for some $\theta_j \in (0, 1)$, and this differs from $h_j(D_j f)(\mathbf{x})$ by less than $|h_j|\varepsilon/n$, using (41). By (42), it follows that

$$\left| f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - \sum_{j=1}^{n} h_j(D_j f)(\mathbf{x}) \right| \le \frac{1}{n} \sum_{j=1}^{n} |h_j| \varepsilon \le |\mathbf{h}| \varepsilon$$

for all $\mathbf{h}$ such that $|\mathbf{h}| < r$.

This says that f is differentiable at $\mathbf{x}$ and that $f'(\mathbf{x})$ is the linear function which assigns the number $\Sigma h_j(D_j f)(\mathbf{x})$ to the vector $\mathbf{h} = \Sigma h_j \mathbf{e}_j$. The matrix $[f'(\mathbf{x})]$ consists of the row $(D_1 f)(\mathbf{x}), \ldots, (D_n f)(\mathbf{x})$; and since $D_1 f, \ldots, D_n f$ are continuous functions on E, the concluding remarks of Sec. 9.9 show that $f \in \mathscr{C}'(E)$.

THE CONTRACTION PRINCIPLE

We now interrupt our discussion of differentiation to insert a fixed point theorem that is valid in arbitrary complete metric spaces. It will be used in the proof of the inverse function theorem.

9.22 Definition Let X be a metric space, with metric d. If φ maps X into X and if there is a number $c < 1$ such that

$$d(\varphi(x), \varphi(y)) \le c\, d(x, y) \tag{43}$$

for all $x, y \in X$, then φ is said to be a *contraction* of X into X.

9.23 Theorem *If X is a complete metric space, and if φ is a contraction of X into X, then there exists one and only one $x \in X$ such that $\varphi(x) = x$.*

In other words, φ has a unique fixed point. The uniqueness is a triviality, for if $\varphi(x) = x$ and $\varphi(y) = y$, then (43) gives $d(x, y) \le c\, d(x, y)$, which can only happen when $d(x, y) = 0$.

The *existence* of a fixed point of φ is the essential part of the theorem. The proof actually furnishes a constructive method for locating the fixed point.

Proof Pick $x_0 \in X$ arbitrarily, and define $\{x_n\}$ recursively, by setting

$$x_{n+1} = \varphi(x_n) \qquad (n = 0, 1, 2, \ldots). \tag{44}$$

Choose $c < 1$ so that (43) holds. For $n \ge 1$ we then have

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \le c\, d(x_n, x_{n-1}).$$

Hence induction gives

$$d(x_{n+1}, x_n) \le c^n\, d(x_1, x_0) \qquad (n = 0, 1, 2, \ldots). \tag{45}$$

If $n < m$, it follows that

$$\begin{aligned} d(x_n, x_m) &\le \sum_{i=n+1}^{m} d(x_i, x_{i-1}) \\ &\le (c^n + c^{n+1} + \cdots + c^{m-1})\, d(x_1, x_0) \\ &\le [(1-c)^{-1}\, d(x_1, x_0)]c^n. \end{aligned}$$

Thus $\{x_n\}$ is a Cauchy sequence. Since X is complete, $\lim x_n = x$ for some $x \in X$.

Since φ is a contraction, φ is continuous (in fact, uniformly continuous) on X. Hence

$$\varphi(x) = \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

THE INVERSE FUNCTION THEOREM

The inverse function theorem states, roughly speaking, that a continuously differentiable mapping $\mathbf{f}$ is invertible in a neighborhood of any point $\mathbf{x}$ at which the linear transformation $\mathbf{f}'(\mathbf{x})$ is invertible:

9.24 Theorem *Suppose* $\mathbf{f}$ *is a* $\mathscr{C}'$*-mapping of an open set* $E \subset R^n$ *into* R^n, $\mathbf{f}'(\mathbf{a})$ *is invertible for some* $\mathbf{a} \in E$, *and* $\mathbf{b} = \mathbf{f}(\mathbf{a})$. *Then*

(*a*) *there exist open sets* U *and* V *in* R^n *such that* $\mathbf{a} \in U$, $\mathbf{b} \in V$, $\mathbf{f}$ *is one-to-one on* U, *and* $\mathbf{f}(U) = V$;

(*b*) *if* $\mathbf{g}$ *is the inverse of* $\mathbf{f}$ [*which exists, by* (*a*)], *defined in* V *by*

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x} \qquad (\mathbf{x} \in U),$$

then $\mathbf{g} \in \mathscr{C}'(V)$.

