JOURNAL OF DIFFERENTIAL EQUATIONS 55, 330–345 (1984)

Differentiation in Abstract Spaces

MARTIN SCHECHTER*

Yeshiva University, New York, New York 10033

Received November 15, 1982

A new definition of differentiation for mappings between topological vector spaces is introduced. It does not require the mapping to be defined on a linear manifold, nor does it require it to be continuous. All of the main theorems of differential calculus hold and all other known definitions of differentiation are included. The new definition can be used for singular mappings and those defined on arbitrary sets. Applications are given. © 1984 Academic Press, Inc.

1. THE DEFINITION

There have been many definitions of differentiation since the time of Hadamard [1], Frechet [2], Gateaux [3] and Levy [4] (cf. Averbukh and Smolyanov [5] for a comprehensive survey). All of them can be described in the following way. One is given a mapping from X to Y, where X, Y are real topological vector spaces. One is then given a set $R(X, Y)$ of maps from X to Y which are considered small in some sense. One then says that the continuous linear map A from X to Y is the derivative of f at x if

$$f(x+h)-f(x)=Ah+r(h), \qquad h\in X \tag{1.1}$$

where $r\in R(X, Y)$. In all cases A was taken as an operator (or the restriction of one) defined on the whole of X and continuous from X to Y. Of course one can define differentiation along a subspace by letting W be a topological vector space contained in Y and considering f as a map from W to Y (cf. [9]).

However, none of these definitions can be used in the study of unbounded functions not defined on linear sets. For instance, let G be a functional defined on a set D which is not a linear manifold and such that G is unbounded from above and below on D. Suppose we are interested in solving the Euler–Lagrange equations for G. This means that we wish to find an element $u\in D$ such that

$$\lim_{t\to 0} t^{-1}[G(u+tq)-G(u)]=0 \tag{1.2}$$

* Current address: Mathematics Department, University of California, Irvine, California 92717.

0022-0396/84 $3.00

for all q in some dense set Q. When (1.2) holds one says that the variation of G vanishes at u. The most convenient way of showing that (1.2) holds is to find a minimum or maximum point for G. However, when G is unbounded from above and below, it is extremely difficult to fund such a point. One approach is to introduce constraints, say, of the form $F(v)=0$, where F is a mapping of D into some space Y. One tries to pick F in such a way that G will have a minimum on the set $S=\{v \in D \mid F(v)=0\}$. However, if

$$G(u)=\min_{S} G(v) \tag{1.3}$$

it does not necessarily follow that (1.2) holds. In order to use the calculus arguments that imply (1.2) at a minimum we would need $F(u+tq)=0$ for all $q \in Q$ and t near 0. This is too much to ask. The most one can expect is that for each $q \in Q$ there is a mapping $q(t)$ from the real numbers to Q such that

$$F(u+tq(t)) \equiv 0 \qquad \text{and} \qquad q(t) \to q \tag{1.4}$$

where the type of convergence depends on F and u. Clearly (1.3) and (1.4) do not imply (1.2).

The limit in (1.2) is called the Gateaux or weak derivative. It is at this point that one might wonder if differentiability of G in some other sense would allow (1.3), (1.4) to imply (1.2). If we attempt to use any of the known definitions we would need $r(tq(t))=o(t)$ for all $r \in R(X, Y)$ no matter how $q(t)$ converges to q. This is clearly impossible to achieve. This has led us to search for a definition which

1. does not require G to be defined on a linear manifold;
2. does not require the derivative to be defined everywhere or by continuous;
3. can be used in proving (1.2);
4. gives rise to the usual theorems of calculus;
5. contains all other known definitions of differentiation.

We introduce the following definition which satisfies all of these requirements

DEFINITION A. Let X be a vector space and let Z, Y be separated topological vector spaces such that $Q \subset X$. Let $G(x)$ be a mapping from a subset D of X to Y and let A be a linear map from X with $D(A)=Q$. We shall say that A is the derivative of G at x, with respect to Q, and write $A=G'_Q(x)$ if

(a) $x \in D$.

(b) For every $q \in Q$ there are sequences $\{q_n\} \subset Q$, $\{t_n\} \subset \mathbb{R}$ such that

$$q_n \to q \quad \text{in } Q, \qquad 0 \neq t_n \to 0 \qquad \text{and} \qquad x + t_n q_n \in D. \tag{1.5}$$

(c) If $\{q_n\}$, $\{t_n\}$ are any sequences satisfying (1.5), then

$$t_n^{-1}[G(x + t_n q_n) - G(x)] \to Aq \quad \text{in } Y \qquad \text{as} \qquad n \to \infty. \tag{1.6}$$

In the next section I shall show that this derivative obeys the usual theorems of calculus. Now we consider several examples.

