MAU23203—Analysis in Several Variables School of Mathematics, Trinity College Michaelmas Term 2019 Section 10: The Inverse and Implicit Function Theorems

David R. Wilkins

10.1. Contraction Mappings on Closed Subsets of Euclidean Spaces

Theorem 10.1

Let F be a closed subset of $\mathbb{R}^n$, let r be a real number satisfying $0 < r < 1$, and let $\varphi\colon F \to F$ be a continuous map from F to itself with the property that

$$|\varphi(\mathbf{x}') - \varphi(\mathbf{x}'')| \leq r|\mathbf{x}' - \mathbf{x}''|$$

for all $\mathbf{x}', \mathbf{x}'' \in F$. Then there exists a unique point $\mathbf{x}^*$ of F for which $\varphi(\mathbf{x}^*) = \mathbf{x}^*$.

Proof

Choose $\mathbf{x}_0 \in F$, and let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ be the infinite sequence of points of F defined such that $\mathbf{x}_j = \varphi(\mathbf{x}_{j-1})$ for all positive integers j. Then

$$|\mathbf{x}_{j+1} - \mathbf{x}_j| \leq r|\mathbf{x}_j - \mathbf{x}_{j-1}|$$

for all positive integers j. It follows that

$$|\mathbf{x}_{j+1} - \mathbf{x}_j| \leq r^j|\mathbf{x}_1 - \mathbf{x}_0|$$

for all positive integers j, and therefore

$$|\mathbf{x}_k - \mathbf{x}_j| \leq \frac{r^j - r^k}{1 - r}|\mathbf{x}_1 - \mathbf{x}_0| \leq \frac{r^j}{1 - r}|\mathbf{x}_1 - \mathbf{x}_0|$$

for all positive integers j and k satisfying $j < k$.

Now the inequality $r < 1$ ensures that, given any positive real number ε, there exists a positive integer N large enough to ensure that $r^j|\mathbf{x}_1 - \mathbf{x}_0| < (1 - r)\varepsilon$ for all integers j satisfying $j \geq N$. Then $|\mathbf{x}_k - \mathbf{x}_j| < \varepsilon$ for all positive integers j and k satisfying $k > j \geq N$. The infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is thus a Cauchy sequence of points of $\mathbb{R}^n$. Now all Cauchy sequences in $\mathbb{R}^n$ are convergent (see Theorem 2.8). We conclude therefore that the infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is convergent. Let $\mathbf{x}^* = \lim_{j\to+\infty} \mathbf{x}_j$. Then $\mathbf{x}^* \in F$, because F is closed in $\mathbb{R}^n$. Moreover

$$\mathbf{x}^* = \lim_{j\to+\infty} \mathbf{x}_{j+1} = \lim_{j\to+\infty} \varphi(\mathbf{x}_j) = \varphi\left(\lim_{j\to+\infty} \mathbf{x}_j\right) = \varphi(\mathbf{x}^*).$$

We have thus proved the existence of a point $\mathbf{x}^*$ of F for which $\varphi(\mathbf{x}^*) = \mathbf{x}^*$.

If $\tilde{\mathbf{x}}$ belongs to F, and if $\varphi(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}$ then

$$|\tilde{\mathbf{x}} - \mathbf{x}^*| = |\varphi(\tilde{\mathbf{x}}) - \varphi(\mathbf{x}^*)| \leq r|\tilde{\mathbf{x}} - \mathbf{x}^*|.$$

But $r < 1$. It follows that the Euclidean distance $|\tilde{\mathbf{x}} - \mathbf{x}^*|$ from $\tilde{\mathbf{x}}$ to $\mathbf{x}^*$ cannot be strictly positive, and therefore $\tilde{\mathbf{x}} = \mathbf{x}^*$. We conclude therefore that $\mathbf{x}^*$ is the unique point of F for which $\varphi(\mathbf{x}^*) = \mathbf{x}^*$, as required. ∎

10.2. The Inverse Function Theorem

Lemma 10.2

Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a differentiable function mapping X into $\mathbb{R}^n$, let $\mathbf{p}$ be a point of X, and let c be a positive real number. Suppose that $|\mathbf{x} - \mathbf{p}| \leq c|\varphi(\mathbf{x}) - \varphi(\mathbf{p})|$ for all points $\mathbf{x}$ of X. Then $|\mathbf{v}| \leq c|(D\varphi)_{\mathbf{p}}\mathbf{v}|$ for all $\mathbf{v} \in \mathbb{R}^m$.