Writing the equation $\mathbf{y} = \mathbf{f}(\mathbf{x})$ in component form, we arrive at the following interpretation of the conclusion of the theorem: The system of n equations

$$y_i = f_i(x_1, \ldots, x_n) \qquad (1 \le i \le n)$$

can be solved for $x_1, \ldots, x_n$ in terms of $y_1, \ldots, y_n$, if we restrict $\mathbf{x}$ and $\mathbf{y}$ to small enough neighborhoods of $\mathbf{a}$ and $\mathbf{b}$; the solutions are unique and continuously differentiable.

Proof

(*a*) Put $\mathbf{f}'(\mathbf{a}) = A$, and choose λ so that

$$2\lambda \|A^{-1}\| = 1. \tag{46}$$

Since $\mathbf{f}'$ is continuous at $\mathbf{a}$, there is an open ball $U \subset E$, with center at $\mathbf{a}$, such that

(47) $$\|\mathbf{f}'(\mathbf{x}) - A\| < \lambda \qquad (\mathbf{x} \in U).$$

We associate to each $\mathbf{y} \in R^n$ a function φ, defined by

(48) $$\varphi(\mathbf{x}) = \mathbf{x} + A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x})) \qquad (\mathbf{x} \in E).$$

Note that $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ *if and only if* $\mathbf{x}$ *is a fixed point of* φ.

Since $\varphi'(\mathbf{x}) = I - A^{-1}\mathbf{f}'(\mathbf{x}) = A^{-1}(A - \mathbf{f}'(\mathbf{x}))$, (46) and (47) imply that

(49) $$\|\varphi'(\mathbf{x})\| < \tfrac{1}{2} \qquad (\mathbf{x} \in U).$$

Hence

(50) $$|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)| \le \tfrac{1}{2}|\mathbf{x}_1 - \mathbf{x}_2| \qquad (\mathbf{x}_1, \mathbf{x}_2 \in U),$$

by Theorem 9.19. It follows that φ has at most one fixed point in U, so that $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ for at most one $\mathbf{x} \in U$.

Thus $\mathbf{f}$ *is* $1 - 1$ *in* U.

Next, put $V = \mathbf{f}(U)$, and pick $\mathbf{y}_0 \in V$. Then $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ for some $\mathbf{x}_0 \in U$. Let B be an open ball with center at $\mathbf{x}_0$ and radius $r > 0$, so small that its closure $\bar{B}$ lies in U. We will show that $\mathbf{y} \in V$ whenever $|\mathbf{y} - \mathbf{y}_0| < \lambda r$. This proves, of course, that V is open.

Fix $\mathbf{y}$, $|\mathbf{y} - \mathbf{y}_0| < \lambda r$. With φ as in (48),

$$|\varphi(\mathbf{x}_0) - \mathbf{x}_0| = |A^{-1}(\mathbf{y} - \mathbf{y}_0)| < \|A^{-1}\|\lambda r = \frac{r}{2}.$$

If $\mathbf{x} \in \bar{B}$, it therefore follows from (50) that

$$\begin{aligned}|\varphi(\mathbf{x}) - \mathbf{x}_0| &\le |\varphi(\mathbf{x}) - \varphi(\mathbf{x}_0)| + |\varphi(\mathbf{x}_0) - \mathbf{x}_0| \\ &< \frac{1}{2}|\mathbf{x} - \mathbf{x}_0| + \frac{r}{2} \le r;\end{aligned}$$

hence $\varphi(\mathbf{x}) \in B$. Note that (50) holds if $\mathbf{x}_1 \in \bar{B}$, $\mathbf{x}_2 \in \bar{B}$.

