ if $M(\bar{K}/\lambda)^{1/\mu}\,\|\mathscr{F}(\bar{f})\|_0 \leqslant \mu(1-q)$. But, applying (12) and (13),

$$M(\bar{K}/\lambda)^{1/\mu}\,\|\mathscr{F}(\bar{f})\|_0 \leqslant MK_0^{1/\mu}e^{(1-q)\bar{s}}\lambda^{-1/\mu}\,\|\mathscr{F}(f_0)\|_0\, e^{-(1-q)\bar{s}} \leqslant \mu(1-q)$$

by (9). Thus (13) holds.

For (14),

$$\begin{aligned} |f^+ - \bar{f}|_p = \xi\,|f_\varepsilon|_p &\leqslant \xi\bar{K}^{(\mu+1)/\mu}\lambda^{-1/\mu}\,\|\mathscr{F}(\bar{f})\|_0 \\ &\leqslant \int_0^{\xi} \bar{K}^{(\mu+1)/\mu}\lambda^{-1/\mu}\,\|\mathscr{F}(\bar{f})\|_0\, e^{\mu(1+q)t}\,dt \\ &= \int_{\bar{s}}^{s^+} \bar{K}^{(\mu+1)/\mu}\lambda^{-1/\mu}\,\|\mathscr{F}(\bar{f})\|_0\, e^{\mu(1-q)(t-\bar{s})}\,dt \end{aligned}$$

and so (14) holds.

Finally, we verify (15) and determine the constant A. Note that, by (H4)(a), $\|\mathscr{F}'(\bar{f})(f_\varepsilon) + \mathscr{F}(\bar{f})\|_0 \leqslant \varepsilon^\mu \bar{K}\,\|\mathscr{F}(\bar{f})\|_0$ and thus $\|\mathscr{F}'(\bar{f})(f_\varepsilon)\| \leqslant (1+\varepsilon^\mu\bar{K})\,\|\mathscr{F}(\bar{f})\|_0 = (1+\lambda)\,\|\mathscr{F}(\bar{f})\|$. Also, by (7), $(\mu+1)\,rp^{-1} < 1$ and so, since $0 \leqslant \xi \leqslant 1$,

$$\xi \leqslant \int_0^{\xi} \exp\left\{(t-1)(1-q)\left[\frac{(\mu+1)\,r}{p} - 1\right]\right\} dt.$$

Applying these observations gives

$$\begin{aligned} |\bar{f} - f^+|_r = \xi\,|f_\varepsilon|_r &\leqslant \xi C_r\,|f_\varepsilon|_0^{(p-r)/p}\,|f_\varepsilon|_p^{r/p} \\ &\leqslant \xi C_r(C\,\|\mathscr{F}'(\bar{f})(f_\varepsilon)\|_0)^{(p-r)/p}(\bar{K}^{(\mu+1)/\mu}\lambda^{-1/\mu}\,\|\mathscr{F}(\bar{f})\|_0)^{r/p} \\ &\leqslant C_r(C(1+\lambda)\,\|\mathscr{F}(\bar{f})\|_0)^{(p-r)/p}(\bar{K}^{(\mu+1)/\mu}\lambda^{-1/\mu}\,\|\mathscr{F}(\bar{f})\|_0)^{r/p} \\ &\qquad \int_0^{\xi} \exp\left\{(t-1)(1-q)\left[\frac{(\mu+1)\,r}{p} - 1\right]\right\} dt \\ &= A\,\|\mathscr{F}(\bar{f})\|_0\,\bar{K}^{(\mu+1)r/\mu p} \\ &\qquad \int_{\bar{s}}^{s^+} \exp\left\{(t-\bar{s})(1-q)\left[\frac{(\mu+1)\,r}{p} - 1\right]\right\} dt \end{aligned}$$

where

$$A = C_r(C(1+\lambda))^{(p-r)/p}\,\lambda^{-r/p\mu}\exp\left\{-(1-q)\left[\frac{(\mu+1)\,r}{p} - 1\right]\right\}.$$

Thus $(f^+, s^+, K^+) \gtrsim (\bar{f}, \bar{s}, \bar{K})$, and the proof of Theorem 1 is complete.