Proof
Let $\mathbf{v} \in \mathbb{R}^m$. Then

$$t|\mathbf{v}| = |(\mathbf{p} + t\mathbf{v}) - \mathbf{p}| \leq c|\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p})|$$

for all positive real numbers t small enough to ensure that $\mathbf{p} + t\mathbf{v} \in X$. Now

$$(D\varphi)_{\mathbf{p}}\mathbf{v} = \lim_{t \to 0^+} \frac{\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p})}{t}$$

(see Proposition 9.13). It follows that

$$\begin{aligned} |\mathbf{v}| &\leq \lim_{t \to 0^+} c \left| \frac{\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p})}{t} \right| = c \left| \lim_{t \to 0^+} \frac{\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p})}{t} \right| \\ &= c|(D\varphi)_{\mathbf{p}}\mathbf{v}|, \end{aligned}$$

as required. ∎

Proposition 10.3

Let X be an open set in $\mathbb{R}^n$, let $\varphi\colon X \to \mathbb{R}^n$ be a differentiable function on X, and let $\mathbf{p}$ be a point of X at which the derivative of φ is both invertible and continuous. Then there exist positive real numbers r, s and c such that the following properties hold:

(i) *if $\mathbf{x} \in \mathbb{R}^n$ satisfies $|\mathbf{x} - \mathbf{p}| \leq r$ then $x \in X$;*

(ii) *if $\mathbf{y} \in \mathbb{R}^n$ satisfies $|\mathbf{y} - \varphi(\mathbf{p})| < s$ then there exists $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < r$ for which $\varphi(\mathbf{x}) = \mathbf{y}$;*

(iii) *$|\mathbf{x}' - \mathbf{x}''| \leq c|\varphi(\mathbf{x}') - \varphi(\mathbf{x}'')|$ for all points $\mathbf{x}'$ and $\mathbf{x}''$ of X for which $|\mathbf{x}' - \mathbf{p}| \leq r$ and $|\mathbf{x}'' - \mathbf{p}| \leq r$.*

Proof

The derivative $(D\varphi)_{\mathbf{p}} \colon \mathbb{R}^n \to \mathbb{R}^n$ of φ at the point $\mathbf{p}$ is an invertible linear transformation, by assumption. Let $T = (D\varphi)_{\mathbf{p}}^{-1}$, let a positive real number c be chosen such that $2|T\mathbf{x}| \leq c$ for all $\mathbf{x} \in \mathbb{R}^n$ satisfying $|\mathbf{x}| = 1$, and let $\psi \colon X \to \mathbb{R}^n$ be defined such that

$$\psi(\mathbf{x}) = \mathbf{x} - T(\varphi(\mathbf{x}) - \mathbf{q})$$

for all $\mathbf{x} \in X$, where $\mathbf{q} = \varphi(\mathbf{p})$.

Now the derivative of any linear transformation at any point is equal to that linear transformation (see Lemma 9.9). It follows from the Chain Rule (Proposition 9.20) that the derivative of the composition function $T \circ \varphi$ at any point $\mathbf{x}$ of X is equal to $T(D\varphi)_{\mathbf{x}}$. It follows that $(D\psi)_{\mathbf{x}} = I - T(D\varphi)_{\mathbf{x}}$ for all $\mathbf{x} \in X$, where I denotes the identity operator on $\mathbb{R}^n$. In particular $(D\psi)_{\mathbf{p}} = I - T(D\varphi)_{\mathbf{p}} = 0$. Moreover $\psi(\mathbf{p}) = \mathbf{p}$. It then follows from a straightforward application of Corollary 9.7 that there exists a positive real number r small enough to ensure both that $\mathbf{x} \in X$ for all elements $\mathbf{x}$ of $\mathbb{R}^n$ satisfying $|\mathbf{x} - \mathbf{p}| \leq r$ and also that

