Module MAU23203: Analysis in Several Real Variables
Michaelmas Term 2021
Section 9: The Inverse and Implicit Function Theorems

D. R. Wilkins

Contents

9 The Inverse and Implicit Function Theorems

9.1 Contraction Mappings on Closed Subsets of Euclidean Spaces

Definition Let F be a subset of $\mathbb{R}^n$ for some positive integer n. A function $\varphi\colon F \to F$ mapping that set F into itself is said to be a *contraction mapping* on F if there exists some non-negative real number λ satisfying $\lambda < 1$ that is such as to ensure that

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v})| \leq \lambda |\mathbf{u} - \mathbf{v}|$$

for all points $\mathbf{u}$ and $\mathbf{v}$ of F.

Theorem 9.1 *Let F be a closed subset of $\mathbb{R}^n$, and let $\varphi\colon F \to F$ be a contraction mapping on the set F. Then there exists a unique point $\mathbf{p}$ of F for which $\varphi(\mathbf{p}) = \mathbf{p}$.*

Proof The function $\varphi\colon F \to F$ is a contraction mapping. Therefore a non-negative real number λ satisfying $\lambda < 1$ can be associated with the function φ so as to ensure that

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v})| \leq \lambda |\mathbf{u} - \mathbf{v}|$$

for all points $\mathbf{u}$ and $\mathbf{v}$ of F.

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

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

for all positive integers j. It follows that

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

for all positive integers j, and therefore

$$\begin{aligned} |\mathbf{x}_k - \mathbf{x}_j| &\leq \left(\sum_{m=j}^{k-1} \lambda^m \right) |\mathbf{x}_1 - \mathbf{x}_0| \leq \frac{\lambda^j - \lambda^k}{1 - \lambda} |\mathbf{x}_1 - \mathbf{x}_0| \\ &\leq \frac{\lambda^j}{1 - \lambda} |\mathbf{x}_1 - \mathbf{x}_0| \end{aligned}$$

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

Now the inequality $\lambda < 1$ ensures that, given any positive real number ε, there exists a positive integer N large enough to ensure that $\lambda^j |\mathbf{x}_1 - \mathbf{x}_0| < (1-\lambda)\varepsilon$ 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 F. Now $F \subset \mathbb{R}^n$ and every Cauchy sequence in $\mathbb{R}^n$ is convergent (see Theorem 3.7). We conclude therefore that the infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is convergent. Let $\mathbf{p} = \lim_{j \to +\infty} \mathbf{x}_j$. Then $\mathbf{p} \in F$, because F is closed in $\mathbb{R}^n$ (see Lemma 4.7). Moreover

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

(This follows on applying Proposition 5.2.) We have thus proved the existence of a point $\mathbf{p}$ of F for which $\varphi(\mathbf{p}) = \mathbf{p}$.

Now let $\mathbf{q}$ be any point of the closed set F with the property that $\varphi(\mathbf{q}) = \mathbf{q}$. Then

$$|\mathbf{q} - \mathbf{p}| = |\varphi(\mathbf{q}) - \varphi(\mathbf{p})| \leq \lambda |\mathbf{q} - \mathbf{p}|.$$

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

9.2 The Inverse Function Theorem

Lemma 9.2 *Let X be an open set in $\mathbb{R}^m$, let $\varphi: X \to \mathbb{R}^n$ be a differentiable function mapping X into $\mathbb{R}^n$, let $\mathbf{p}$ be a point of X, and let K be a positive real number. Suppose that $|\mathbf{x} - \mathbf{p}| \leq K|\varphi(\mathbf{x}) - \varphi(\mathbf{p})|$ for all points $\mathbf{x}$ of X. Then $|\mathbf{w}| \leq K|(D\varphi)_{\mathbf{p}}\mathbf{w}|$ for all $\mathbf{w} \in \mathbb{R}^m$.*

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

$$t|\mathbf{w}| = |(\mathbf{p} + t\mathbf{w}) - \mathbf{p}| \leq K|\varphi(\mathbf{p} + t\mathbf{w}) - \varphi(\mathbf{p})|$$