The proof of Theorem 2 follows along the same lines as Theorem 1; we omit many of the details.

Proof of Theorem 2. We may choose $\beta > 0$ so small and $\delta \in (0, \mu p(r+\mu r)^{-1})$ so large that

$$\|\mathscr{F}(f_0)\|_0 \leqslant \delta\mu \left(\frac{1-\beta}{1+\mu}\right)^{(\mu+1)/\mu} M^{-1}K_0^{-1/\mu}$$

$$= \delta \left(1-\beta-\frac{1-\beta}{1+\mu}\right)\left(\frac{1-\beta}{1+\mu}\right)^{1/\mu} M^{-1}K_0^{-1/\mu}.$$

Then with λ and q defined as in Theorem 1, we see that

$$\|\mathscr{F}(f_0)\|_0 M(K_0/\lambda)^{1/\mu} \leqslant \delta(1-q). \tag{17}$$

Also, in view of (8), $\delta \leqslant \mu$.

With (E, d) and ψ defined as in Theorem 1, define a partial order "$\lesssim$" on E by saying $(f_1, s_1, K_1) \lesssim (f_2, s_2, K_2)$ if and only if (10), (11) and (12) hold and

(18) $K_2 = K_1 \exp(\delta(1-q)(s_2-s_1))$;

(19) $|f_1 - f_2|_p \leqslant \lambda^{-1} \|\mathscr{F}(f_1)\|_0 K_1^{(\mu+1)/\mu} \int_{s_1}^{s_2} \exp(\delta(1-q)(t-s_1))\,dt$;

and

(20) $|f_1 - f_2|_r \leqslant A \|\mathscr{F}(f_1)\|_0 K_1^{(\mu+1)r/\mu p} \int_{s_1}^{s_2} \exp\{(t-s_1)(1-q)[(\delta(1+\mu)r)/\mu p - 1]\}\,dt$.

In exactly the same fashion as before one shows that "$\lesssim$" is a partial order. Assumptions (1) and (2) of Proposition B also hold exactly as before; for assumption (3), let $\{(f_n, s_n, K_n)\}$ be a nondecreasing sequence in $\tilde{E}$. If we can show $\{f_n\}$ is a Cauchy sequence in F_r, the remainder of the verification of (3) will follow as before. However, by (12) and (17), for $m > n$,

$$|f_n - f_m|_r \leqslant A \|\mathscr{F}(f_0)\|_0 K_0^{(\mu+1)r/\mu p}$$

$$\int_{s_n}^{s_m} \exp\left\{t(1-q)\left[\frac{\delta(\mu+1)r}{\mu p} - 1\right]\right\} dt,$$

and $\{f_n\}$ is a Cauchy sequence in F_r since, by our choice of δ, $\delta(\mu+1)r(\mu p)^{-1} < 1$.

Thus, by Proposition B, there is an $(\bar{f}, \bar{s}, \bar{K}) \in \tilde{E}$ which is maximal; we obtain a contradiction by selecting f_ε as in Theorem 1. Assumptions (10), (11) and (12) follow exactly as before; for (18):

$$M|f^+|_p \leqslant M(|\bar{f}|_p + \xi|f_\varepsilon|_p)$$

$$\leqslant \bar{K}(1 + \xi M(\bar{K}/\lambda)^{1/\mu} \|\mathscr{F}(\bar{f})\|_0)$$

and $M|f^+|_p \leqslant K^+$ if $M(\bar{K}/\lambda)^{1/\mu} \|\mathscr{F}(\bar{f})\|_0 \leqslant \delta(1-q)$. Applying (12) and (18) and using the fact that $\delta \leqslant \mu$,

$$M(\bar{K}/\lambda)^{1/\mu} \|\mathscr{F}(\bar{f})\|_0 \leqslant MK_0^{1/\mu} \exp\left(\frac{\delta(1-q)\,\bar{s}}{\mu}\right) \lambda^{-1/\mu} \|\mathscr{F}(f_0)\|_0$$

$$e^{-(1-q)\bar{s}} \leqslant \delta(1-q)$$

by (17), and thus (18) holds. Relations (19) and (20) follow exactly as in Theorem 1, and so Theorem 2 is verified.

