CHAPTER XII

NONLINEAR FUNCTIONAL ANALYSIS, INFINITE-DIMENSIONAL CALCULUS

DEFINITION Let E and F be (possibly infinite dimensional) real or complex Banach spaces, and let f be a map from a subset D of E into F. We say that f is *differentiable* at a point $x \in D$ if:

(1) x belongs to the interior of D; i.e., there exists an $\epsilon > 0$ such that $B_\epsilon(x) \subseteq D$.

(2) There exists a continuous linear transformation $L : E \to F$ and a function $\theta : B_\epsilon(0) \to F$ such that

$$f(x+h) - f(x) = L(h) + \theta(h), \tag{12.1}$$

for all $h \in B_\epsilon(0)$, and

$$\lim_{h \to 0} \|\theta(h)\| / \|h\| = 0. \tag{12.2}$$

The function f is said to be *differentiable on D* if it is differentiable at every point of D.

If $E = \mathbb{R}$, i.e., if f is a map from a subset D of $\mathbb{R}$ into a Banach space F, then f is said to have a *derivative* at a point $x \in D$ if $\lim_{t \to 0}[f(x+t) - f(x)]/t$ exists, in which case we write

$$f'(x) = \lim_{t \to 0} \frac{f(x+t) - f(x)}{t}. \tag{12.3}$$

If $D \subseteq E$, $D' \subseteq F$, and $f : D \to D'$, then f is called a *diffeomorphism* of D onto D' if f is a homeomorphism of D onto D' and f and f^{-1} are differentiable on D and D' respectively.

EXERCISE 12.1. (a) Suppose $f : D \to F$ is differentiable at a point $x \in D$, and write

$$f(x+h) - f(x) = L(h) + \theta(h)$$

as in Equation (12.1). Prove that $\theta(0) = 0$.

(b) Let $D \subseteq \mathbb{R}$, and suppose f is a function from D into a Banach space F. Show that f is differentiable at a point $x \in D$ if and only if f has a derivative at x. If f has a derivative at x, what is the continuous linear transformation $L : \mathbb{R} \to F$ and what is the map θ that satisfy Equation (12.1)?

THEOREM 12.1. *Suppose $f : D \to F$ is differentiable at a point x. Then both the continuous linear transformation L and the map θ of Equation (12.1) are unique.*

PROOF. Suppose, as in Equations (12.1) and (12.2), that

$$f(x+h) - f(x) = L_1(h) + \theta_1(h),$$

$$f(x+h) - f(x) = L_2(h) + \theta_2(h),$$

$$\lim_{h \to 0} \|\theta_1(h)\| / \|h\| = 0,$$

and

$$\lim_{h \to 0} \|\theta_2(h)\| / \|h\| = 0.$$

Then

$$L_1(h) - L_2(h) = \theta_2(h) - \theta_1(h).$$

If $L_1 \neq L_2$, choose a unit vector $u \in E$ such that $\|L_1(u) - L_2(u)\| = c > 0$. But then,

$$\begin{aligned}
0 &= \lim_{t \to 0} (\|\theta_2(tu)\| / \|tu\| + \|\theta_1(tu)\| / \|tu\|) \\
&\geq \lim_{t \to 0} \|\theta_2(tu) - \theta_1(tu)\| / \|tu\| \\
&= \lim_{t \to 0} \|L_1(tu) - L_2(tu)\| / \|tu\| \\
&= \lim_{t \to 0} |t| c / (|t| \|u\|) \\
&= c \\
&> 0,
\end{aligned}$$

which is a contradiction. Therefore, $L_1 = L_2$, whence $\theta_1 = \theta_2$ as well.

DEFINITION. Suppose $f : D \to F$ is differentiable at a point x. The (unique) continuous linear transformation L is called the *differential of f at x*, and is denoted by df_x. The differential is also called the *Fréchet derivative* of f at x.

THEOREM 12.2. *Let E and F be real or complex Banach spaces.*

(1) *Let $f : E \to F$ be a constant function; i.e., $f(x) \equiv y_0$. Then f is differentiable at every $x \in E$, and df_x is the zero linear transformation for all x.*
(2) *Let f be a continuous linear transformation from E into F. Then f is differentiable at every $x \in E$, and $df_x = f$ for all $x \in E$.*
(3) *Suppose $f : D \to F$ and $g : D' \to F$ are both differentiable at a point x. Then $f + g : D \cap D' \to F$ is differentiable at x, and $d(f + g)_x = df_x + dg_x$.*
(4) *If $f : D \to F$ is differentiable at a point x, and if c is a scalar, then the function $g = cf$ is differentiable at x and $dg_x = cdf_x$.*
(5) *If $f : D \to F$ is differentiable at a point x, and if v is a vector in E, then*

$$df_x(v) = \lim_{t\to 0} \frac{f(x + tv) - f(x)}{t}.$$

(6) *Suppose f is a function from a subset $D \subseteq \mathbb{R}$ into F. If f is differentiable at a point x (equivalently, f has a derivative at x), then*

$$f'(x) = df_x(1).$$

PROOF. If $f(x) \equiv y_0$, then we have

$$f(x + h) - f(x) = 0 + 0;$$

i.e., we may take both L and θ to be 0. Both Equations (12.1) and (12.2) are satisfied, and $df_x = 0$ for every x.