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

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

(see Proposition 8.13). It follows that

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

as required. ∎

Proposition 9.3 *Let X and Y be open sets in $\mathbb{R}^n$, let $\varphi: X \to \mathbb{R}^n$ be a differentiable function mapping X into $\mathbb{R}^n$, and let K be a positive real number. Suppose that $Y \subset \varphi(X)$. Suppose also that $|\mathbf{u} - \mathbf{v}| \leq K|\varphi(\mathbf{u}) - \varphi(\mathbf{v})|$ for all points $\mathbf{u}$ and $\mathbf{v}$ of X. Then there is a differentiable function $\mu: Y \to \mathbb{R}^n$ characterized by the property that, for any point $\mathbf{y}$ of Y, $\mu(\mathbf{y})$ is the unique point of X for which $\varphi(\mu(\mathbf{y})) = \mathbf{y}$. Moreover $\mu(Y)$ is an open set in $\mathbb{R}^n$, and $(D\mu)_{\varphi(\mathbf{p})} = (D\varphi)_{\mathbf{p}}^{-1}$ for all $\mathbf{p} \in \mu(Y)$.*

Proof Given any point $\mathbf{y}$ of Y, there exists at least one point $\mathbf{x}$ of X for which $\varphi(\mathbf{x}) = \mathbf{y}$, because $Y \subset \varphi(X)$. Also the stated inequality in the statement of the lemma ensures that, given any point $\mathbf{y}$ of Y, there cannot exist more than one point $\mathbf{x}$ of X for which $\varphi(\mathbf{x}) = \mathbf{y}$. Consequently there is a well-defined function $\mu: Y \to \mathbb{R}^n$ characterized by the property that, for all points $\mathbf{y}$ of the open set Y, the point $\mu(\mathbf{y})$ is the unique point of the open set X for which $\varphi(\mathbf{x}) = \mathbf{y}$. We must prove that this function μ is differentiable and that it maps the open set Y onto an open set in $\mathbb{R}^n$.

First we show that $\mu(Y)$ is an open set in $\mathbb{R}^n$. Let $\mathbf{p}$ be a point of $\mu(Y)$. The continuity of the function φ ensures that $\varphi^{-1}(Y)$ is open in X. Therefore there exists some positive real number δ that is small enough to ensure both that all points $\mathbf{x}$ of $\mathbb{R}^n$ that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$ belong to the open set X and also that all points $\mathbf{x}$ of that open set that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$ are mapped by φ into the open set Y. Consequently all points of the open ball of radius δ in $\mathbb{R}^n$ centred on the point $\mathbf{p}$ are mapped by φ into the set Y and therefore belong to $\mu(Y)$. Consequently $\mu(Y)$ is an open set in $\mathbb{R}^n$.

Let $\mathbf{q} \in Y$, and let $\mathbf{p} = \mu(\mathbf{q})$. Also let some positive real number ε be given. The differentiability of the function φ at $\mathbf{p}$ ensures the existence of a positive real number δ that is small enough to ensure that all points $\mathbf{x}$ of $\mathbb{R}^n$ that satisfy the inequality $|\mathbf{x} - \mathbf{p}| \leq K\delta$ belong to the open set X and also satisfy the inequality

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| \leq \frac{\varepsilon}{K^2}|\mathbf{x} - \mathbf{p}|.$$

Reducing the value of δ if necessary, we can also ensure that the open ball of radius δ centred on the point $\mathbf{q}$ is contained in the open set Y. Let $\mathbf{y} \in Y$ satisfy $|\mathbf{y} - \mathbf{q}| < \delta$, and let $\mathbf{x} = \mu(\mathbf{y})$. Then $\varphi(\mathbf{x}) = \mathbf{y}$ and $\varphi(\mathbf{p}) = \mathbf{q}$, and therefore

$$|\mathbf{x} - \mathbf{p}| \leq K|\varphi(\mathbf{x}) - \varphi(\mathbf{p})| = K|\mathbf{y} - \mathbf{q}| < K\delta.$$

It follows that

$$\begin{aligned} |\mathbf{y} - \mathbf{q} - (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| &= |\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| \\ &\leq \frac{\varepsilon}{K^2}|\mathbf{x} - \mathbf{p}| \leq \frac{\varepsilon}{K}|\mathbf{y} - \mathbf{q}|. \end{aligned}$$