Thus φ is a contraction of $\bar{B}$ into $\bar{B}$. Being a closed subset of R^n, $\bar{B}$ is complete. Theorem 9.23 implies therefore that φ has a fixed point $\mathbf{x} \in \bar{B}$. For this $\mathbf{x}$, $f(\mathbf{x}) = \mathbf{y}$. Thus $\mathbf{y} \in \mathbf{f}(\bar{B}) \subset \mathbf{f}(U) = V$.

This proves part (*a*) of the theorem.

(*b*) Pick $\mathbf{y} \in V$, $\mathbf{y} + \mathbf{k} \in V$. Then there exist $\mathbf{x} \in U$, $\mathbf{x} + \mathbf{h} \in U$, so that $\mathbf{y} = \mathbf{f}(\mathbf{x})$, $\mathbf{y} + \mathbf{k} = \mathbf{f}(\mathbf{x} + \mathbf{h})$. With φ as in (48),

$$\varphi(\mathbf{x} + \mathbf{h}) - \varphi(\mathbf{x}) = \mathbf{h} + A^{-1}[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x} + \mathbf{h})] = \mathbf{h} - A^{-1}\mathbf{k}.$$

By (50), $|\mathbf{h} - A^{-1}\mathbf{k}| \le \frac{1}{2}|\mathbf{h}|$. Hence $|A^{-1}\mathbf{k}| \ge \frac{1}{2}|\mathbf{h}|$, and

(51) $$|\mathbf{h}| \le 2\|A^{-1}\|\,|\mathbf{k}| = \lambda^{-1}|\mathbf{k}|.$$

By (46), (47), and Theorem 9.8, $\mathbf{f}'(\mathbf{x})$ has an inverse, say T. Since

$$\mathbf{g}(\mathbf{y}+\mathbf{k})-\mathbf{g}(\mathbf{y})-T\mathbf{k}=\mathbf{h}-T\mathbf{k}=-T[\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x})-\mathbf{f}'(\mathbf{x})\mathbf{h}],$$

(51) implies

$$\frac{|\mathbf{g}(\mathbf{y}+\mathbf{k})-\mathbf{g}(\mathbf{y})-T\mathbf{k}|}{|\mathbf{k}|}\le\frac{\|T\|}{\lambda}\cdot\frac{|\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x})-\mathbf{f}'(\mathbf{x})\mathbf{h}|}{|\mathbf{h}|}.$$

As $\mathbf{k}\to\mathbf{0}$, (51) shows that $\mathbf{h}\to\mathbf{0}$. The right side of the last inequality thus tends to 0. Hence the same is true of the left. We have thus proved that $\mathbf{g}'(\mathbf{y})=T$. But T was chosen to be the inverse of $\mathbf{f}'(\mathbf{x})=\mathbf{f}'(\mathbf{g}(\mathbf{y}))$. Thus

$$\mathbf{g}'(\mathbf{y})=\{\mathbf{f}'(\mathbf{g}(\mathbf{y}))\}^{-1}\qquad(\mathbf{y}\in V). \tag{52}$$

Finally, note that $\mathbf{g}$ is a continuous mapping of V onto U (since $\mathbf{g}$ is differentiable), that $\mathbf{f}'$ is a continuous mapping of U into the set Ω of all invertible elements of $L(R^n)$, and that inversion is a continuous mapping of Ω onto Ω, by Theorem 9.8. If we combine these facts with (52), we see that $\mathbf{g}\in\mathscr{C}'(V)$.

This completes the proof.

Remark. The full force of the assumption that $\mathbf{f}\in\mathscr{C}'(E)$ was only used in the last paragraph of the preceding proof. Everything else, down to Eq. (52), was derived from the existence of $\mathbf{f}'(\mathbf{x})$ for $\mathbf{x}\in E$, the invertibility of $\mathbf{f}'(\mathbf{a})$, and the continuity of $\mathbf{f}'$ at just the point $\mathbf{a}$. In this connection, we refer to the article by A. Nijenhuis in *Amer. Math. Monthly*, vol. 81, 1974, pp. 969–980.

The following is an immediate consequence of part (*a*) of the inverse function theorem.