If f is itself a continuous linear transformation of E into F, then

$$f(x + h) - f(x) = f(h) + 0;$$

i.e., we may take $L = f$ and $\theta = 0$. Then both Equations (12.1) and (12.2) are satisfied, whence $df_x = f$ for every x.

To prove part 3, write

$$f(x+h) - f(x) = df_x(h) + \theta_f(h)$$

and

$$g(x+h) - g(x) = dg_x(h) + \theta_g(h).$$

Then we have

$$(f+g)(x+h) - (f+g)(x) = [df_x + dg_x](h) + [\theta_f(h) + \theta_g(h)],$$

and we may set $L = df_x + dg_x$ and $\theta = \theta_f + \theta_g$. Again, Equations (12.1) and (12.2) are satisfied, and $d(f+g)_x = df_x + dg_x$.

Part 4 is immediate.

To see part 5, suppose f is differentiable at x and that v is a vector in E. Then we have

$$\begin{aligned} df_x(v) &= \lim_{t\to 0} df_x(tv)/t \\ &= \lim_{t\to 0} \frac{f(x+tv) - f(x) - \theta(tv)}{t} \\ &= \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} + \lim_{t\to 0} \frac{\theta(tv)}{t} \\ &= \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t}, \end{aligned}$$

showing part 5.

Finally, if f is a map from a subset D of $\mathbb{R}$ into a Banach space F, and if f is differentiable at a point x, then we have from part 5 that

$$df_x(1) = \lim_{t\to 0} \frac{f(x+t) - f(x)}{t},$$

which proves that $f'(x) = df_x(1)$.

EXERCISE 12.2. Show that the following functions are differentiable at the indicated points, and verify that their differentials are as given below in parentheses.

(a) $f : B(H) \to B(H)$ is given by $f(T) = T^2$.
($df_T(S) = TS + ST.$)

(b) $f : B(H) \to B(H)$ is given by $f(T) = T^n$.
($df_T(S) = \sum_{j=0}^{n-1} T^j S T^{n-1-j}$.)

(c) f maps the invertible elements of $B(H)$ into themselves and is given by $f(T) = T^{-1}$.
($df_T(S) = -T^{-1}ST^{-1}$.)

(d) Let μ be a σ-finite measure, let p be an integer > 1, and let $f : L^p(\mu) \to L^1(\mu)$ be given by $f(g) = g^p$.
($df_g(h) = pg^{p-1}h$.)

(e) Suppose E, F, and G are Banach spaces, and let $f : E \times F \to G$ be continuous and bilinear.
($df_{x,y}(z, w) = f(x, w) + f(z, y)$.)

(f) Let E, F and G be Banach spaces, let D be a subset of E, let $f : D \to F$, let $g : D \to G$, and assume that f and g are differentiable at a point $x \in D$. Define $h : D \to F \oplus G$ by $h(y) = (f(y), g(y))$. Show that h is differentiable at x.
($dh_x(v) = (df_x(v), dg_x(v))$.)

EXERCISE 12.3. Suppose D is a subset of $\mathbb{R}^n$ and that $f : D \to \mathbb{R}^k$ is differentiable at a point $x \in D$. If we express each element of $\mathbb{R}^k$ in terms of the standard basis for $\mathbb{R}^k$, then we may write f in component form as $\{f_1, \dots, f_k\}$.

(a) Prove that each component function f_i of f is differentiable at x.

(b) If we express the linear transformation df_x as a matrix $J(x)$ with respect to the standard bases in $\mathbb{R}^n$ and $\mathbb{R}^k$, show that the ijth entry of $J(x)$ is the partial derivative of f_i with respect to the jth variable x_j evaluated at x. That is, show that

$$J(x)_{ij} = \frac{\partial f_i}{\partial x_j}(x).$$

The matrix $J(x)$ is called the *Jacobian* of f at x.

EXERCISE 12.4. Let A be a Banach algebra with identity I, and define $f : A \to A$ by $f(x) = e^x$.

(a) Prove that f is differentiable at 0, and compute df_0.

(b) Prove that f is differentiable at every $x \in A$, and compute $df_x(y)$ for arbitrary x and y.

THEOREM 12.3. *If $f : D \to F$ is differentiable at a point x, then f is continuous at x.*

PROOF. Suppose $\epsilon > 0$ is such that $B_\epsilon(x) \subseteq D$, and let y satisfy $0 < \|y - x\| < \epsilon$. Then

$$\begin{aligned}\|f(y) - f(x)\| &= \|f(x + (y - x)) - f(x)\| \\ &= \|df_x(y - x) + \theta(y - x)\| \\ &\leq \|df_x\|\|y - x\| + \|y - x\|\|\theta(y - x)\|/\|y - x\|,\end{aligned}$$

which tends to 0 as y tends to x. This shows the continuity of f at x.