Consequently it follows (on applying Lemma 9.2) that

$$\begin{aligned} \left|(D\varphi)_{\mathbf{p}}^{-1}(\mathbf{y}-\mathbf{q})-(\mathbf{x}-\mathbf{p})\right| &\leq K\left|(D\varphi)_{\mathbf{p}}((D\varphi)_{\mathbf{p}}^{-1}(\mathbf{y}-\mathbf{q})-(\mathbf{x}-\mathbf{p}))\right| \\ &\leq K\left|\mathbf{y}-\mathbf{q}-(D\varphi)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})\right| \\ &\leq \varepsilon|\mathbf{y}-\mathbf{q}|. \end{aligned}$$

But $\mathbf{x}=\mu(\mathbf{y})$ and $\mathbf{p}=\mu(\mathbf{q})$. We conclude therefore that, given any positive real number ε, there exists some positive real number δ such that $\mathbf{y}\in Y$ and

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

for all points $\mathbf{y}$ of $\mathbb{R}^n$ satisfying $|\mathbf{y}-\mathbf{q}|<\delta$. It follows that the function $\mu\colon Y\to\mathbb{R}^n$ is differentiable at $\mathbf{q}$, and moreover

$$(D\mu)_{\mathbf{q}}=(D\varphi)_{\mathbf{p}}^{-1}=(D\varphi)_{\mu(\mathbf{q})}^{-1}.$$

The result follows. ∎

Definition A vector-valued function, defined over an open set in some Euclidean space, is said to be *continuously differentiable* if it is differentiable, with continuous first order partial derivatives throughout its domain.

It follows directly from a result previously established that if a vector-valued function defined over an open set in a Euclidean space has continuous first order partial derivatives then that function must necessarily be differentiable (see Proposition 8.12). Thus the existence of continuous first order partial derivatives throughout the domain of such a function is sufficient to ensure that the function is continuously differentiable over its domain. No additional differentiability criterion is required in order to ensure continuous differentiability.

Theorem 9.4 (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 function φ at the point $\mathbf{p}$ is an invertible linear transformation. Then there exists an open set Y in $\mathbb{R}^n$ and a continuously differentiable function $\mu\colon Y\to\mathbb{R}^n$ that satisfies the following conditions:—*

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

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

Proof The derivative $(D\varphi)_{\mathbf{p}}: \mathbb{R}^n \to \mathbb{R}^n$ of φ at the point $\mathbf{p}$ is an invertible linear operator on the real vector space $\mathbb{R}^n$. In other words, it is an invertible linear transformation mapping $\mathbb{R}^n$ onto itself. Let $T = (D\varphi)_{\mathbf{p}}^{-1}$, and let a positive real number K be chosen such that $2|T\mathbf{w}| \leq K$ for all $\mathbf{w} \in \mathbb{R}^n$ satisfying $|\mathbf{w}| = 1$. Then $|T\mathbf{w}| \leq \frac{1}{2}K|\mathbf{w}|$ for all $\mathbf{w} \in \mathbb{R}^n$.

Also let $\psi: X \to \mathbb{R}^n$ 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 8.9). It follows on applying the Chain Rule (Proposition 8.20) that the derivative of the composition function $T \circ \varphi$ at any point $\mathbf{x}$ of X is equal to $T(D\varphi)_{\mathbf{x}}$. Consequently $(D\psi)_{\mathbf{x}} = I - T(D\varphi)_{\mathbf{x}}$ 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 $\psi(\mathbf{p}) = \mathbf{p}$. Now the first order derivatives of the function φ are continuous at the point $\mathbf{p}$. Therefore, given that $(D\psi)_{\mathbf{p}} = 0$, we can choose some positive constant r that is 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

$$|\psi(\mathbf{u}) - \psi(\mathbf{v})| \leq \tfrac{1}{2}|\mathbf{u} - \mathbf{v}|$$

for all points $\mathbf{u}$ and $\mathbf{v}$ of X for which $|\mathbf{u} - \mathbf{p}| \leq r$ and $|\mathbf{v} - \mathbf{p}| \leq r$ (see Corollary 8.7).