$$|\psi(\mathbf{x}') - \psi(\mathbf{x}'')| \leq \tfrac{1}{2}|\mathbf{x}' - \mathbf{x}''|$$

for all points $\mathbf{x}'$ and $\mathbf{x}''$ of X for which $|\mathbf{x}' - \mathbf{p}| \leq r$ and $|\mathbf{x}'' - \mathbf{p}| \leq r$.

Let $\mathbf{x}'$ and $\mathbf{x}''$ be points of X for which $|\mathbf{x}' - \mathbf{p}| \leq r$ and $|\mathbf{x}'' - \mathbf{p}| \leq r$. Then

$$\psi(\mathbf{x}') - \psi(\mathbf{x}'') = \mathbf{x}' - \mathbf{x}'' - T(\varphi(\mathbf{x}') - \varphi(\mathbf{x}'')),$$

because T is a linear transformation, and therefore

$$\begin{aligned} |\mathbf{x}' - \mathbf{x}''| &= \left|\psi(\mathbf{x}') - \psi(\mathbf{x}'') + T(\varphi(\mathbf{x}') - \varphi(\mathbf{x}''))\right| \\ &\leq |\psi(\mathbf{x}') - \psi(\mathbf{x}'')| + \left|T(\varphi(\mathbf{x}') - \varphi(\mathbf{x}''))\right| \\ &\leq \tfrac{1}{2}|\mathbf{x}' - \mathbf{x}''| + \left|T(\varphi(\mathbf{x}') - \varphi(\mathbf{x}''))\right|. \end{aligned}$$

Subtracting $\frac{1}{2}|\mathbf{x}' - \mathbf{x}''|$ from both sides of this inequality, and multiplying by 2, we deduce that

$$|\mathbf{x}' - \mathbf{x}''| \leq 2\left|T(\varphi(\mathbf{x}') - \varphi(\mathbf{x}''))\right| \leq c|\varphi(\mathbf{x}') - \varphi(\mathbf{x}'')|,$$

for all points $\mathbf{x}'$ and $\mathbf{x}''$ of X satisfying $|\mathbf{x}' - \mathbf{p}| \leq r$ and $|\mathbf{x}'' - \mathbf{p}| \leq r$.

Now let

$$F = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{p}| \leq r\}.$$

Then F is a closed subset of $\mathbb{R}^n$, and $F \subset X$. Moreover $|\psi(\mathbf{x}') - \psi(\mathbf{x}'')| \leq \frac{1}{2}|\mathbf{x}' - \mathbf{x}''|$ for all $\mathbf{x}' \in F$ and $\mathbf{x}'' \in F$.

Let $\mathbf{y} \in \mathbb{R}^n$ satisfy $|\mathbf{y} - \mathbf{q}| < s$, where $s = r/c$, let $\mathbf{z} = \mathbf{p} + T(\mathbf{y} - \mathbf{q})$, and let

$$\theta(\mathbf{x}) = \psi(\mathbf{x}) + \mathbf{z} - \mathbf{p}$$

for all $\mathbf{x} \in X$. Now the choice of c then ensures that

$$|\mathbf{z} - \mathbf{p}| \leq \tfrac{1}{2} c |\mathbf{y} - \mathbf{q}| \leq \tfrac{1}{2} cs = \tfrac{1}{2} r.$$

If $\mathbf{x} \in X$ satisfies $|\mathbf{x} - \mathbf{p}| \leq r$, and if

$$\mathbf{x}' = \psi(\mathbf{x}) + \mathbf{z} - \mathbf{p},$$

then

$$|\mathbf{x}' - \mathbf{z}| = |\psi(\mathbf{x}) - \mathbf{p}| = |\psi(\mathbf{x}) - \psi(\mathbf{p})| \leq \tfrac{1}{2} |\mathbf{x} - \mathbf{p}| \leq \tfrac{1}{2} r,$$

and therefore

$$|\mathbf{x}' - \mathbf{p}| \leq |\mathbf{x}' - \mathbf{z}| + |\mathbf{z} - \mathbf{p}| < r.$$