9.25 Theorem *If $\mathbf{f}$ is a $\mathscr{C}'$-mapping of an open set $E\subset R^n$ into R^n and if $\mathbf{f}'(\mathbf{x})$ is invertible for every $\mathbf{x}\in E$, then $\mathbf{f}(W)$ is an open subset of R^n for every open set $W\subset E$.*

In other words, $\mathbf{f}$ is an *open mapping* of E into R^n.

The hypotheses made in this theorem ensure that each point $\mathbf{x}\in E$ has a neighborhood in which $\mathbf{f}$ is 1-1. This may be expressed by saying that $\mathbf{f}$ is *locally* one-to-one in E. But $\mathbf{f}$ need not be 1-1 in E under these circumstances. For an example, see Exercise 17.

THE IMPLICIT FUNCTION THEOREM

If f is a continuously differentiable real function in the plane, then the equation $f(x, y)=0$ can be solved for y in terms of x in a neighborhood of any point

(a, b) at which $f(a, b) = 0$ and $\partial f/\partial y \neq 0$. Likewise, one can solve for x in terms of y near (a, b) if $\partial f/\partial x \neq 0$ at (a, b). For a simple example which illustrates the need for assuming $\partial f/\partial y \neq 0$, consider $f(x, y) = x^2 + y^2 - 1$.

The preceding very informal statement is the simplest case (the case $m = n = 1$ of Theorem 9.28) of the so-called "implicit function theorem." Its proof makes strong use of the fact that continuously differentiable transformations behave locally very much like their derivatives. Accordingly, we first prove Theorem 9.27, the linear version of Theorem 9.28.

9.26 Notation If $\mathbf{x} = (x_1, \ldots, x_n) \in R^n$ and $\mathbf{y} = (y_1, \ldots, y_m) \in R^m$, let us write $(\mathbf{x}, \mathbf{y})$ for the point (or vector)

$$(x_1, \ldots, x_n, y_1, \ldots, y_m) \in R^{n+m}.$$

In what follows, the first entry in $(\mathbf{x}, \mathbf{y})$ or in a similar symbol will always be a vector in R^n, the second will be a vector in R^m.

Every $A \in L(R^{n+m}, R^n)$ can be split into two linear transformations A_x and A_y, defined by

$$A_x \mathbf{h} = A(\mathbf{h}, \mathbf{0}), \qquad A_y \mathbf{k} = A(\mathbf{0}, \mathbf{k}) \tag{53}$$

for any $\mathbf{h} \in R^n$, $\mathbf{k} \in R^m$. Then $A_x \in L(R^n)$, $A_y \in L(R^m, R^n)$, and

$$A(\mathbf{h}, \mathbf{k}) = A_x \mathbf{h} + A_y \mathbf{k}. \tag{54}$$

The linear version of the implicit function theorem is now almost obvious.

9.27 Theorem *If $A \in L(R^{n+m}, R^n)$ and if A_x is invertible, then there corresponds to every $\mathbf{k} \in R^m$ a unique $\mathbf{h} \in R^n$ such that $A(\mathbf{h}, \mathbf{k}) = \mathbf{0}$.*

This $\mathbf{h}$ can be computed from $\mathbf{k}$ by the formula

$$\mathbf{h} = -(A_x)^{-1} A_y \mathbf{k}. \tag{55}$$

Proof By (54), $A(\mathbf{h}, \mathbf{k}) = \mathbf{0}$ if and only if

$$A_x \mathbf{h} + A_y \mathbf{k} = \mathbf{0},$$

which is the same as (55) when A_x is invertible.

The conclusion of Theorem 9.27 is, in other words, that the equation $A(\mathbf{h}, \mathbf{k}) = \mathbf{0}$ can be solved (uniquely) for $\mathbf{h}$ if $\mathbf{k}$ is given, and that the solution $\mathbf{h}$ is a linear function of $\mathbf{k}$. Those who have some acquaintance with linear algebra will recognize this as a very familiar statement about systems of linear equations.