Let $\mathbf{u}$ and $\mathbf{v}$ be points of X for which $|\mathbf{u} - \mathbf{p}| \leq r$ and $|\mathbf{v} - \mathbf{p}| \leq r$. Now $\psi(\mathbf{x}) = \mathbf{x} - T(\varphi(\mathbf{x}) - \mathbf{q})$ for all $\mathbf{x} \in X$, and moreover T is a linear operator. It follows that

$$\psi(\mathbf{u}) - \psi(\mathbf{v}) = \mathbf{u} - \mathbf{v} - T(\varphi(\mathbf{u}) - \varphi(\mathbf{v})).$$

Therefore

$$\begin{aligned} |\mathbf{u} - \mathbf{v}| &= |\psi(\mathbf{u}) - \psi(\mathbf{v}) + T(\varphi(\mathbf{u}) - \varphi(\mathbf{v}))| \\ &\leq |\psi(\mathbf{u}) - \psi(\mathbf{v})| + |T(\varphi(\mathbf{u}) - \varphi(\mathbf{v}))| \\ &\leq \tfrac{1}{2}|\mathbf{u} - \mathbf{v}| + |T(\varphi(\mathbf{u}) - \varphi(\mathbf{v}))|. \end{aligned}$$

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

$$|\mathbf{u} - \mathbf{v}| \leq 2\,|T(\varphi(\mathbf{u}) - \varphi(\mathbf{v}))| \leq K|\varphi(\mathbf{u}) - \varphi(\mathbf{v})|,$$

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

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 $|\psi(\mathbf{u}) - \psi(\mathbf{v})| \leq \frac{1}{2}|\mathbf{u} - \mathbf{v}|$ for all $\mathbf{u} \in F$ and $\mathbf{v} \in F$.

Let $\mathbf{y} \in \mathbb{R}^n$ satisfy $|\mathbf{y} - \mathbf{q}| < s$, where $\mathbf{q} = \varphi(\mathbf{p})$ and $s = r/K$. Also 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 $\mathbf{z} - \mathbf{p} = T(\mathbf{y} - \mathbf{q})$ and $\psi(\mathbf{x}) = \mathbf{x} - T(\varphi(\mathbf{x}) - \mathbf{q})$ for all $\mathbf{x} \in X$. It follows from the definition of $\theta(\mathbf{x})$ and the linearity of T that

$$\begin{aligned} \theta(\mathbf{x}) - \mathbf{x} &= \mathbf{z} - \mathbf{p} + \psi(\mathbf{x}) - \mathbf{x} \\ &= T(\mathbf{y} - \mathbf{q}) - T(\varphi(\mathbf{x}) - \mathbf{q}) \\ &= T(\mathbf{y} - \varphi(\mathbf{x})) \end{aligned}$$

for all $\mathbf{x} \in X$. Moreover the linear operator T is invertible. Consequently a point $\mathbf{x}$ of X satisfies the equation $\mathbf{x} = \theta(\mathbf{x})$ if and only if $\varphi(\mathbf{x}) = \mathbf{y}$. Accordingly if we can show that the restriction of the function θ to the closed set F maps that closed set into itself, where

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

and if we can also show that the restriction of the function θ to the closed set F is a contraction mapping on that closed set, then we can use the result (Theorem 9.1) concerning contraction mappings on closed sets previously established to deduce the existence of a fixed point $\mathbf{x}$ for θ located within the closed set F. That fixed point $\mathbf{x}$ will then satisfy the equation $\varphi(\mathbf{x}) = \mathbf{y}$.

Now the positive constant K was chosen at the beginning of the proof so as to ensure that $|T\mathbf{w}| \leq \frac{1}{2}K|\mathbf{w}|$ for all $\mathbf{w} \in \mathbb{R}^n$. Also $|\mathbf{y} - \mathbf{q}| < s$, where $s = r/K$. Consequently

$$|\mathbf{z} - \mathbf{p}| = |T(\mathbf{y} - \mathbf{q})| \leq \tfrac{1}{2}K|\mathbf{y} - \mathbf{q}| < \tfrac{1}{2}Ks = \tfrac{1}{2}r.$$