THEOREM 12.4. (Chain Rule) *Let $E, F,$ and G be Banach spaces and let $D \subseteq E$ and $D' \subseteq F$. Suppose $f : D \to F$, that $g : D' \to G$, that f is differentiable at a point $x \in D$, and that g is differentiable at the point $f(x) \in D'$. Then the composition $g \circ f$ is differentiable at x, and*

$$d(g \circ f)_x = dg_{f(x)} \circ df_x.$$

PROOF. Write y for the point $f(x) \in D'$, and define the functions θ_f and θ_g by

$$f(x+h) - f(x) = df_x(h) + \theta_f(h), \tag{12.4}$$

and

$$g(y+k) - g(y) = dg_y(k) + \theta_g(k). \tag{12.5}$$

Let $\epsilon > 0$ be such that $B_\epsilon(y) \subseteq D'$, and let $\delta > 0$ be such that $B_\delta(x) \subseteq D$, that $f(B_\delta(x)) \subseteq B_\epsilon(y)$, and that

$$\|\theta_f(h)\|/\|h\| \leq 1 \tag{12.6}$$

if $\|h\| < \delta$. For $\|h\| < \delta$, define $k(h) = f(x+h) - f(x)$, and observe from Equations (12.4) and (12.6) that $\|k(h)\| \leq M\|h\|$, where $M = \|df_x\| + 1$.

To prove the chain rule, we must show that

$$\lim_{h \to 0} \frac{\|g(f(x+h)) - g(f(x)) - dg_{f(x)}(df_x(h))\|}{\|h\|} = 0.$$

But,

$$\begin{aligned}
&g(f(x+h)) - g(f(x)) - dg_{f(x)}(df_x(h)) \\
&= g(y + k(h)) - g(y) - dg_y(df_x(h)) \\
&= dg_y(k(h)) + \theta_g(k(h)) - dg_y(df_x(h)) \\
&= dg_y(f(x+h) - f(x)) - dg_y(df_x(h)) \\
&\quad + \theta_g(k(h)) \\
&= dg_y(\theta_f(h)) + \theta_g(k(h)),
\end{aligned}$$

so,

$$\|g(f(x+h)) - g(f(x)) - dg_{f(x)}(df_x(h))\| \leq \|dg_y\|\|\theta_f(h)\| + \|\theta_g(k(h))\|,$$

so that it will suffice to show that

$$\lim_{h \to 0} \|\theta_g(k(h))\|/\|h\| = 0.$$

If $k(h) = 0$, then $\|\theta_g(k(h))\|/\|h\| = 0$. Otherwise,

$$\begin{aligned}\frac{\|\theta_g(k(h))\|}{\|h\|} &= \frac{\|k(h)\|}{\|h\|}\frac{\|\theta_g(k(h))\|}{\|k(h)\|}\\ &\leq M\frac{\|\theta_g(k(h))\|}{\|k(h)\|},\end{aligned}$$

so we need only show that

$$\lim_{h\to 0}\frac{\|\theta_g(k(h))\|}{\|k(h)\|} = 0.$$

But, since f is continuous at x, we have that $k(h)$ approaches 0 as h approaches 0, so that the desired result follows from Equation (12.5).

EXERCISE 12.5. Let E, F, and G be Banach spaces, and let D be a subset of E.

(a) Let $f : D \to F$ and $g : D \to G$, and suppose B is a continuous bilinear map of $F \times G$ into a Banach space H. Define $p : D \to H$ by $p(y) = B(f(y), g(y))$. Assume that f and g are both differentiable at a point $x \in D$. Show that p is differentiable at x and compute $dp_x(y)$.

(b) Derive the "Product Formula" for differentials. That is, let A be a Banach algebra, let $f : D \to A$ and $g : D \to A$, and suppose both f and g are differentiable at a point $x \in D$. Show that the product function $f(y)g(y)$ is differentiable at x, and derive the formula for its differential.

(c) Suppose E is a Hilbert space and that $f : E \to \mathbb{R}$ is defined by $f(x) = \|x\|$. Prove that f is differentiable at every nonzero x.

(d) Let $E = L^1(\mathbb{R})$, and define $f : E \to \mathbb{R}$ by $f(x) = \|x\|_1$. Show that f is not differentiable at any point.

THEOREM 12.5. (First Derivative Test) *Let E be a Banach space, let D be a subset of E, and suppose $f : D \to \mathbb{R}$ is differentiable at a point $x \in D$. Assume that the point $f(x)$ is an extreme point of the set $f(D)$. Then df_x is the 0 linear transformation. That is, if a function achieves an extreme value at a point where it is differentiable, then the differential at that point must be 0.*

PROOF. Let v be a vector in E. Since x belongs to the interior of D, we let $\epsilon > 0$ be such that $x + tv \in D$ if $|t| < \epsilon$, and define a function $h : (-\epsilon, \epsilon) \to \mathbb{R}$ by $h(t) = f(x + tv)$. Then, by the chain rule, h is differentiable at 0. Furthermore, since $f(x)$ is an extreme point of the set $f(D)$, it follows that h attains either a local maximum or a local minimum at 0. From the first derivative test in elementary calculus, we then have that $h'(0) = dh_0(1) = 0$, implying that $df_x(v) = 0$. Since this is true for arbitrary elements $v \in E$, we see that $df_x = 0$.