9.28 Theorem *Let $\mathbf{f}$ be a $\mathscr{C}'$-mapping of an open set $E \subset R^{n+m}$ into R^n, such that $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ for some point $(\mathbf{a}, \mathbf{b}) \in E$.*

Put $A = \mathbf{f}'(\mathbf{a}, \mathbf{b})$ and assume that A_x is invertible.

Then there exist open sets $U \subset R^{n+m}$ *and* $W \subset R^m$, *with* $(\mathbf{a}, \mathbf{b}) \in U$ *and* $\mathbf{b} \in W$, *having the following property:*

To every $\mathbf{y} \in W$ *corresponds a unique* $\mathbf{x}$ *such that*

$$(\mathbf{x}, \mathbf{y}) \in U \qquad \text{and} \qquad \mathbf{f}(\mathbf{x}, \mathbf{y}) = 0. \tag{56}$$

If this $\mathbf{x}$ *is defined to be* $\mathbf{g}(\mathbf{y})$, *then* $\mathbf{g}$ *is a* $\mathscr{C}'$*-mapping of* W *into* R^n, $\mathbf{g}(\mathbf{b}) = \mathbf{a}$,

$$\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{0} \qquad (\mathbf{y} \in W), \tag{57}$$

and

$$\mathbf{g}'(\mathbf{b}) = -(A_x)^{-1}A_y. \tag{58}$$

The function g is "implicitly" defined by (57). Hence the name of the theorem.

The equation $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ can be written as a system of n equations in $n + m$ variables:

$$\begin{aligned} &f_1(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &f_n(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0. \end{aligned} \tag{59}$$

The assumption that A_x is invertible means that the n by n matrix

$$\begin{bmatrix} D_1 f_1 & \cdots & D_n f_1 \\ \cdots & \cdots & \cdots \\ D_1 f_n & \cdots & D_n f_n \end{bmatrix}$$

evaluated at $(\mathbf{a}, \mathbf{b})$ defines an invertible linear operator in R^n; in other words, its column vectors should be independent, or, equivalently, its determinant should be $\neq 0$. (See Theorem 9.36.) If, furthermore, (59) holds when $\mathbf{x} = \mathbf{a}$ and $\mathbf{y} = \mathbf{b}$, then the conclusion of the theorem is that (59) can be solved for $x_1, \ldots, x_n$ in terms of $y_1, \ldots, y_m$, for every $\mathbf{y}$ near $\mathbf{b}$, and that these solutions are continuously differentiable functions of $\mathbf{y}$.

Proof Define $\mathbf{F}$ by

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y}) \qquad ((\mathbf{x}, \mathbf{y}) \in E). \tag{60}$$

Then $\mathbf{F}$ is a $\mathscr{C}'$-mapping of E into R^{n+m}. We claim that $\mathbf{F}'(\mathbf{a}, \mathbf{b})$ is an invertible element of $L(R^{n+m})$:

Since $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$, we have

$$\mathbf{f}(\mathbf{a} + \mathbf{h}, \mathbf{b} + \mathbf{k}) = A(\mathbf{h}, \mathbf{k}) + \mathbf{r}(\mathbf{h}, \mathbf{k}),$$

where $\mathbf{r}$ is the remainder that occurs in the definition of $\mathbf{f}'(\mathbf{a}, \mathbf{b})$. Since

$$\begin{aligned} \mathbf{F}(\mathbf{a} + \mathbf{h}, \mathbf{b} + \mathbf{k}) - \mathbf{F}(\mathbf{a}, \mathbf{b}) &= (\mathbf{f}(\mathbf{a} + \mathbf{h}, \mathbf{b} + \mathbf{k}), \mathbf{k}) \\ &= (\mathbf{A}(\mathbf{h}, \mathbf{k}), \mathbf{k}) + (\mathbf{r}(\mathbf{h}, \mathbf{k}), \mathbf{0}) \end{aligned}$$

it follows that $\mathbf{F}'(\mathbf{a}, \mathbf{b})$ is the linear operator on R^{n+m} that maps $(\mathbf{h}, \mathbf{k})$ to $(A(\mathbf{h}, \mathbf{k}), \mathbf{k})$. If this image vector is $\mathbf{0}$, then $A(\mathbf{h}, \mathbf{k}) = \mathbf{0}$ and $\mathbf{k} = \mathbf{0}$, hence $A(\mathbf{h}, \mathbf{0}) = \mathbf{0}$, and Theorem 9.27 implies that $\mathbf{h} = \mathbf{0}$. It follows that $\mathbf{F}'(\mathbf{a}, \mathbf{b})$ is 1-1; hence it is invertible (Theorem 9.5).