We conclude therefore that θ maps the closed set F into itself, where

$$F = \{\mathbf{x} \in X : |\mathbf{x} - \mathbf{p}| \leq r\}.$$

Moreover $|\theta(\mathbf{x})| < r$ for all $\mathbf{x} \in F$ and

$$|\theta(\mathbf{x}') - \theta(\mathbf{x}'')| = |\psi(\mathbf{x}') - \psi(\mathbf{x}'')| \leq \tfrac{1}{2}|\mathbf{x}' - \mathbf{x}''|$$

for all $\mathbf{x}' \in F$ and $\mathbf{x}'' \in F$. It then follows from Theorem 10.1 that there exists a point $\mathbf{x}$ of F for which $\theta(\mathbf{x}) = \mathbf{x}$. Then $|\mathbf{x} - \mathbf{p}| < r$. Also

$$\mathbf{x} = \theta(\mathbf{x}) = \psi(\mathbf{x}) + \mathbf{z} - \mathbf{p} = \mathbf{x} - T(\varphi(\mathbf{x}) - \mathbf{q}) + \mathbf{z} - \mathbf{p},$$

where $\mathbf{q} = \varphi(\mathbf{p})$, and thus $\mathbf{z} - \mathbf{p} = T(\varphi(\mathbf{x}) - \mathbf{q})$. But $\mathbf{z} - \mathbf{p} = T(\mathbf{y} - \mathbf{q})$. It follows that $T\mathbf{y} = T(\varphi(\mathbf{x}))$, and therefore

$$\mathbf{y} = (D\varphi)_{\mathbf{p}}(T\mathbf{y}) = (D\varphi)_{\mathbf{p}}(T(\varphi(\mathbf{x})) = \varphi(\mathbf{x}).$$

We have thus shown that, given any element $\mathbf{y}$ of $\mathbb{R}^n$ satisfying $|\mathbf{y} - \mathbf{q}| < s$, there exists $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < r$ for which $\varphi(\mathbf{x}) = \mathbf{y}$. This completes the proof. $\blacksquare$

Theorem 10.4 (Inverse Function Theorem)

Let $\varphi\colon X \to \mathbb{R}^n$ be a continuously differentiable function defined over an open set X in n-dimensional Euclidean space $\mathbb{R}^n$ and mapping X into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X. Suppose that the derivative $(D\varphi)_{\mathbf{p}}\colon \mathbb{R}^n \to \mathbb{R}^n$ of the map φ at the point $\mathbf{p}$ is an invertible linear transformation. Then there exists an open set W in $\mathbb{R}^n$ and a continuously differentiable function $\mu\colon W \to X$ that satisfies the following conditions:—

(i) *$\mu(W)$ is an open set in $\mathbb{R}^n$ contained in X, and $\mathbf{p} \in \mu(W)$;*

(ii) *$\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$.*

Proof

It follows from Proposition 10.3 that there exist positive real numbers r, s and c such that the following properties hold: if $\mathbf{x} \in \mathbb{R}^n$ satisfies $|\mathbf{x} - \mathbf{p}| \leq r$ then $x \in X$; if $\mathbf{y} \in \mathbb{R}^n$ satisfies $|\mathbf{y} - \varphi(\mathbf{p})| < s$ then there exists $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < r$ for which $\varphi(\mathbf{x}) = \mathbf{y}$; $|\mathbf{x}' - \mathbf{x}''| \leq c|\varphi(\mathbf{x}') - \varphi(\mathbf{x}'')|$ for all points $\mathbf{x}'$ and $\mathbf{x}''$ of X for which $|\mathbf{x}' - \mathbf{p}| \leq r$ and $|\mathbf{x}'' - \mathbf{p}| \leq r$. It then follows from Lemma 10.2 that $|(D\varphi)_{\mathbf{u}}\mathbf{v}| \geq c|\mathbf{v}|$ for all $\mathbf{u} \in X$ satisfying $|\mathbf{u} - \mathbf{p}| < r$ and for all $\mathbf{v} \in \mathbb{R}^n$.