The constant A of (20) will be

$$A = C_r(C(1+\lambda))^{(p-r)/p} \lambda^{-r/p} \exp\left\{-(1-q)\left[\frac{\delta(\mu+1)\,r}{\mu p} - 1\right]\right\}$$

in Theorem 2.

Remarks. (i) Since the solution f obtained in Theorem 1 must lie in $\mathscr{S}(f_0, 0, K_0)$, it follows that

$$|f - f_0|_r \leqslant A\, \|\mathscr{F}(f_0)\|_0\, K_0 \int_0^\infty \exp\left\{t(1-q)\left[\frac{(1+\mu)\,r}{p} - 1\right]\right\} dt$$

and thus we need to assume only that $\mathscr{F}$ is defined on a ball $B(f_0, R)$ where R is sufficiently large. A similar remark applies to Theorem 2.

(ii) W. Kirk and J. Caristi [6] and Kirk and D. Downing [3] obtain mapping theorems for Gâteaux differentiable operators $\mathscr{F}: X \to Y$ under the assumption that $\mathscr{F}'(x)(X)$ is dense in Y for each x and $\mathscr{F}(X)$ is closed. The hypothesis (H4) enables us to weaken their assumption on the range of $\mathscr{F}$.

(iii) In [1] M. Altman obtains results similar in character to the above. His proofs rely on transfinite induction. We remark that, while (5) is not explicitly assumed in [1], it appears to be necessary to the argument (in particular, to derive Eq. (3.28), p. 136 of [1]).

(iv) We have used the Brezis–Browder principle in order to avoid transfinite induction; the above proof carries over, essentially without change, if one uses Zorn's Lemma instead.

(v) While our argument is not constructive, it is a consequence of our proof that there exists a nondecreasing sequence $\{(f_n, s_n, K_n)\}$ which converges to a solution of $\mathscr{F}(f) = 0$. Unfortunately, the proof does not suggest how to select this sequence even when $\mathscr{F}'(x)$ is injective.

(vi) In [7] it is assumed that $\mathscr{F}$ admits the estimate

$$\|\mathscr{F}(f+h) - \mathscr{F}(f) - \mathscr{F}'(f)(h)\|_0 \leqslant M\,|h|_0^{2-\beta}\,|h|_p^\beta$$

for some $\beta \in (0, 1)$; this estimate is not required by our approach. Also, solutions are obtained in [7] in F_r only for $r < p$ sufficiently small, in contrast to our theorems.

(vii) As we observed in [8] and [9], theorems of the above type can be formulated in the context of "normal solvability." This reformulation is routine, and we leave it to the interested reader.

REFERENCES

1. M. ALTMAN, "Contractors and Contractor Directions Theory and Applications," Dekker, New York, 1977.
2. H. BREZIS AND F. E. BROWDER, A general principle on ordered sets in nonlinear functional analysis, *Adv. in Math.* **21** (1976), 777–787.
3. D. DOWNING AND W. A. KIRK, A generalization of Caristi's theorem with applications to nonlinear operators, *Pacific J. Math.* **69** (1977), 339–346.
4. I. EKELAND, Surr les problèmes variationnels, *C. R. Acad. Sci. Paris* **275** (1972), 1057–1059.
5. R. S. HAMILTON, The inverse function theorem of Nash and Moser, *Bull. Amer. Math. Soc.* (*N.S.*) (1982), 65–222.
6. W. A. KIRK AND J. V. CARISTI, A mapping theorem in metric and Banach spaces, *Bull. Pol. Acad. Sci. Ser. Math. Astron. Phys.* **23** (1975), 891–894.
7. J. MOSER, A rapidly convergent iteration method and nonlinear differential equations, *Ann. Scuola Norm. Sup. Pisa* **20** (1966), 499–535.
8. W. O. RAY AND W. J. CRAMER, Solvability of nonlinear operator equations, *Pacific J. Math.* **95** (1981), 37–50.
9. W. O. RAY AND A. M. WALKER, Mapping theorems for Gâteaux differentiable and accretive operators, *J. Nonlinear Anal.* **6** (1982), 423–433.