THEOREM 12.6. (Mean Value Theorem) *Suppose E and F are Banach spaces, D is a subset of E, and $f : D \to F$. Suppose x and y are elements of D and that the closed line segment joining x and y is contained in D. Assume that f is continuous at each point of the closed line segment joining x to y, i.e., at each point $(1-t)x+ty$ for $0 \leq t \leq 1$, and assume that f is differentiable at each point on the open segment joining x and y, i.e., at each point $(1-t)x+ty$ for $0 < t < 1$. Then:*

(1) *There exists a $t^* \in (0,1)$ such that*

$$\|f(y) - f(x)\| \leq \|df_z(y-x)\| \leq \|df_z\|\|y-x\|,$$

for $z = (1-t^*)x + t^*y$.

(2) *If $F = \mathbb{R}$, then there exists a t^* in (0,1) such that*

$$f(y) - f(x) = df_z(y-x)$$

for $z = (1-t^*)x + t^*y$.

PROOF. Using the Hahn-Banach Theorem, choose ϕ in the conjugate space F^* of F so that $\|\phi\| = 1$ and

$$\|f(y) - f(x)\| = \phi(f(y) - f(x)).$$

Let h be the map of [0,1] into E defined by $h(t) = (1-t)x + ty$, and observe that

$$\|f(y) - f(x)\| = \phi(f(h(1))) - \phi(f(h(0))).$$

Defining $j = \phi \circ f \circ h$, we have from the chain rule that j is continuous on [0,1] and differentiable on (0,1). Then, using the Mean Value Theorem from elementary calculus, we have:

$$\begin{aligned}
\|f(y) - f(x)\| &= j(1) - j(0) \\
&= j'(t^*) \\
&= dj_{t^*}(1) \\
&= d(\phi \circ f \circ h)_{t^*}(1) \\
&= d\phi_{f(h(t^*))}(df_{h(t^*)}(dh_{t^*}(1))) \\
&= \phi(df_{h(t^*)}(dh_{t^*}(1))) \\
&= \phi(df_{h(t^*)}(y-x)),
\end{aligned}$$

whence

$$\begin{aligned}\|f(y)-f(x)\| &\leq \|\phi\|\|df_{h(t^*)}(y-x)\| \\ &= \|df_z(y-x)\|,\end{aligned}$$

as desired.

We leave the proof of part 2 to the exercises.

EXERCISE 12.6. (a) Prove part 2 of the preceding theorem.

(b) Define $f : [0,1] \to \mathbb{R}^2$ by

$$f(x) = (x^3, x^2).$$

Show that part 1 of the Mean Value Theorem cannot be strengthened to an equality. That is, show that there is no t^* between 0 and 1 satisfying $f(1) - f(0) = df_{t^*}(1)$.

(c) Define D to be the subset of $\mathbb{R}^2$ given by $0 \leq x \leq 1, 0 \leq y \leq 1$, and define $f : D \to \mathbb{R}^2$ by

$$f(x,y) = (y \cos x, y \sin x).$$

Show that every point $f(x,1)$ is an extreme point of the set $f(D)$ but that $df_{(x,1)} \neq 0$. Conclude that the first derivative test only works when the range space is $\mathbb{R}$.

DEFINITION. Let f be a map from a subset D of a Banach space E into a Banach space F. We say that f is *continuously differentiable* at a point x if f is differentiable at each point y in a neighborhood of x and if the map $y \to df_y$ is continuous at x. ($y \to df_y$ is a map from a neighborhood of $x \in E$ into the Banach space $L(E,F)$.)

The map f is *twice differentiable* at x if it is continuously differentiable at x and the map $y \to df_y$ is differentiable at x. The differential of this map $y \to df_y$ at the point x is denoted by $d^2 f_x$. The map f is *2 times continuously differentiable* at x if the map $y \to df_y$ is continuously differentiable at x.

The notions of n times continuously differentiable are defined by induction.

EXERCISE 12.7. (a) Let E and F be Banach spaces, let D be a subset of E, and suppose $f : D \to F$ is twice differentiable at a point $x \in D$. For each $v \in E$, show that $d^2 f_x(v)$ is an element of $L(E,F)$, whence for each pair (v,w) of elements in E, $[d^2 f_x(v)](w)$ is an element of F.

(b) Let f be as in part a. Show that d^2f_x represents a continuous bilinear map of $E \oplus E$ into F.

(c) Suppose f is a continuous linear transformation of E into F. Show that f is twice differentiable everywhere, and compute d^2f_x for any x.

(d) Suppose H is a Hilbert space, that $E = F = B(H)$ and that $f(T) = T^{-1}$. Show that f is twice differentiable at each invertible T, and compute d^2f_T.

THEOREM 12.7. (Theorem on Mixed Partials) *Suppose E and F are Banach spaces, D is a subset of E, and $f : D \to F$ is twice differentiable at each point of D. Suppose further that f is 2 times continuously differentiable at a point $x \in D$. Then*

$$[d^2f_x(v)](w) = [d^2f_x(w)](v);$$

i.e., the bilinear map d^2f_x is symmetric.