Let

$$W = \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{y} - \varphi(\mathbf{p})| < s\}.$$

If $\mathbf{y}$ is a point of W, there exists a point $\mathbf{x}$ of X such that $|\mathbf{x} - \mathbf{p}| < r$ and $\varphi(\mathbf{x}) = \mathbf{y}$. There cannot exist more than one point of X with this property because if $\mathbf{x}'$ is a point of X distinct from $\mathbf{x}$, and if $|\mathbf{x}' - \mathbf{p}| < r$, then

$$|\varphi(\mathbf{x}') - \mathbf{y}| \geq c|\mathbf{x}' - \mathbf{x}| > 0.$$

Therefore there is a well-defined function $\mu \colon W \to \mathbb{R}^n$ characterized by the property that, for each $\mathbf{y} \in W$, $\mu(\mathbf{y})$ is the unique point of X for which $|\mu(\mathbf{y}) - \mathbf{p}| < r$ and $\varphi(\mu(\mathbf{y})) = \mathbf{y}$.

We next show that $\mu(W)$ is an open subset of $\mathbb{R}^n$. Let $\mathbf{u} \in \mu(W)$. Then $|\mathbf{u} - \mathbf{p}| < r$, and there exists $\mathbf{w} \in W$ for which $\mu(\mathbf{w}) = \mathbf{u}$. But then $\varphi(\mathbf{u}) = \mathbf{w}$, and thus $\mathbf{u} \in \varphi^{-1}(W)$. We conclude that

$$\mu(W) \subset \varphi^{-1}(W) \cap \{\mathbf{x} \in X : |\mathbf{x} - \mathbf{p}| < r\}.$$

Conversely let $\mathbf{u}$ be a point of $\varphi^{-1}(W)$ satisfying $|\mathbf{u} - \mathbf{p}| < r$, and let $\mathbf{w} = \varphi(\mathbf{u})$. Then $\mathbf{w} \in W$ and $\mu(\mathbf{w}) = \mathbf{u}$, and therefore $\mathbf{u} \in \mu(W)$. We conclude from this that

$$\mu(W) = \varphi^{-1}(W) \cap \{\mathbf{x} \in X : |\mathbf{x} - \mathbf{p}| < r\}.$$

It follows that $\mu(W)$ is the intersection of two open subsets of X, and must therefore itself be open in X. Now X itself is open in $\mathbb{R}^n$. It follows that $\mu(W)$ is indeed an open subset of $\mathbb{R}^n$.

Let $\mathbf{w} \in W$, and let $\mathbf{u} = \mu(\mathbf{w})$. Then $|\mathbf{u} - \mathbf{p}| < r$. Let some positive real number ε be given. The differentiability of the map φ at $\mathbf{u}$ ensures the existence of a positive real number δ such that $\eta + |\mathbf{u} - \mathbf{p}| \leq r$ and

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{u}) - (D\varphi)_{\mathbf{u}}(\mathbf{x} - \mathbf{u})| \leq \frac{\varepsilon}{c^2}|\mathbf{x} - \mathbf{u}|$$

for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{u}| \leq c\delta$. Let $\mathbf{y} \in W$ satisfy $|\mathbf{y} - \mathbf{w}| < \delta$, and let $\mathbf{x} = \mu(\mathbf{y})$. Then $\varphi(\mathbf{x}) = \mathbf{y}$ and $\varphi(\mathbf{u}) = \mathbf{w}$, and therefore

$$|\mathbf{x} - \mathbf{u}| \leq c|\varphi(\mathbf{x}) - \varphi(\mathbf{u})| = c|\mathbf{y} - \mathbf{w}| < c\delta.$$