Also $\psi(\mathbf{p}) = \mathbf{p}$, and consequently

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

Moreover $|\psi(\mathbf{u}) - \psi(\mathbf{v})| \leq \frac{1}{2}|\mathbf{u} - \mathbf{v}|$ for all points $\mathbf{u}$ and $\mathbf{v}$ of X that satisfy $|\mathbf{u} - \mathbf{p}| \leq r$ and $|\mathbf{v} - \mathbf{p}| \leq r$. Consequently if $|\mathbf{x} - \mathbf{p}| \leq r$ then

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

and therefore

$$|\theta(\mathbf{x}) - \mathbf{p}| \leq |\theta(\mathbf{x}) - \mathbf{z}| + |\mathbf{z} - \mathbf{p}| < r.$$

We have thus shown that if $\mathbf{x} \in \mathbb{R}^n$ satisfies $|\mathbf{x} - \mathbf{p}| \leq r$ then $\mathbf{x} \in X$ and $|\theta(\mathbf{x}) - \mathbf{p}| < r$. We conclude therefore that θ maps the closed set F into its interior, where

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

Moreover

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

for all $\mathbf{u} \in F$ and $\mathbf{v} \in F$. It then follows from Theorem 9.1 that there exists a point $\mathbf{x}$ of F for which $\theta(\mathbf{x}) = \mathbf{x}$. It then follows from results previously established that $|\mathbf{x} - \mathbf{p}| < r$ and $\varphi(\mathbf{x}) = \mathbf{y}$.

We have now established that, given any point $\mathbf{y}$ of $\mathbb{R}^n$ satisfying $|\mathbf{y}-\mathbf{q}| < s$, where $\mathbf{q} = \varphi(\mathbf{p})$, there exists a point $\mathbf{x}$ of X satisfying $|\mathbf{x} - \mathbf{p}| < r$ for which $\varphi(\mathbf{x}) = \mathbf{y}$. Accordingly let

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

Then

$$Y \subset \varphi(\{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{p}| < r\}).$$

It therefore follows (on applying Proposition 9.3) that there is a well-defined function $\mu: Y \to \mathbb{R}^n$ characterized by the properties that $|\mu(\mathbf{y}) - \mathbf{p}| < r$ and $\mathbf{y} = \varphi(\mu(\mathbf{y}))$ for all $\mathbf{y} \in Y$. Moreover this function μ is differentiable, and $(D\mu)_{\varphi(\mathbf{x})} = (D\varphi)_{\mathbf{x}}^{-1}$ for all $\mathbf{x} \in \mu(Y)$.

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

9.3 The Implicit Function Theorem

Theorem 9.5 (Implicit Function Theorem) *Let X be an open set in $\mathbb{R}^n$, let $f_1, f_2, \ldots, f_m$ be continuously differentiable real-valued functions on X, where $m < n$, let*

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

and let $\mathbf{p}$ *be a point of* S*. Suppose that the matrix*

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

is invertible at the point $\mathbf{p}$*. Then there exists an open neighbourhood* V *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*

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

Proof Let $\varphi: X \to \mathbb{R}^n$ 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 $j \neq i$.

We show that the derivative $(D\varphi)_{\mathbf{p}}$ of the function φ at the point $\mathbf{p}$ is an invertible linear operator on $\mathbb{R}^n$. Let $v_1, v_2, \ldots, v_n$ be real numbers, and let $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. Consider the the $m \times m$ matrix whose coefficient in the ith row and jth column is the corresponding coefficient $J_{i,j}$ of the Jacobian matrix of the function φ at the point $\mathbf{p}$. The hypotheses of the Implicit Function Theorem require that this $m \times m$ matrix be invertible. Consequently there exist real numbers $w_1, w_2, \ldots, w_m$ such that, for each integer i between 1 and m,

$$\sum_{j=1}^{m} J_{i,j} w_j = v_i - \sum_{j=m+1}^{n} J_{i,j} v_j.$$

Let $w_j = v_j$ for all integers j for which $m+1 \leq j \leq n$. Then

$$v_i = \sum_{j=1}^{m} J_{i,j} w_j + \sum_{j=m+1}^{n} J_{i,j} w_j = \sum_{j=1}^{n} J_{i,j} w_j$$