PROOF. Let v and w be in E, and let $\phi \in F^*$. Write $\phi = U + iV$ in its real and imaginary parts. Then

$$\begin{aligned}
&U([d^2f_x(v)](w)) \\
&= \lim_{t\to 0} U(\frac{[df_{x+tv} - df_x](w)}{t}) \\
&= \lim_{t\to 0}\lim_{s\to 0} U(\frac{f(x+tv+sw) - f(x+tv) - f(x+sw) + f(x)}{st}) \\
&= \lim_{t\to 0}\lim_{s\to 0} \frac{J_s(t) - J_s(0)}{st},
\end{aligned}$$

where $J_s(t) = U(f(x+sw+tv)-f(x+tv))$. Therefore, using the ordinary Mean Value Theorem on the real-valued function J_s, we have that

$$\begin{aligned}
U([d^2f_x(v)](w)) &= \lim_{t\to 0}\lim_{s\to 0} J_s'(t^*)/s \\
&= \lim_{t\to 0}\lim_{s\to 0} U(df_{x+sw+t^*v}(v) - df_{x+t^*v}(v))/s \\
&= \lim_{t\to 0}\lim_{s\to 0} U([df_{x+t^*v+sw} - df_{x+t^*v}](v))/s \\
&= \lim_{t\to 0} U([d^2f_{x+t^*v}(w)](v)) \\
&= U([d^2f_x(w)](v)),
\end{aligned}$$

because of the continuity of d^2f_y at $y = x$. A similar computation shows that

$$V([d^2f_x(v)](w)) = V([d^2f_x(w)](v)),$$

which implies that

$$\phi([d^2 f_x(v)](w)) = \phi([d^2 f_x(w)](v)).$$

This equality being valid for every $\phi \in F^*$ implies that

$$[d^2 f_x(v)](w) = [d^2 f_x(w)](v),$$

as desired.

EXERCISE 12.8. (Second Derivative Test) Let E and F be Banach spaces, let D be a subset of E, and suppose $f : D \to F$ is 2 times continuously differentiable at a point $x \in D$.

(a) Show that for each pair v, w of elements in E, the function

$$y \to [d^2 f_y(v)](w)$$

is continuous at x.

(b) Suppose $F = \mathbb{R}$, that f is 2 times continuously differentiable at x, that $df_x = 0$, and that the bilinear form $d^2 f_x$ is positive definite; i.e., there exists a $\delta > 0$ such that $[d^2 f_x(v)](v) \geq \delta$ for every unit vector $v \in E$. Prove that f attains a local minimum at x. That is, show that there exists an $\epsilon > 0$ such that if $\|y - x\| < \epsilon$ then $f(x) < f(y)$. HINT: Use the Mean Value Theorem twice to show that $f(y) - f(x) > 0$ for all y in a sufficiently small ball around x.

EXERCISE 12.9. Let (X, d) be a metric space. A map $\phi : X \to X$ is called a *contraction map* on X if there exists an α with $0 \leq \alpha < 1$ such that

$$d(\phi(x), \phi(y)) \leq \alpha d(x, y)$$

for all $x, y \in X$.

(a) If ϕ is a contraction map on (X, d), $x_0 \in X$, and $k < n$ are positive integers, show that

$$\begin{aligned} d(\phi^n(x_0), \phi^k(x_0)) &\leq \sum_{j=k}^{n-1} d(\phi^{j+1}(x_0), \phi^j(x_0)) \\ &\leq \sum_{j=k}^{n-1} \alpha^j d(\phi(x_0), x_0) \\ &= d(\phi(x_0), x_0)\alpha^k \frac{1 - \alpha^{n-k}}{1 - \alpha}, \end{aligned}$$

where ϕ^i denotes the composition of ϕ with itself i times.

(b) If ϕ is a contraction map on a complete metric space (X,d), and $x_0 \in X$, show that the sequence $\{\phi^n(x_0)\}$ has a limit in X.

(c) If ϕ is a contraction map on a complete metric space (X,d), and $x_0 \in X$, show that the limit y_0 of the sequence $\{\phi^n(x_0)\}$ is a fixed point of ϕ; i.e., $\phi(y_0) = y_0$.

(d) (Contraction mapping theorem) Show that a contraction map on a complete metric space (X,d) has one and only one fixed point y_0, and that $y_0 = \lim_n \phi^n(x)$ for each $x \in X$.

THEOREM 12.8. (Implicit Function Theorem) *Let E and F be Banach spaces, and equip $E \oplus F$ with the max norm. Let f be a map of an open subset O in $E \oplus F$ into F, and suppose f is continuously differentiable at a point $x = (x_1, x_2) \in O$. Assume further that the linear transformation $T : F \to F$, defined by $T(w) = df_x(0, w)$, is 1-1 and onto F. Then there exists a neighborhood U_1 of x_1 in E, a neighborhood U_2 of x_2 in F, and a unique continuous function $g : U_1 \to U_2$ such that*