It follows that

$$|\mathbf{y} - \mathbf{w} - (D\varphi)_{\mathbf{u}}(\mathbf{x} - \mathbf{u})| \leq \frac{\varepsilon}{c^2}|\mathbf{x} - \mathbf{u}| \leq \frac{\varepsilon}{c}|\mathbf{y} - \mathbf{w}|,$$

and therefore

$$\begin{aligned} \left|(D\varphi)_{\mathbf{u}}^{-1}(\mathbf{y} - \mathbf{w}) - (\mathbf{x} - \mathbf{u})\right| &\leq c\,|\mathbf{y} - \mathbf{w} - (D\varphi)_{\mathbf{u}}(\mathbf{x} - \mathbf{u})| \\ &\leq \varepsilon|\mathbf{y} - \mathbf{w}|. \end{aligned}$$

But $\mathbf{x} - \mathbf{u} = \mu(\mathbf{y}) - \mu(\mathbf{w})$. We conclude therefore that, given any positive real number ε, there exists some positive real number δ such that

$$\left|\mu(\mathbf{y}) - \mu(\mathbf{w}) - (D\varphi)_{\mathbf{u}}^{-1}(\mathbf{y} - \mathbf{w})\right| \leq \varepsilon|\mathbf{y} - \mathbf{w}|$$

for all points $\mathbf{y}$ of W satisfying $|\mathbf{y} - \mathbf{w}| < \delta$. It follows that the map $\mu\colon W \to X$ is differentiable at $\mathbf{w}$, and moreover

$$(D\mu)_{\mathbf{w}} = (D\varphi)_{\mathbf{u}}^{-1} = (D\varphi)_{\mu(\mathbf{y})}^{-1}.$$

Now the map $\mu\colon W \to X$ is continuous, because it is differentiable. Also the coefficients of the Jacobian matrix representing the derivative of φ at points $\mathbf{x}$ of $\mu(W)$ are continuous functions of $\mathbf{x}$ on $\mu(W)$. It follows that the coefficients of the inverse of the Jacobian matrix of the map φ are also continuous functions of $\mathbf{x}$ on $\mu(W)$. Each coefficient of the Jacobian matrix of the map μ is thus the composition of the continuous map μ with a continuous real-valued function on $\mu(W)$, and must therefore itself be a continuous real-valued function on W. It follows that the map $\mu\colon W \to X$ is continuously differentiable on W. This completes the proof. ∎

10.3. The Implicit Function Theorem

Theorem 10.5

Let X be an open set in $\mathbb{R}^n$, let $f_1, f_2, \ldots, f_m$ be a continuously differentiable real-valued functions on X, where $m < n$, let

$$M = \{\mathbf{x} \in X : f_i(\mathbf{x}) = 0 \text{ for } i = 1, 2, \ldots, m\},$$

and let $\mathbf{p}$ be a point of M.

Suppose that $f_1, f_2, \ldots, f_m$ are zero at $\mathbf{p}$ and that the matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_m} \end{pmatrix}$$

is invertible at the point $\mathbf{p}$. Then there exists an open neighbourhood U of $\mathbf{p}$ and continuously differentiable functions $h_1, h_2, \ldots, h_m$ of $n - m$ real variables, defined around $(p_{m+1}, \ldots, p_n)$ in $\mathbb{R}^{n-m}$, such that

$$\begin{aligned} M \cap U \quad = \quad & \{(x_1, x_2, \ldots, x_n) \in U : \\ & \qquad x_i = h_i(x_{m+1}, \ldots, x_n) \text{ for } i = 1, 2, \ldots, m\}. \end{aligned}$$

Proof

Let $\varphi\colon X \to \mathbb{R}^n$ be the continuously differentiable function defined such that

$$\varphi(\mathbf{x}) = \Bigl(f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}), x_{m+1}, \ldots, x_n\Bigr)$$

for all $\mathbf{x} \in X$. (Thus the ith Cartesian component of the function φ is equal to f_i for $i \leq m$, but is equal to x_i for $m < i \leq n$.) Let J be the Jacobian matrix of φ at the point $\mathbf{p}$, and let $J_{i,j}$ denote the coefficient in the ith row and jth column of J. Then

$$J_{i,j} = \frac{\partial f_i}{\partial x_j}$$

for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Also $J_{i,i} = 1$ if $i > m$, and $J_{i,j} = 0$ if $i > m$ and $j \neq i$.