1. *The Frechet derivative.* Take $D = Q = X$, Y Banach spaces, $B(X, Y)$ all bounded linear operators from X to Y. Then $A \in B(X, Y)$ is the Frechet derivative of f at x if (1.1) holds and

$$\frac{\|r(h)\|}{\|h\|} \to 0 \qquad \text{as} \quad \|h\| \to 0. \tag{1.7}$$

In this case f is clearly differentiable at x in the sense of Definition A. For example, let $\{q_n\}$, $\{t_n\}$ be any sequences satisfying (1.5) (which always exist), then by (1.1)

$$t_n^{-1}[f(x + t_n q_n) - f(x)] = Aq_n + t_n^{-1} r(t_n q_n) \to Aq$$

since

$$\|t_n^{-1} r(t_n q_n)\| = \|q_n\| \frac{\|r(t_n q_n)\|}{\|t_n q_n\|} \to 0$$

by (1.4). Thus $A = f'_X(x)$.

2. *The Gateaux derivative.* Take $D = X$. Then A is called the Gateaux derivative of f at x if

$$f(x + th) - f(x) = tAh + r(t), \qquad h \in X \tag{1.8}$$

where

$$t^{-1} r(t) \to 0 \qquad \text{as} \quad t \to 0.$$

(Note that A need not be linear on the whole of X.) We can consider A as a *partial derivative* in the following way. Let h be a fixed vector in X and let Q be the one-dimensional subspace of X containing h and let Q have the topology of $\mathbb{R}$. Then if $q \in Q$, there are sequences $\{q_n\}$, $\{t_n\}$ satisfying (1.5). Moreover, for such sequences we have $q_n = s_n h$, $q = sh$ with $s_n \to s$. Thus

$$\begin{aligned} f(x + t_n q_n) &= f(x + t_n s_n h) \\ &= f(x) + t_n s_n Ah + r(t_n s_n) \end{aligned} \tag{1.9}$$

and consequently $A = f'_Q(x)$. Of course A is linear on Q. Thus the problems associated with the Gateaux derivative can be obviated by considering it a partial derivative with respect to h.

3. *Hadamard (compact) differentiation.* Take $D = Q = X$, $L(X, Y)$ the set of continuous linear operators from X to Y. We say that $A \in L(X, Y)$ is the compact derivative of f at x if

$$f(x + t_n h_n) - f(x) = Ah + o(t_n)$$

whenever $t_n \to 0$ and $h_n \to h$ in X. Clearly A is the derivative in the sense of Definition A as well.

There have been many more definitions of differentiation proposed by various authors. Approximately twenty-five have been catalogued by Averbukh and Smolyanov in their survey article [5]. They show that if a mapping is differentiable in any sense listed there, then it is differentiable in the Gateaux sense (and all but the Gateaux derivative are differentiable in the compact sense). Hence, *if a mapping is differentiable in any sense mentioned there, it is differentiable in the sense of Definition A*. On the other hand, this definition is strong enough for all of the main theorems of differential calculus to hold. We prove some of them in Section 2.

In Section 3 we give an application in which the definition of derivative is critical. In it the constraints are of the form $F(u) = 0$. In order to show that a minimum under the constraints is stationary point without constraints, we find a Banach space N continuously embedded in Q such that

$$F(u + tq + h(t)) \equiv 0, \qquad q \in Q \text{ fixed} \tag{1.10}$$

has a solution $h(t) \in N$ such that $t^{-1}h(t)$ converges in N as $t \to 0$ (cf. Theorem 3.4). One of the requirements that (1.10) have a solution is that $F'_N(u)$ have a bounded inverse on N. One of the main difficulties in the application is that there does not exist a Banach space N for which $F'_N(u)$ exists and has a bounded inverse for all u under consideration. Our approach is to pick N and Q to depend on u. We then proceed to show that $G'_Q(u) = 0$. This can be achieved only if the topology of Q is just right. If it is too strong, the derivative will not exist. If it is too weak, it will not vanish.

The method presented here is a generalization of methods developed with M. Berger in [6] and R. Weder in [7]. In the former case we applied the technique to various problems in partial differential equations. In the latter we proved the (theoretical) existence of dyons-subatomic particles with both electric and magnetic charge. (The existence of dyons was conjectured by Schwinger [8].) The reason for the present generalization is twofold. First, I wanted to obtain a definition of derivative that would include all others. Second, the methods of [6] and [7] were not powerful enough to handle the application considered here.

2. Properties of the Derivative

We now show that the derivative given by Definition A has the usual properties desired for derivatives. Throughout we assume that X is a vector space, Q, Y are topological vector spaces, $Q \subset X$ and G maps $D \subset X$ into Y. We let Y^* denote the space of continuous linear functionals y^* on Y.

First we give a slightly more general definition of derivative.