(1) *The level set $f^{-1}(f(x)) \cap U$ coincides with the graph of g, where $U = U_1 \times U_2$.*
(2) *g is differentiable at x_1, and*

$$dg_{x_1}(h) = -T^{-1}(df_x(h, 0)).$$

PROOF. We will use the contraction mapping theorem. (See the previous exercise.) By the Isomorphism Theorem for continuous linear transformations on Banach spaces, we know that the inverse T^{-1} of T is an element of the Banach space $L(F, F)$. From the hypothesis of continuous differentiability at x, we may assume then that O is a sufficiently small neighborhood of x so that

$$\|df_z - df_x\| < 1/2\|T^{-1}\| \tag{12.7}$$

if $z \in O$. Write

$$f(x + h) - f(x) = df_x(h) + \theta(h).$$

We may assume also that O is sufficiently small so that

$$\|\theta(h)\| \leq \|h\|/2\|T^{-1}\| \tag{12.8}$$

if $x+h \in O$. Now there exist neighborhoods O_1 of x_1 and O_2 of x_2 such that $O_1 \times O_2 \subseteq O$. Choose $\epsilon > 0$ such that the closed ball $\bar{B}_\epsilon(x_2) \subseteq O_2$, and then choose $\delta > 0$ such that $B_\delta(x_1) \subseteq O_1$ and such that

$$\delta < \max(\epsilon, \epsilon/2\|T^{-1}\|\|df_x\|). \tag{12.9}$$

Set $U_1 = B_\delta(x_1),\ U_2 = \bar{B}_\epsilon(x_2)$, and $U = U_1 \times U_2$.

Let X be the set of all continuous functions from U_1 into U_2, and make X into a metric space by defining

$$d(g_1, g_2) = \sup_{v \in U_1} \|g_1(v) - g_2(v)\|.$$

Then, in fact, X is a complete metric space. (See the following exercise.)

Define a map ϕ, from X into the set of functions from U_1 into F, by

$$[\phi(g)](v) = g(v) - T^{-1}(f(v, g(v)) - f(x)).$$

Notice that each function $\phi(g)$ is continuous on U_1. Further, if $v \in U_1$, i.e., if $\|v - x_1\| < \delta$, then using inequalities (12.8) and (12.9) we have that

$$\begin{aligned}
&\|[\phi(g)](v) - x_2\| \\
&= \|g(v) - x_2 - T^{-1}(f(v, g(v)) - f(x))\| \\
&\leq \|T^{-1}\|\|T(g(v) - x_2) - f(v, g(v)) + f(x)\| \\
&= \|T^{-1}\| \\
&\quad \times \|df_x(0, g(v) - x_2) - df_x(v - x_1, g(v) - x_2) - \theta(v - x_1, g(v) - x_2)\| \\
&= \|T^{-1}\|\|df_x(v - x_1, 0) + \theta(v - x_1, g(v) - x_2)\| \\
&\leq \|T^{-1}\|\|df_x\|\delta + \|T^{-1}\|\|\theta(v - x_1, g(v) - x_2)\| \\
&< \|T^{-1}\|\|df_x\|\delta + \|(v - x_1, g(v) - x_2)\|/2 \\
&< \|T^{-1}\|\|df_x\|\delta + \max(\|v - x_1\|, \|g(v) - x_2\|)/2 \\
&< \|T^{-1}\|\|df_x\|\delta + \epsilon/2 \\
&< \epsilon,
\end{aligned}$$

showing that $\phi(g) \in X$.

Next, for $g_1, g_2 \in X$, we have:

$$\begin{aligned}
&d(\phi(g_1), \phi(g_2)) \\
&= \sup_{v \in U_1} \|g_1(v) - g_2(v) - T^{-1}(f(v, g_1(v)) - f(v, g_2(v)))\| \\
&\leq \sup_{v \in U_1} \|T^{-1}\| \\
&\qquad \times \|T(g_1(v) - g_2(v)) - [f(v, g_1(v)) - f(v, g_2(v))]\| \\
&= \sup_{v \in U_1} \|T^{-1}\| \\
&\qquad \times \|[T(g_1(v)) - f(v, g_1(v))] - [T(g_2(v)) - f(v, g_2(v))]\| \\
&\leq \sup_{v \in U_1} \|T^{-1}\| \\
&\qquad \times \|J^v(w_1) - J^v(w_2)\|,
\end{aligned}$$

where $w_i = g_i(v)$, and where J^v is the function defined on O_2 by

$$J^v(w) = T(w) - f(v, w).$$

So, by the Mean Value Theorem and inequality (12.7), we have

$$\begin{aligned}
d(\phi(g_1), \phi(g_2)) &\leq \sup_{v \in U_1} \|T^{-1}\| \|d(J^v)_z(w_1 - w_2)\| \\
&= \sup_{v \in U_1} \|T^{-1}\| \|[T - df_{(v,z)}](g_1(v) - g_2(v))\| \\
&\leq \sup_{v \in U_1} \|T^{-1}\| \|df_x - df_{(v,z)}\| \|g_1(v) - g_2(v)\| \\
&\leq d(g_1, g_2)/2,
\end{aligned}$$

showing that ϕ is a contraction mapping on X.

Let g be the unique fixed point of ϕ. Then, $\phi(g) = g$, whence $f(v, g(v)) = f(x)$ for all $v \in U_1$, which shows that the graph of g is contained in the level set $f^{-1}(f(x)) \cap U$. On the other hand, if $(v_0, w_0) \in U$ satisfies $f(v_0, w_0) = f(x)$, we may set $g_0(v) \equiv w_0$, and observe that $[\phi^n(g_0)](v_0) = w_0$ for all n. Therefore, the unique fixed point g of ϕ must satisfy $g(v_0) = w_0$, because $g = \lim \phi^n(g_0)$. Hence, any element (v_0, w_0) of the level set $f^{-1}(f(x)) \cap U$ belongs to the graph of g.