for each integer i between 1 and m. Moreover $v_i = \sum_{j=1}^{n} J_{i,j} w_j$ for each integer i between $m+1$ and n because, as already noted, $J_{i,i} = 1$ if $i > m$, and $J_{i,j} = 0$ if $i > m$ and $j \neq i$. It follows that $(D\varphi)_{\mathbf{p}} \mathbf{w} = \mathbf{v}$, where $\mathbf{w} = (w_1, w_2, \ldots, w_n)$. Now if $\mathbf{u}$ is any vector in $\mathbb{R}^n$ satisfying the equation $(D\varphi)_{\mathbf{p}} \mathbf{u} = \mathbf{v}$, and if $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, then $u_i = v_i = w_i$ for all integers i greater than m, and consequently

$$\sum_{j=1}^{m} J_{i,j} u_j = v_i - \sum_{j=m+1}^{n} J_{i,j} v_j = \sum_{j=1}^{m} J_{i,j} w_j.$$

It then follows from the invertibility of the $m \times m$ matrix with coefficient $J_{i,j}$ in the ith row and jth column that $u_i = w_i$ for all integers i between 1 and m. We have already noted that $u_i = w_i$ for all integers i between $m+1$ and n. Consequently $\mathbf{u} = \mathbf{w}$. We conclude therefore that the vector $\mathbf{w}$ is the unique vector in $\mathbb{R}^n$ that satisfies the equation $(D\varphi)_{\mathbf{p}} \mathbf{w} = \mathbf{v}$. We have accordingly established that the derivative $(D\varphi)_{\mathbf{p}} \colon \mathbb{R}^n \to \mathbb{R}^n$ of the function φ at the point $\mathbf{p}$ is an invertible linear operator on $\mathbb{R}^n$.

The Inverse Function Theorem (Theorem 9.4) now ensures the existence of a continuously differentiable function $\mu \colon Y \to \mathbb{R}^n$, defined over an open set Y in $\mathbb{R}^n$, with the properties that $\mu(Y)$ is an open subset of X, $\mathbf{p} \in \mu(Y)$ and $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in Y$.

Let $\mathbf{y}$ be a point of Y, 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 $i = 1, 2, \ldots, m$. Also y_i is equal to the ith component of $\mu(\mathbf{y})$ for all integers i between $m+1$ and n.

Now $\mathbf{p} \in \mu(Y)$. Therefore there exists some point $\mathbf{q}$ of Y satisfying $\mu(\mathbf{q}) = \mathbf{p}$. Now $\mathbf{p} \in S$, 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 the definitions of the functions φ and μ ensure that $q_i = p_i$ for each integer i between $m+1$ and n.

Let g_i denote the ith Cartesian component of the continuously differentiable function $\mu \colon Y \to \mathbb{R}^n$ for $i = 1, 2, \ldots, n$. Then $g_i \colon Y \to \mathbb{R}$ is a continuously differentiable real-valued function on Y for $i = 1, 2, \ldots, n$. If $(y_1, y_2, \ldots, y_n) \in Y$ 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 function φ that y_i is the ith Cartesian component of $\mu(y_1, y_2, \ldots, y_n)$ when $i > m$, and thus

$$y_i = g_i(y_1, y_2, \ldots, y_n) \quad \text{when} \quad m + 1 \leq i \leq n.$$

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

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

Let $V = H(\mathbf{p}, \delta)$ and let

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

Also 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 $\mathbf{x} \in V$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Then $\mathbf{x} \in \mu(Y)$. There therefore exists $\mathbf{q}' \in Y$ for which $\mu(\mathbf{q}') = \mathbf{x}$. But the properties of the function μ ensure that $\mathbf{q}' = \varphi(\mu(\mathbf{q}'))$. It follows that

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

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 $\mathbf{x} \in V \cap S$ 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

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

Now let $\mathbf{x}$ be a point of V 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 $x_i = g_i(\mathbf{y})$ for $i = 1, 2, \ldots, m$, where

$$\mathbf{y} = (0, 0, \ldots, 0, x_{m+1}, \ldots, x_n).$$

However it follows from an identity established at an earlier point of the proof that

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

for all integers i between $m+1$ and n. Consequently $x_i = g_i(\mathbf{y})$ for all integers i between 1 and n, and therefore $\mathbf{x} = \mu(\mathbf{y})$.