The matrix J can therefore be represented in block form as

$$J = \left(\begin{array}{c|c} J_0 & A \\ \hline 0 & I_{n-m} \end{array} \right),$$

where J_0 is the leading $m \times m$ minor of the matrix J, A is an $m \times (n-m)$ minor of the matrix J and I_{n-m} is the identity $(n-m) \times (n-m)$ matrix. It follows from standard properties of determinants that $\det J = \det J_0$. Moreover the hypotheses of the theorem require that $\det J_0 \neq 0$. Therefore $\det J \neq 0$. The derivative $(D\varphi)_{\mathbf{p}}$ of φ at the point $\mathbf{p}$ is represented by the Jacobian matrix J. It follows that $(D\varphi)_{\mathbf{p}} \colon \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear transformation.

The Inverse Function Theorem (Theorem 10.4) now ensures the existence of a continuously differentiable map $\mu\colon W \to X$ with the properties that $\mu(W)$ is an open subset of X and $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$.

Let $\mathbf{y}$ be a point of W, and let $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. Then $\mathbf{y} = \varphi(\mu(\mathbf{y}))$, and therefore $y_i = f_i(\mu(\mathbf{y}))$ for $i = 1, 2, \ldots, m$, and y_i is equal to the ith component of $\mu(\mathbf{y})$ when $m < i \leq n$.

Now $\mathbf{p} \in \mu(W)$. Therefore there exists some point $\mathbf{q}$ of W satisfying $\mu(\mathbf{q}) = \mathbf{p}$. Now $\mathbf{p} \in M$, and therefore $f_i(\mathbf{p}) = 0$ for $i = 1, 2, \ldots, m$. But $q_i = f_i(\mu(\mathbf{q})) = f_i(\mathbf{p})$ when $1 \leq i \leq m$. It follows that $q_i = 0$ when $1 \leq i \leq m$. Also $q_i = p_i$ when $i > m$.

Let g_i denote the ith Cartesian component of the continuously differentiable map $\mu\colon W \to \mathbb{R}^n$ for $i = 1, 2, \ldots, n$. Then $g_i\colon W \to \mathbb{R}$ is a continuously differentiable real-valued function on W for $i = 1, 2, \ldots, n$. If $(y_1, y_2, \ldots, y_n) \in W$ then

$$(y_1, y_2, \ldots, y_n) = \varphi(\mu(y_1, y_2, \ldots, y_n)).$$

It then follows from the definition of the map φ that y_i is the ith Cartesian component of $\mu(y_1, y_2, \ldots, y_n)$ when $i > m$, and thus

$$y_i = g_i(y_1, y_2, \ldots, y_n) \quad \text{when} \quad i > m.$$

Now $\mu(W)$ is an open set, and $\mathbf{p} \in \mu(W)$. It follows that there exists some positive real number δ such that $H(\mathbf{p}, \delta) \subset \mu(W)$. where

$$\begin{aligned} H(\mathbf{p}, \delta) &= \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : \\ &\qquad p_i - \delta < x_i < p_i + \delta \text{ for } i = 1, 2, \ldots, n\}. \end{aligned}$$

Let

$$\begin{aligned} D &= \{(z_1, z_2, \ldots, z_{n-m}) \in \mathbb{R}^{n-m} : p_{m+j} - \delta < z_j < p_{m+j} + \delta \\ &\qquad \text{for } j = 1, 2, \ldots, n - m\}, \end{aligned}$$

and let $h_i \colon D \to \mathbb{R}$ be defined so that

$$h_i(z_1, z_2, \ldots, z_{n-m}) = g_i(0, 0, \ldots, 0, z_1, z_2, \ldots, z_{n-m})$$

for $i = 1, 2, \ldots, m$.