Finally, to see that g is differentiable at x_1 and has the prescribed differential, it will suffice to show that

$$\lim_{h \to 0} \|g(x_1 + h) - g(x_1) + T^{-1}(df_x(h, 0))\| / \|h\| = 0.$$

Now, because

$$f(x_1+h, x_2+(g(x_1+h)-x_2))-f(x_1,x_2)=0,$$

we have that

$$0=df_x(h,0)+df_x(0,g(x_1+h)-x_2)+\theta(h,g(x_1+h)-x_2),$$

or

$$g(x_1+h)-g(x_1)=-T^{-1}(df_x(h,0))-T^{-1}(\theta(h,g(x_1+h)-g(x_1))).$$

Hence, there exists a constant $M\geq 1$ such that

$$\|g(x_1+h)-g(x_1)\|\leq M\|h\|$$

whenever $x_1+h\in U_1$. (How?) But then

$$\begin{aligned}&\frac{\|g(x_1+h)-g(x_1)+T^{-1}(df_x(h,0))\|}{\|h\|}\\ \leq\ &\frac{\|T^{-1}\|\|\theta(h,g(x_1+h)-g(x_1))\|}{\|h\|}\\ \leq\ &\frac{\|T^{-1}\|M\|\theta(h,g(x_1+h)-g(x_1))\|}{\|(h,g(x_1+h)-g(x_1))\|},\end{aligned}$$

and this tends to 0 as h tends to 0 since g is continuous at x_1.

This completes the proof.

EXERCISE 12.10. Verify that the set X used in the preceding proof is a complete metric space with respect to the function d defined there.

THEOREM 12.9. (Inverse Function Theorem) *Let f be a mapping from an open subset O of a Banach space E into E, and assume that f is continuously differentiable at a point $x\in O$. Suppose further that the differential df_x of f at x is 1-1 from E onto E. Then there exist neighborhoods O_1 of x and O_2 of $f(x)$ such that f is a homeomorphism of O_1 onto O_2. Further, the inverse f^{-1} of the restriction of f to O_1 is differentiable at the point $f(x)$, whence*

$$d(f^{-1})_{f(x)}=(df_x)^{-1}.$$

PROOF. Define a map $J : E \times O \to E$ by $J(v, w) = v - f(w)$. Then J is continuously differentiable at the point $(f(x), x)$, and

$$dJ_{(f(x),x)}(0, y) = -df_x(y),$$

which is 1-1 from E onto E. Applying the implicit function theorem to J, there exist neighborhoods U_1 of the point $f(x)$, U_2 of the point x, and a continuous function $g : U_1 \to U_2$ whose graph coincides with the level set $J^{-1}(0) \cap (U_1 \times U_2)$. But this level set consists precisely of the pairs (v, w) in $U_1 \times U_2$ for which $v = f(w)$, while the graph of g consists precisely of the pairs (v, w) in $U_1 \times U_2$ for which $w = g(v)$. Clearly, then, g is the inverse of the restriction of f to U_2. Setting $O_1 = U_2$ and $O_2 = U_1$ gives the first part of the theorem. Also, from the implicit function theorem, $g = f^{-1}$ is differentiable at $f(x)$, and then the fact that $d(f^{-1})_{f(x)} = (df_x)^{-1}$ follows directly from the chain rule.

EXERCISE 12.11. Let H be a Hilbert space and let $E = B(H)$.

(a) Show that the exponential map $T \to e^T$ is 1-1 from a neighborhood $U = B_\epsilon(0)$ of 0 onto a neighborhood V of I.

(b) Let U and V be as in part a. Show that, for $T \in U$, we have e^T is a positive operator if and only if T is selfadjoint, and e^T is unitary if and only if T is *skewadjoint,* i.e., $T^* = -T$.

THEOREM 12.10. (Foliated Implicit Function Theorem) *Let E and F be Banach spaces, let O be an open subset of $E \times F$, and let $f : O \to F$ be continuously differentiable at every point $y \in O$. Suppose $x = (x_1, x_2)$ is a point in O for which the map $w \to df_x(0, w)$ is 1-1 from F onto F. Then there exist neighborhoods U_1 of x_1, U_2 of $f(x)$, U of x, and a diffeomorphism $J : U_1 \times U_2 \to U$ such that $J(U_1 \times \{z\})$ coincides with the level set $f^{-1}(z) \cap U$ for all $z \in U_2$.*

PROOF. For each $y \in O$, define $T_y : F \to F$ by $T_y(w) = df_y(0, w)$. Because T_x is an invertible element in $L(F, F)$, and because f is continuously differentiable at x, we may assume that O is small enough so that T_y is 1-1 and onto for every $y \in O$.