The properties that characterize of the function μ then ensure that

$$\varphi(\mathbf{x}) = \mathbf{y} = (0, \ldots, 0, x_{m+1}, \ldots, x_n).$$

Moreover, for each integer i between 1 and m, the ith component function of φ is the function f_i. It follows therefore that $f_i(\mathbf{x}) = 0$ for each integer i between 1 and m, and therefore $\mathbf{x} \in V \cap S$. We can conclude therefore that

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

This completes the proof of the Implicit Function Theorem. ∎

The three following results are special cases of the Implicit Function Theorem, and cover those standard cases in which the theorem is applied to continuously differentiable scalar-valued and vector-valued functions of two or three real variables.

These results are basic building blocks for establishing secure logical foundations for that part of the field of differential geometry that is concerned with the theory of curves and surfaces in low-dimensional Euclidean spaces. Curves and surfaces specified in terms of continuously differentiable functions, and their higher-dimensional analogues in finite-dimensional Euclidean spaces, are examples of *submanifolds* of the Euclidean spaces that contain them. The Implicit Function Theorem generalizes the results concerning curves and surfaces expressed in the following corollaries so as to apply to submanifolds of Euclidean spaces of any finite dimension.

Corollary 9.6 *Let f be a continuously differentiable real-valued function defined over an open set in $\mathbb{R}^2$, and let (p,q) be a point of the domain of the function f. Suppose that $f(p,q) = 0$ and*

$$\frac{\partial f}{\partial y} \neq 0$$

at the point (p,q). Then there exists an open set V in $\mathbb{R}^2$, where $(p,q) \in V$, and a continuously differentiable function h of a single real variable, defined around the real number p, such that

$$\{(x,y) \in V : f(x,y) = 0\} \quad = \quad \{(x,y) \in V : y = h(x)\}.$$

Corollary 9.7 *Let f be a continuously differentiable real-valued function defined over an open set in $\mathbb{R}^3$, and let (p,q,r) be a point of the domain of the function f. Suppose that $f(p,q,r) = 0$ and*

$$\frac{\partial f}{\partial z} \neq 0$$

at the point (p,q,r). Then there exists an open set V in $\mathbb{R}^3$, where $(p,q,r) \in V$, and a continuously differentiable function h of two real variables, defined around the point $(p,q) \in \mathbb{R}^2$, such that

$$\{(x,y,z) \in V : f(x,y,z) = 0\} \quad = \quad \{(x,y,z) \in V : z = h(x,y)\}.$$

Corollary 9.8 *Let v and w be continuously differentiable real-valued functions defined over an open set in $\mathbb{R}^3$, and let (p,q,r) be a point of the common domain of the functions v and w. Suppose that $v(p,q,r) = 0$, $w(p,q,r) = 0$ and*

$$\frac{\partial v}{\partial y}\frac{\partial w}{\partial z} - \frac{\partial v}{\partial z}\frac{\partial w}{\partial y} \neq 0$$

at the point (p,q,r). Then there exists an open set V in $\mathbb{R}^3$, where $(p,q,r) \in V$, and continuously differentiable functions f and g of a single real variable, defined around the real number p, such that

$$\begin{aligned} &\{(x,y,z) \in V : v(x,y,z) = w(x,y,z) = 0\} \\ &\quad = \quad \{(x,y,z) \in V : y = f(x) \text{ and } z = g(x)\}. \end{aligned}$$

Note that the condition imposed on the first order partial derivatives of the function v and w in the statement of Corollary 9.8, requiring the value of

$$\frac{\partial v}{\partial y}\frac{\partial w}{\partial z} - \frac{\partial v}{\partial z}\frac{\partial w}{\partial y}$$

to be non-zero at the point (p, q, r) is a necessary and sufficient condition for ensuring that the matrix

$$\begin{pmatrix} \dfrac{\partial v}{\partial y} & \dfrac{\partial v}{\partial z} \\ \dfrac{\partial w}{\partial y} & \dfrac{\partial w}{\partial z} \end{pmatrix}$$

of functions is an invertible matrix when those functions are evaluated at the point (p, q, r).