Let $\mathbf{x} \in H(\mathbf{p}, \delta)$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Then $\mathbf{x} \in \mu(W)$. There therefore exists $\mathbf{w} \in W$ for which $\mu(\mathbf{w}) = \mathbf{x}$. But the properties of the map μ ensure that $\mathbf{w} = \varphi(\mu(\mathbf{w}))$. It follows that

$$\mathbf{x} = \mu(\mathbf{w}) = \mu(\varphi(\mu(\mathbf{w}))) = \mu(\varphi(\mathbf{x})).$$

Thus

$$\begin{aligned}(x_1, x_2, \ldots, x_n) &= \mu(\varphi(\mathbf{x})) \\ &= \mu\Big(f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}), x_{m+1}, \ldots, x_n\Big).\end{aligned}$$

On equating Cartesian components we find that

$$x_i = g_i\Big(f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}), x_{m+1}, \ldots, x_n\Big).$$

for $i = 1, 2, \ldots, n$.

In particular, if $\mathbf{x} \in H(\mathbf{p}, \delta) \cap M$ then

$$f_1(\mathbf{x}) = f_2(\mathbf{x}) = \cdots = f_m(\mathbf{x}) = 0,$$

and therefore

$$\begin{aligned} x_i &= g_i\Big(0, 0, \ldots, 0, x_{m+1}, \ldots, x_n\Big) \\ &= h_i\Big(x_{m+1}, \ldots, x_n\Big). \end{aligned}$$

for $i = 1, 2, \ldots, m$. It follows that

$$\begin{aligned} M \cap H(\mathbf{p}, \delta) \quad \subset \quad & \{(x_1, x_2, \ldots, x_n) \in H(\mathbf{p}, \delta) : \\ & \qquad x_i = h_i(x_{m+1}, \ldots, x_n) \text{ for } i = 1, 2, \ldots, m\}. \end{aligned}$$

Now let $\mathbf{x}$ be a point of $H(\mathbf{x}, \delta)$ whose Cartesian components $x_1, x_2, \ldots, x_n$ satisfy the equations

$$x_i = h_i(x_{m+1}, \ldots, x_n)$$

for $i = 1, 2, \ldots, m$. Then

$$x_i = g_i(0, 0, \ldots, 0, x_{m+1}, \ldots, x_n)$$

for $i = 1, 2, \ldots, m$. Now it was shown earlier that

$$y_i = g_i(y_1, y_2, \ldots, y_n)$$

for all $(y_1, y_2, \ldots, y_n) \in W$ when $i > m$. It follows from this that

$$x_i = g_i(0, 0, \ldots, 0, x_{m+1}, \ldots, x_n)$$

when $m < i \leq n$. The functions $g_1, g_2, \ldots, g_n$ are the Cartesian components of the map $\mu \colon W \to X$. We conclude therefore that

$$(x_1, x_2, \ldots, x_n) = \mu(0, 0, \ldots, 0, x_{m+1}, \ldots, x_n),$$

Applying the function φ to both sides of this equation we see that

$$\begin{aligned}\varphi(x_1, x_2, \ldots, x_n) &= \varphi(\mu(0, 0, \ldots, 0, x_{m+1}, \ldots, x_n)) \\ &= (0, 0, \ldots, 0, x_{m+1}, \ldots, x_n).\end{aligned}$$

It then follows from the definition of the map φ that

$$f_i(x_1, x_2, \ldots, x_n) = 0,$$

for $i = 1, 2, \ldots, m$. We have thus shown that if $\mathbf{x}$ is a point of $H(\mathbf{x}, \delta)$ whose Cartesian components $x_1, x_2, \ldots, x_n$ satisfy the equations

$$x_i = h_i(x_{m+1}, \ldots, x_n)$$

for $i = 1, 2, \ldots, m$ then $\mathbf{x} \in M$. The converse of this result was proved earlier. The proof of the theorem is therefore completed on taking $U = H(\mathbf{p}, \delta)$. ∎