The inverse function theorem can therefore be applied to $\mathbf{F}$. It shows that there exist open sets U and V in R^{n+m}, with $(\mathbf{a}, \mathbf{b}) \in U$, $(\mathbf{0}, \mathbf{b}) \in V$, such that $\mathbf{F}$ is a 1-1 mapping of U onto V.

We let W be the set of all $\mathbf{y} \in R^m$ such that $(\mathbf{0}, \mathbf{y}) \in V$. Note that $\mathbf{b} \in W$.

It is clear that W is open since V is open.

If $\mathbf{y} \in W$, then $(\mathbf{0}, \mathbf{y}) = \mathbf{F}(\mathbf{x}, \mathbf{y})$ for some $(\mathbf{x}, \mathbf{y}) \in U$. By (60), $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ for this $\mathbf{x}$.

Suppose, with the same $\mathbf{y}$, that $(\mathbf{x}', \mathbf{y}) \in U$ and $\mathbf{f}(\mathbf{x}', \mathbf{y}) = \mathbf{0}$. Then

$$\mathbf{F}(\mathbf{x}', \mathbf{y}) = (\mathbf{f}(\mathbf{x}', \mathbf{y}), \mathbf{y}) = (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y}) = \mathbf{F}(\mathbf{x}, \mathbf{y}).$$

Since $\mathbf{F}$ is 1-1 in U, it follows that $\mathbf{x}' = \mathbf{x}$.

This proves the first part of the theorem.

For the second part, define $\mathbf{g}(\mathbf{y})$, for $\mathbf{y} \in W$, so that $(\mathbf{g}(\mathbf{y}), \mathbf{y}) \in U$ and (57) holds. Then

(61) $$\mathbf{F}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = (\mathbf{0}, \mathbf{y}) \qquad (\mathbf{y} \in W).$$

If $\mathbf{G}$ is the mapping of V onto U that inverts $\mathbf{F}$, then $\mathbf{G} \in \mathscr{C}'$, by the inverse function theorem, and (61) gives

(62) $$(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{G}(\mathbf{0}, \mathbf{y}) \qquad (\mathbf{y} \in W).$$

Since $\mathbf{G} \in \mathscr{C}'$, (62) shows that $\mathbf{g} \in \mathscr{C}'$.

Finally, to compute $\mathbf{g}'(\mathbf{b})$, put $(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \Phi(\mathbf{y})$. Then

(63) $$\Phi'(\mathbf{y})\mathbf{k} = (\mathbf{g}'(\mathbf{y})\mathbf{k}, \mathbf{k}) \qquad (\mathbf{y} \in W, \mathbf{k} \in R^m).$$

By (57), $\mathbf{f}(\Phi(\mathbf{y})) = \mathbf{0}$ in W. The chain rule shows therefore that

$$\mathbf{f}'(\Phi(\mathbf{y}))\Phi'(\mathbf{y}) = 0.$$

When $\mathbf{y} = \mathbf{b}$, then $\Phi(\mathbf{y}) = (\mathbf{a}, \mathbf{b})$, and $\mathbf{f}'(\Phi(\mathbf{y})) = A$. Thus

(64) $$A\Phi'(\mathbf{b}) = 0.$$

It now follows from (64), (63), and (54), that

$$A_x\mathbf{g}'(\mathbf{b})\mathbf{k} + A_y\mathbf{k} = A(\mathbf{g}'(\mathbf{b})\mathbf{k}, \mathbf{k}) = A\Phi'(\mathbf{b})\mathbf{k} = \mathbf{0}$$

for every $\mathbf{k} \in R^m$. Thus

(65) $$A_x\mathbf{g}'(\mathbf{b}) + A_y = 0.$$