Definition B. If A is a linear operator from X to Y with $D(A) = Q$, we shall say that A is the weak derivative of G at x with respect to Q if for each $y^* \in Y^*$ the mapping y^*G from D to $\mathbb{R}$ has a derivative at x with respect to Q which equals y^*A.

Clearly any derivatve under Definition A is a weak derivative under Definition B. We have

Lemma 2.1 (The Mean Value Theorem). *Assume that* $q \in Q$, $y^* \in Y^*$ *and that* $u + sq$ *is in* D *for* $0 \leqslant s \leqslant 1$. *Assume also that*

$$y^*G(u+sq) \to y^*G(u) \quad \text{as} \quad s \to 0 \tag{2.1}$$

$$y^*G(u+sq) \to y^*G(u+q) \quad \text{as} \quad s \to 1 \tag{2.2}$$

and that the weak derivative $G'_Q(u+sq)$ *exists for* $0 < s < 1$. *Then there is a* θ *satisfying* $0 < \theta < 1$ *and such that*

$$y^*[G(u+q) - G(u)] = y^*G'_Q(u+\theta q)\, q. \tag{2.3}$$

Proof. Put $w(s) = y^*G(x+sq)$ and let $\{t_n\}$ be a sequence of real numbers converging to 0. Then for $0 < s < 1$

$$\begin{aligned} &t_n^{-1}[w(s+t_n) - w(s)] \\ &\quad = t_n^{-1} y^*[G(u+sq+t_n q) - G(u+sq)] \to y^*G'_Q(u+sq)\, q. \end{aligned}$$

Thus $w'(s)$ exists in $(0, 1)$. Moreover, $w(s)$ is continuous in $[0, 1]$ by (2.1) and (2.2). Hence by the mean value theorem there is a θ in $(0, 1)$ such that $w(1) - w(0) = w'(\theta)$. This gives (2.3). ∎

THEOREM 2.2 (The Chain Rule). *Let Z be a topological vector space, and let R be a closed subspace of Y. Let F be a map of $V \subset Y$ to Z, and put $W = \{x \in D \mid G(x) \in V\}$. Assume that $u \in W$, $v = G(u)$ and that $G'_Q(u)$, $F'_R(v)$ exist. Assume further that for every q in Q there are sequences $\{q_n\}$, $\{t_n\}$ such that*

$$q_n \to q \quad \text{in } Q, \qquad 0 \neq t_n \to 0, \qquad u + t_n q_n \in W \tag{2.4}$$

and for all sequences satisfying (2.4), $G(u + t_n q_n) - v$ is in R. Put $H(x) = F(G(x))$. Then $H'_Q(u)$ exists and equals $F'_R(v)\, G'_Q(u)$.

Proof. Let $\{q_n\}$, $\{t_n\}$ satisfy (2.4). Then

$$\begin{aligned} H(u + t_n q_n) - H(u) &= F(G(u + t_n q_n)) - F(v) \\ &= F(r + t_n r_n) - F(v) \end{aligned}$$

where

$$r_n = t_n^{-1}[G(u + t_n q_n) - G(u)] \to G'_Q(u)\, q \qquad \text{in} \quad Y.$$

By hypothesis, r_n is in R and consequently $r_n \to G'_Q(u)\, q$ in R. Hence

$$t_n^{-1}[F(v + t_n r_n) - F(v)] \to F'_R(v)\, G'_Q(n)\, q \qquad \text{in} \quad Z.$$

This gives the theorem. ∎

LEMMA 2.3. *Suppose W is a topological vector space contained in Q with continuous injection. Assume that u is in D and that for each h in W there are sequences $\{h_n\} \subset W$, $\{t_n\} \subset \mathbb{R}$ such that*

$$h_n \to h \quad \textit{in} \quad W, \qquad 0 \neq t_n \to 0, \qquad u + t_n h_n \in D \tag{2.5}$$

If $G'_Q(u)$ exists, then $G'_W(u)$ exists and equals the restriction of $G'_Q(u)$ to W.

Proof. Suppose (2.5) is satisfied. Then $h_n \to h$ in Q. Since $G'_Q(u)$ exists, we have

$$t_n^{-1}[G(u + t_n h_n) - G(u)] \to G'_Q(u)\, h \qquad \text{in} \quad Y.$$

Thus $G'_W(u)$ exists and

$$G'_W(u)\, h = G'_Q(u)\, h, \qquad h \in W. \quad ∎$$

THEOREM 2.4 (IMPLICIT FUNCTION THEOREM). *Let F map $V \subset X$ into a Banach space $N \subset X$, and let $u \in V$, $q \in Q$ be given. Assume that there is an $m > 0$ such that if $v = u + tg + h$ with h in N, $\|h\| < m$ and $|t| < m$, then v is* *in V, the weak derivative $T(v)=F'_N(v)$ exists, $T=T(u)$ has a bounded inverse on N and*

$$\|T(v)-T\| \to 0 \qquad as \quad t\to 0 \text{ and } \|h\|\to 0. \tag{2.6}$$

If the derivative $F'_Q(u)$ exists, then there is a mapping $h(t)$ from an open interval $(-t_0, t_0)$ to N such that

$$F(u+tq+h(t))=F(u), \qquad |t|<t_0 \tag{2.7}$$

and

$$t^{-1}h(t)\to -T^{-1}F'_Q(u)\,q \quad in\ N \qquad as \quad t\to 0. \tag{2.8}$$