Define $h : O \to E \times F$ by

$$h(y) = h(y_1, y_2) = (y_1, f(y)).$$

Observe that h is continuously differentiable on O, and that

$$dh_x(v, w) = (v, df_x(v, w)),$$

whence, if $dh_x(v_1, w_1) = dh_x(v_2, w_2)$, then $v_1 = v_2$. But then $df_x(0, w_1 - w_2) = 0$, implying that $w_1 = w_2$, and therefore dh_x is 1-1 from $E \times F$ into $E \times F$. The exercise that follows this proof shows that dh_x is also onto, so we may apply the inverse function theorem to h. Thus, there exist neighborhoods O_1 of x and O_2 of $h(x)$ such that h is a homeomorphism of O_1 onto O_2. Now, there exist neighborhoods U_1 of x_1 and U_2 of $f(x)$ such that $U_1 \times U_2 \subseteq O_2$, and we define U to be the neighborhood $h^{-1}(U_1 \times U_2)$ of x. Define J to be the restriction of h^{-1} to $U_1 \times U_2$. Just as in the above argument for dh_x, we see that dh_y is 1-1 and onto if $y \in U$, whence, again by the inverse function theorem, J is differentiable at each point of its domain and is therefore a diffeomorphism of $U_1 \times U_2$ onto U.

We leave the last part of the proof to the following exercise.

EXERCISE 12.12. (a) Show that the linear transformation dh_x of the preceding proof is onto.

(b) Prove the last part of Theorem 12.10; i.e., show that $J(U_1 \times \{z\})$ coincides with the level set $f^{-1}(z) \cap U$.

We close this chapter with some exercises that examine the important special case when the Banach space E is actually a (real) Hilbert space.

EXERCISE 12.13. (Implicit Function Theorem in Hilbert Space) Suppose E is a Hilbert space, F is a Banach space, D is a subset of E, $f : D \to F$ is continuously differentiable on D, and that the differential df_x maps E onto F for each $x \in D$. Let c be an element of the range of f, let S denote the level set $f^{-1}(c)$, let x be in S, and write M for the kernel of df_x. Prove that there exists a neighborhood U_x of $0 \in M$, a neighborhood V_x of $x \in E$, and a continuously differentiable 1-1 function $g_x : U_x \to V_x$ such that the range of g_x coincides with the intersection $V_x \cap S$ of V_x and S. HINT: Write $E = M \oplus M^{\perp}$. Show also that $d(g_x)_0(h) = h$. We say that the level set $S = f^{-1}(c)$ is *locally parameterized* by an open subset of M.

DEFINITION. Suppose E is a Hilbert space, F is a Banach space, D is a subset of E, $f : D \to F$ is continuously differentiable on D, and that the differential df_x maps E onto F for each $x \in D$. Let c be an element of the range of f, and let S denote the level set $f^{-1}(c)$. We say that S is a *differentiable manifold*, and if $x \in S$, then a vector $v \in E$ is called a *tangent vector* to S at x if there exists an $\epsilon > 0$ and a continuously differentiable function $\phi : [-\epsilon, \epsilon] \to S \subseteq E$ such that $\phi(0) = x$ and $\phi'(0) = v$.

EXERCISE 12.14. Let x be a point in a differentiable manifold S, and write M for the kernel of df_x. Prove that v is a tangent vector to S at x if and only if $v \in M$. HINT: If $v \in M$, use Exercise 12.13 to define $\phi(t) = g_x(tv)$.

DEFINITION. Let D be a subset of a Banach space E, and suppose $f : D \to \mathbb{R}$ is differentiable at a point $x \in D$. We identify the conjugate space $\mathbb{R}^*$ with $\mathbb{R}$. By the *gradient* of f at x we mean the element of E^* defined by $\operatorname{grad} f(x) = df_x^*(1)$, where df_x^* denotes the adjoint of the continuous linear transformation df_x.

If E is a Hilbert space, then $\operatorname{grad} f(x)$ can by the Riesz representation theorem for Hilbert spaces be identified with an element of $E \equiv E^*$.

EXERCISE 12.15. Let S be a manifold in a Hilbert space E, and let g be a real-valued function that is differentiable at each point of an open set D that contains S. Suppose $x \in S$ is such that $g(x) \geq g(y)$ for all $y \in S$, and write $M = \ker(df_x)$. Prove that the vector $\operatorname{grad} g(x)$ is orthogonal to M.

EXERCISE 12.16. (Method of Lagrange Multipliers) Let E be a Hilbert space, let D be an open subset of E, let $f = \{f_1, \ldots , f_n\} : D \to \mathbb{R}^n$ be continuously differentiable at each point of D, and assume that each differential df_x for $x \in D$ maps onto $\mathbb{R}^n$. Let S be the level set $f^{-1}(c)$ for $c \in \mathbb{R}^n$. Suppose g is a real-valued differentiable function on D and that g attains a maximum on S at the point x. Prove that there exist real constants $\{\lambda_1, \ldots , \lambda_n\}$ such that

$$\operatorname{grad} g(x) = \sum_{i=1}^{n} \lambda_i \operatorname{grad} f_i(x).$$

The constants $\{\lambda_i\}$ are called the *Lagrange multipliers*.

EXERCISE 12.17. Let S be the unit sphere in $L^2([0, 1])$; i.e., S is the manifold consisting of the functions $f \in L^2([0, 1])$ for which $\|f\|_2 = 1$.

(a) Define g on S by $g(f) = \int_0^1 f(x)\,dx$. Use the method of Lagrange multipliers to find all points where g attains its maximum value on S.

(b) Define g on S by $g(f) = \int_0^1 |f|^{3/2}(x)\,dx$. Find the maximum value of g on S.