Proof. By replacing $F(v)$ by $T^{-1}[F(v)-F(u)]$ we may assume that $F(u)=0$ and $T=1$. Put $R(t,h)=F(u+tq+h)-h$ and let n^* be a bounded linear functional on N. Then by Lemma 2.1,

$$\begin{aligned} n^*[R(t,h_1)-R(t,h_2)] &= n^*[F(u+tq+h_1)-F(u+tq+h_2)] \\ -n^*(h_1-h_2) &= n^*[T(u+tq+h_\theta)-1](h_1-h_2) \end{aligned} \tag{2.9}$$

provided $|t<m$, $\|h_i\|<m$, where $0<\theta<1$ and $h_\theta=h_1-\theta(h_1-h_2)$. Take $\delta>0$ so small that

$$\|T(v)-1\|<\varepsilon, \qquad |t|<\delta, \qquad \|h\|<\delta. \tag{2.10}$$

Thus if $\|h_i\|<\delta$, then $\|h_\theta\|<\delta$, and consequently

$$|n^*[R(t,h_1)-R(t,h_2)]| \leqslant \varepsilon\,\|n^*\|\,\|h_1-h_2\|.$$

Since n^* was arbitrary, this implies

$$\|R(t,h_1)-R(t,h_2)\| \leqslant \varepsilon\,\|h_1-h_2\|. \tag{2.11}$$

Thus for each t in $|t|<\delta$ there is an element $h_t\in N$ such that $\|h_t\|<\delta$ and

$$R(t,h_t)+h_t=0$$

that is

$$F(u+tq+h_t)=0. \tag{2.12}$$

Moreover,

$$\begin{aligned} &n^*[t^{-1}h_t+F'_Q(u)\,q] \\ &\quad = t^{-1}n^*\{h_t-[F(u+tq+h_t)-F(u+tq)] \\ &\qquad -[F(u+tq)-F(u)-tF'_Q(u)\,q]\} \end{aligned}$$

$$= t^{-1} n^* \{h_t - [F(u + tq + h_t) - F(u + tq)]\}$$
$$- n^* \{t^{-1}[F(u + tq) - F(u)] - F'_Q(u)\, q\}$$
$$= t^{-1} n^* [1 - T(u + tq + \theta h_t)]\, h_t$$
$$- n^* \{t^{-1}[F(u + tq) - F(u)] - F'_Q(u)\, q\}.$$

The last term converges to 0 as $t \to 0$. Thus for t sufficiently small

$$\|t^{-1} h_t + F'_Q(u)\, q\| \leqslant \varepsilon(\|t^{-1} h_t\| + 1). \tag{2.13}$$

This shows that $\|t^{-1} h_t\| \leqslant C$ and consequently that the left-hand side of (2.13) converges to 0 as $t \to 0$. We take $h(t) = h_t$. Thus (2.7) and (2.8) hold. ▮

THEOREM 2.5. *Let F, u, q satisfy the hypothesis of Theorem 2.3. Let G map $D \subset X$ into $\mathbb{R}$ and put*

$$R = \{v \in V \cap D \mid F(v) = F(u)\}.$$

Assume that u is in R and that

$$G(u) = \min_R G(v).$$

Assume further that $G'_Q(u)$ exists and that $u + tq + h \in D$ whenever $h \in N$, t is small and $F(u + tq + h) = F(u)$. If N is continuously imbedded in Q, then

$$G'_Q(u)(1 - T^{-1} F'_Q(u))\, q = 0. \tag{2.14}$$

Proof. Let $\{t_n\}$ be any sequence in $\mathbb{R}$ convergent to 0, and put $q_n = q + t_n^{-1} h(t_n)$, where $h(t)$ is the function given by Theorem 2.3. Then by (2.8)

$$q_n \to (1 - T^{-1} F'_Q(u))\, q \qquad \text{in} \quad Q.$$

Since $G'_Q(u)$ exists, we have

$$t_n^{-1}[G(u + t_n q_n) - G(u)] \to G'_Q(u)(1 - T^{-1} F'_Q(u))\, q \qquad \text{in} \quad \mathbb{R}. \tag{2.15}$$

Moreover, $F(u + t_n q_n) = F(u + t_n q + h(t_n)) = F(u)$. Thus $u + t_n q_n \in R$, and consequently $G(u + t_n q_n) \geqslant G(u)$. If we take $t_n > 0$, we see that the limit in (2.15) is $\geqslant 0$. If we take $t_n < 0$, we see that it is $\leqslant 0$. Thus (2.14) holds. ▮

3. An Application

Suppose we wish to find functions $x(r)$, $y(r)$ defined in $(0, \infty)$ such that

$$\ddot{x} = g(x) - 2xy - xy^2 \tag{3.1}$$

$$\ddot{y} = x^2 + x^2 y \tag{3.2}$$

$$x(0) = 1, \qquad y(0) = 0 \tag{3.3}$$

and

$$\dot{x}, \dot{y}, xy \in L^2 = L^2(0, \infty), \qquad x^2 y, f(x) \in L^1 = L^1(0, \infty) \tag{3.4}$$

where $g(t) = \frac{1}{2} f'(t)$ is a continuous function on $\mathbb{R}$ and $\dot{x} = dx/dr$, $\ddot{x} = d\dot{x}/dr$. It is easily verified that (3.1), (3.2) are the Euler–Lagrange equations corresponding to the functional

$$G(u) = \int_0^\infty (\dot{x}^2 - \dot{y}^2 + f(x) - x^2 y^2 - 2x^2 y)\, dr, \qquad u = \{x, y\}. \tag{3.5}$$

Thus it would suffice to find an extremal of (3.5) subject to (3.3). However, two difficulties present themselves immediately. The first is that $G(u)$ is not defined on a linear set and the second is that $G(u)$ is not bounded from above or below. The combination of these two facts makes the dealing with (3.5) extremely difficult. Our approach is to minimize (3.5) under certain constraints. We must then show that the minimum of $G(u)$ with the constraints is a staionary point of $G(u)$ without the constraints. It is at this point that the definition of derivative plays an important role. If the definition is too restrictive, it will not exist. If it is too weak, it will not vanish.

Our method of attacking the problem is as follows. We fix $x \in L^2$ and try to minimize the functional

$$H(y) = \int_0^\infty (\dot{y}^2 + x^2 y^2 + 2x^2 y + x^2)\, dr. \tag{3.6}$$

It turns out that for each fixed $x \in L^2$, (3.6) has a unique minimum $y = y_x(r)$ in the set

$$D_x = \{ y \mid \dot{y}, xy \in L^2, y(0) = 0 \}. \tag{3.7}$$

We then consider $G(u)$ not for all x, y satisfying (3.3), (3.4) (where it is not bounded below) but for those x, y satisfying (3.3), (3.4) for which $y = y_x(r)$. On this subset $G(u)$ is bounded from below and we are able to obtain a minimum. The problem now is to show that this minimum is a stationary point of $G(u)$. Once this is established, it follows from standard methods that the stationary point is a solution of (3.1)–(3.4). We proceed with a series of lemmas.

LEMMA 3.1. *For each $x \in L^2$ there is a unique function $y_x(r) \in D_x$ such that $H(y_x) = \beta = \inf_{D_x} H(z)$. It is the only solution of*

$$(\dot{y}, \dot{z}) + (x(y+1), xz) = 0, \qquad z \in D_x. \tag{3.8}$$

Proof. Let $\{y_n\}$ be a sequence of functions in D_x such that $H(y_n) \to \beta$. Now

$$H(y) = \|\dot{y}\|^2 + \|x(y+1)\|^2$$

where the norm is that of L^2. Since $H(y_n) \leqslant C$ there is a subsequence (also denoted by $\{y_n\}$) such that $\{\dot{y}_n\}$ and $\{xy_n\}$ converge weakly in L^2. This implies that there is a $y \in D_x$ such that

$$\dot{y}_n \to \dot{y}, \qquad xy_n \to xy, \qquad \text{weakly in} \quad L^2. \tag{3.9}$$

To see this, let $v(r)$ be the weak limit of $\dot{y}_n$ in L^2. Put

$$y(r) = \int_0^r v(t)\, dt.$$

Then $\dot{y} = v$. Let w be a function in L^2 which vanishes for r large. Then

$$(x(y_n - y), w) = \int_0^\infty x(r)\, w(r) \int_0^r (\dot{y}_n - v)\, dt\, dr$$

$$= \int_0^\infty (\dot{y}_n - v) h\ dt \to 0$$

since

$$h(t) = \int_t^\infty x(r)\, w(r)\, dr \in L^2.$$

Since $\{xy_n\}$ converges weakly in L^2, the limit must be xy. Thus $y \in D_x$. Hence we have

$$\|\dot{y}_n - \dot{y}\|^2 + \|x(y_n - y)\|^2$$
$$= H(y_n) + H(y) - 2(\dot{y}_n, \dot{y}) - 2(x(y_n+1), x(y+1)) \to \beta - H(y) \leqslant 0.$$

This shows that the limits in (3.9) exist in the strong sense and $H(y) = \beta$. Taking the derivative of H (in any sense) gives

$$H'(y)\,z = 2(\dot{y}, \dot{z}) + 2(x(y+1), xz).$$

Since y is a minimum point, $H'(y) = 0$. This gives (3.8). Next we note that (3.8) has only one solution. For the difference of two solutions satisfies

$$(\dot{y}, \dot{z}) + (xy, xz) = 0, \qquad z \in D_x$$

which implies

$$\|\dot{y}\|^2 + \|xy\|^2 = 0. \quad \blacksquare$$

Next we put

$$S = \{u = \{x, y\} \mid x, \dot{x} \in L^2, f(x) \in L^1, x(0) = 1, y = y_x\}.$$

Note that S is not empty. If we let x be any smooth function which vanishes for r large and satisfies $x(0) = 1$ and take $y = y_x$ (which exists by Lemma 3.1), then $u = \{x, y\}$ is in S. Put

$$D = \{u = \{x, y\} \mid x, \dot{x} \in L^2, f(x) \in L^1, x(0) = 1, y \in D_x\}$$

and

$$\hat{D} = \{z \mid \dot{z} \in L^2, z(0) = 0, z(r) = 0 \text{ for } r \text{ large}\}.$$

We shall need

LEMMA 3.2. *If $u \in D$ and* (3.8) *holds for all $z \in \hat{D}$, then it holds for all $z \in D_x$.*

Proof. Let $\varphi(r) \in C^\infty$ satisfy $\varphi(r) = 1$ for $r < 1$, $\varphi(r) = 0$ for $r > 2$, $0 \leqslant \varphi \leqslant 1$, and put $\varphi_a(r) = \varphi(r/a)$. Let z be any function in D_x, and put $z_a = z\varphi_a$. Then clearly $z_a \in \hat{D}$ for any a. Now $x^2 y z_a$, $x^2 z_a$, $\dot{y}\dot{z}z_a$ converge pointwise to $x^2 yz$, $x^2 z$, $\dot{y}\dot{z}$, respectively, and they are majorized by them as well. Since the latter functions are in L^1, we have

$$(xy, xz_a) \to (xy, xz), \qquad (x, xz_a) \to (x, xz), \qquad (\dot{y}, \dot{z}\varphi_a) \to (\dot{y}, \dot{z})$$

as $a \to \infty$. Moreover, I claim that the function $z\dot{\varphi}_a$ converges weakly to 0 in L^2. To see this note that

$$|z(r) - z(r')|^2 \leqslant |r - r'|\,\|\dot{z}\|^2 \tag{3.10}$$

by a simple application of the Schwarz inequality. Since $z(0)=0$ we have

$$|z(r)|^2 \leqslant r\,\|\dot{z}\|^2. \tag{3.11}$$

Thus

$$\|z\dot{\varphi}_a\|^2 = a^{-2}\int_0^\infty z(r)^2\,\dot{\varphi}(r/a)^2\,dr \leqslant a^{-2}\,\|\dot{z}\|^2 \int_0^\infty r\dot{\varphi}(r/a)^2\,dr$$
$$= \|\dot{z}\|^2 \int_0^\infty s\dot{\varphi}(s)^2\,ds.$$

Thus the L^2 norm of $z\dot{\varphi}_a$ is bounded uniformly in a. Consequently, there is a subsequence that converges weakly. Since $z\dot{\varphi}_a$ converges to 0 pointwise, the weak limit must vanish. Now, suppose (3.8) holds for $z \in \hat{D}$, and let z be any element of D_x. Since z_a is in $\hat{D}$, we have

$$(\dot{y}, \dot{z}_a) + (x(y+1), xz_a) = 0$$

or

$$(\dot{y}, z\dot{\varphi}_a + \dot{z}\varphi_a) + (x(y+1), xz_a) = 0.$$

Taking the limits, we see that (3.8) holds for z as well. ∎

Now we make a basic assumption on $f(t)$.

Hypothesis A. $f(t) \geqslant c_0 t^2$ for some $c_0 > 0$.

LEMMA 3.3. *Under Hypothesis* A *there is a* $u \in S$ *such that*

$$G(u) = \alpha = \inf_S G(v). \tag{3.12}$$

Proof. If $y \in D_x$ satisfies (3.8), then

$$\|\dot{y}\|^2 + \|xy\|^2 + (x, xy) = 0.$$

Thus for $u \in S$,

$$G(u) = \int_0^\infty (\dot{x}^2 + f(x) + \dot{y}^2 + x^2y^2)\,dr$$
$$\geqslant c_0\,\|\dot{x}\|^2 + \|x\|^2 + \|\dot{y}\|^2 + \|xy\|^2 \tag{3.13}$$

by Hypothesis A. Recall that S is not empty. Let $\{u_n\} = \{x_n, y_n\}$ be a sequence such that $G(u_n) \to \alpha$. By (3.10) there is a subsequence (also denoted by $\{u_n\}$) such that $x_n, \dot{x}_n, \dot{y}_n, x_n y_n$ all converge weakly in L^2. Thus there are functions x, y such that

$$x_n \to x, \qquad \dot{x}_n \to \dot{x}, \qquad \dot{y}_n \to y_n, \qquad \text{weakly in} \quad L^2.$$

By (3.10) the $\{x_n\}$ are uniformly bounded and equicontinuous on any compact interval. Thus there is a subsequence (again denoted by $\{x_n\}$) which converges uniformly on any compact interval. The same is true of $\{y_n\}$. Since $x_n y_n$ converges weakly in L^2, it must converge to xy. Since $f(t) \geqslant 0$, we have by Fatou's lemma that $f(x) \in L^1$ and

$$\int_0^\infty f(x)\, dr \leqslant \liminf \int_0^\infty f(x_n)\, dr.$$

Thus $G(u)$ is weakly lower semicontinuous and we have $G(u) \leqslant \alpha$. We must show that $u \in S$. Let z by any function in $\hat{D}$. Then z is in D_x for any x. Hence

$$(\dot{y}_n, \dot{z}) + (x_n(y_n + 1), x_n z) = 0.$$

Since the x_n, z_n converge uniformly on the support of z and $\dot{y}_n$ converges weakly, we have in the limit

$$(\dot{y}, \dot{z}) + (x(y+1), xz) = 0. \tag{3.14}$$

Note that $u \in D$. Since (3.8) holds for all $z \in \hat{D}$, we can conclude by Lemma 3.2 that it holds for all $z \in D_x$. Thus $y = y_x$, and $u \in S$. ∎

Next put

$$W = \{w = \{\sigma, \tau\} \mid \sigma \in \mathscr{D}(0, \infty), \tau, \dot{\tau} \in L^2, \tau(0) = 0\}. \tag{3.15}$$

We have

THEOREM 3.4. *If u is in S and satisfies* (3.12), *then* $G'_W(u) = 0$.

Proof. Put $u = \{x, y\}$,

$$Q = \{q = \{\mu, \nu\} \mid \mu \in \mathscr{D}(0, \infty), \nu \in D_x\} \tag{3.16}$$

and

$$N = \{q \in Q \mid \mu = 0\}.$$

With the norm given by

$$\|(0, z)\|^2 = \|\dot{z}\|^2 + \|xz\|^2, \qquad (0, z) \in N \tag{3.17}$$

N becomes a Hilbert space continuously imbedded in Q. Let V be the set of those $v = (\sigma, \tau)$ in D such that the expression

$$M(w, v) = (\dot{z}, \dot{\tau}) + (\sigma z, \sigma(\tau + 1)), \qquad w = (0, z) \in N$$

is a bounded linear functional on N. Note that u is in V. Moreover, if $q = (\mu, \nu)$ is in Q, then $v = u + q$ is in V. To see this note that $\sigma = x + \mu$, $\tau = y + \nu$. Thus

$$M(w, v) = M(w, u) + (\dot{z}, \dot{\nu}) + (xz, x\nu)$$
$$+ ([2x + \mu]\, z, \mu(y + \nu + 1)).$$

Since ν is in D_x and μ is in $\mathscr{D}$, it is easily checked that this is a bounded linear functional on N. For $v \in V$ there is an element $F(v)$ in N such that

$$(F(v), w)_N = (\dot{\tau}, \dot{z}) + (\sigma(\tau + 1), \sigma z), \qquad w \in N. \tag{3.18}$$

Thus F is a mapping from V to N. The computation

$$\begin{aligned} &t^{-1}(F(v + tq) - F(v), w)_N \\ &\quad = (\dot{\nu}, \dot{z}) + (\sigma\nu, \sigma z) + 2(\sigma(\tau + 1), \mu z) + 2t(\sigma\nu, \mu z) \\ &\quad\quad + t(\mu(\tau + 1), \mu z) + t^2(\mu\nu, \mu z) \end{aligned}$$

shows that the derivative $F'_Q(v)$ exists for $v = u$ plus an element of Q. It is given by

$$(F'_Q(v)\, q, w)_N = (\dot{\nu}, \dot{z}) + (\sigma\nu, \sigma z) + 2(\sigma(\tau + 1), \mu z). \tag{3.19}$$

In particular, we have

$$(F'_N(v), h, w)_N = (\dot{\rho}, \dot{z}) + (\sigma\rho, \sigma z),$$
$$h = (0, \rho) \in N, \qquad w = (0, z) \in N$$

and

$$(F'_N(u)\, h, w)_N = (h, w)_N, \qquad h, w \in N. \tag{3.20}$$

Thus $F'_N(u) = 1$ and

$$(F'_N(v)\, h - h, w)_N = ([\sigma + x]\, \rho, [\sigma - x]\, z)$$

which shows that (2.6) holds. Let

$$R = \{v \in V \mid F(v) = 0\}.$$

Then by (3.18), the definitions and Lemma 3.1 it follows that $R \subset S$. Since u is in R, we have

$$G(u) = \min_S G(v) = \min_R G(v).$$

Finally, we note that $D + Q \subset D$ and

$$\begin{aligned} &G(u+tq) - G(u) \\ &\quad = 2t[(\dot{x}, \dot{\mu}) - (xy, vy) - 2(xy, \mu) - (\dot{y}, \dot{v}) - (x(y+1), xv)] \\ &\qquad + t^2[\|\dot{\mu}\|^2 - \|\dot{v}\|^2 - 2(x\mu, yv) - \|xv\|^2 \\ &\qquad - \|\mu y\|^2 - 2(\mu, \mu y) - 4(x\mu, v)] \\ &\qquad - 2t^3[(\mu y, \mu v) + (xv, \mu v) + (\mu, \mu v)] \\ &\qquad - t^4 \|\mu v\|^2 + \int_0^\infty [f(x+t\mu) - f(x)]\, dr. \end{aligned}$$

Now

$$t^{-1} \int_0^\infty [f(x+t\mu) - f(x)]\, dr = 2 \int_0^1 \int_0^\infty g(x + st\mu)\, \mu\, dt\, ds.$$

If μ converges in the topology of $\mathscr{D}$, then this expression converges to $2 \int_0^\infty g(x)\, \mu\, dr$ as $t \to 0$. It is now easily checked that if $t \to 0$ and q converges in Q, then

$$\begin{aligned} G'_Q(u)\, q = 2(\dot{x}, \dot{\mu}) + 2(g(x), \mu) - 2(xy, \mu y) \\ - 4(xy, \mu) - 2[(\dot{y}, \dot{v}) + (x(y+1), xv)]. \end{aligned} \tag{3.21}$$

Thus all of the hypotheses of Theorem 2.5 are satisfied. Moreover, since u is in S, we see from (3.8) and (3.21) that

$$G'_Q(u)\, q = 0 \tag{3.22}$$

if q is in N. Thus

$$G'_Q(u)\, T^{-1} F'_Q(u)\, q = 0, \qquad q \in Q.$$

We can now apply (2.14) to conclude that (3.22) holds for all q in Q. This is not quite what we wanted. However, if we note that W is continuously imbedded in Q and recall that $D + Q \subset D$, we can apply Lemma 2.3 to conclude that $G'_W(u)$ exists and is the restriction of $G'_Q(u)$ to W. Thus Theorem 3.4 is proved. ∎

We can now apply standard methods to the functional $G(u)$ to obtain

THEOREM 3.5. *Under Hypothesis* A, *there exists a solution of* (3.1)–(3.4).

REFERENCES

1. J. HADAMARD, La notion de differentielle dans l'enseignement, *Scripta Univ. Ab. Hierosolymitanrum* **1** (1923), 3.
2. M. FRECHET, Sur la notion de differentielle, *C. R. Acad. Sci. Paris* **152** (1911), 845–847.
3. R. GATEAUX, Sur les fonctionelles continues et les fonctionelles analytiques, *C. R. Acad. Paris* **157** (1913), 325–327.
4. P. LEVY, "Leçons d'analyse fonctionelle," Gauthier–Villars, Paris, 1922.
5. V. I. AVERBUKH AND O. G. SMOLYANOV, The various definitions of the derivative in linear topological spaces, *Russian Math. Surveys* **23** (1968), 76–112.
6. M. BERGER AND M. SCHECHTER, On the solvability of semilinear gradient operator equations, *Adv. in Math.* **25** (1977), 97–132.
7. M. SCHECHTER AND R. WEDER, A theorem on the existence of dyon solutions, *Ann. Phys. (N. Y.)* **132** (1981), 292–327.
8. J. SCHWINGER, A magnetic model of matter, *Science* **165** (1969), 757–761.
9. V. I. AVERBUKH AND O. G. SMOLYANOV, The theory of differentiation in linear topological spaces, *Russian Math. Surveys* **22** (1967), 201–258.