Module MA2321: Analysis in Several Real Variables
Michaelmas Term 2015
Section 8

D. R. Wilkins

Contents

8 Curvilinear Coordinates and the Inverse Function Theorem

8.1 Higher Order Derivatives and Smoothness

Definition Let V be an open set in $\mathbb{R}^n$. A function $\varphi: V \to \mathbb{R}^m$ is *continuously differentiable* if the function sending each point $\mathbf{x}$ of V to the derivative $(D\varphi)$ of φ at the point $\mathbf{x}$ is a continuous function from V to $L(\mathbb{R}^n, \mathbb{R}^m)$.

Lemma 8.1 *Let V be an open set in $\mathbb{R}^n$. A function $\varphi: V \to \mathbb{R}^m$ is continuously differentiable if and only if the first order partial derivatives of the components of φ exist and are continuous throughout V.*

Proof The result follows directly from Theorem 7.11. ∎

A function of several real variables is said to be "C^1" if and only if it is continuously differentiable.

The process of differentiation can be repeated. Let $\varphi: V \to \mathbb{R}^m$ be a differentiable function defined over an open set V in $\mathbb{R}^m$. Suppose that the function φ is differentiable at each point $\mathbf{p}$. Then the derivative of φ can itself be regarded as a function on V taking values in the real vector space $L(\mathbb{R}^n, \mathbb{R}^m)$ of linear transformations between the real vector spaces $\mathbb{R}^n$ and $\mathbb{R}^m$. Moreover $L(\mathbb{R}^n, \mathbb{R}^m)$ can itself be regarded as a Euclidean space whose Euclidean norm is the Hilbert-Schmidt norm on $L(\mathbb{R}^n, \mathbb{R}^m)$. It follows that the definition of differentiability can be applied to derivative of a differentiable function of several real variables to obtain the second derivative of a twice-differentiable function. Continuing the process, one can obtain the kth derivative of a k-times differentiable function for any positive integer k. A more detailed analysis of this process shows that if φ is a k-times differentiable function, and if the Cartesian components of φ are $f_1, f_2, \ldots, f_m$, so that

$$\varphi(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))$$

for all $\mathbf{x} \in V$, then the kth derivative of φ at each point of V is represented by the multilinear transformation that maps each k-tuple $(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots, \mathbf{v}^{(k)})$ of vectors in $\mathbb{R}^n$ to the vector in $\mathbb{R}^m$ whose ith component is

$$\sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_k=1}^{n} \frac{\partial^k f_i}{\partial x_{j_1}\, \partial x_{j_2} \cdots \partial x_{j_k}} v_{j_1}^{(1)} v_{j_2}^{(2)} \cdots v_{j_k}^{(k)},$$

where $v_j^{(s)}$ denotes the jth component of the vector $\mathbf{v}^{(s)}$ for $j = 1, 2, \ldots, n$ and $s = 1, 2, \ldots, k$. The kth derivative of the function φ is thus represented by a function from the open set V to some real vector space of multilinear transformations. Such a function is said to be a (Cartesian) *tensor field* on V. Such tensor fields are ubiquitous in differential geometry and theoretical physics.

We can formally define the concept of functions of several variables being *differentiable* of order k by recursion on k.

Definition Let V be an open set in $\mathbb{R}^n$. A function $\varphi: V \to \mathbb{R}^m$ is *k-times differentiable*, where $k > 1$, if it is differentiable and the $D\varphi: V \to L(\mathbb{R}^n, \mathbb{R}^m)$ that maps each point $\mathbf{x}$ of V to the derivative of φ at that point is a $(k-1)$-times differentiable function on V.

Definition Let V be an open set in $\mathbb{R}^n$. A function $\varphi: V \to \mathbb{R}^m$ is *k-times continuously differentiable*, where $k > 1$, if the function $D\varphi: V \to L(\mathbb{R}^n, \mathbb{R}^m)$ that maps each point $\mathbf{x}$ of V to the derivative of φ at that point is a $(k-1)$-times continuously differentiable function on V.

A function of several real variables is said to be "C^k" for some positive integer k if and only if it is k-times continuously differentiable.

Definition A function $\varphi: V \to \mathbb{R}^m$ is said to be *smooth* (or C^∞) if it is k-times differentiable for all positive integers k.

If a function of several real variables is $(k+1)$-times differentiable, then the components of its kth order derivative must be continuous functions, because differentiability implies continuity (see Lemma 7.6). It follows that a function of several real variables is smooth if and only if it is C^k for all positive integers k.

Lemma 8.2 *Let V be an open set in $\mathbb{R}^n$. A function $\varphi: V \to \mathbb{R}^m$ is k-times continuously differentiable (or C^k) if and only if the partial derivatives of the components of φ of all orders up to and including k exist and are continuous throughout V.*

Proof The result can be proved by induction on k. The result is true for $k = 1$ by Lemma 8.1. Suppose as our induction hypothesis that $k > 1$ and that continuously differentiable vector-valued functions on V are C^{k-1} if and only if their partial derivatives of orders up to and including $k-1$ exist and are continuous throughout V. Now a vector-valued function is continuously differentiable if and only if its components are continuously differentiable. Moreover a vector-valued function is C^{k-1} if and only if its components are all C^{k-1}. It follows that the function φ is C^k if and only if the components of its derivative are C^{k-1}. These components are the first-order partial derivatives of φ. The induction hypothesis ensures that these first order partial derivatives of φ are C^{k-1} if and only if their partial derivatives of orders less than or equal to $k-1$ exist and are continuous throughout V. It follows that the function φ itself is C^k if and only if its partial derivatives of orders less than or equal to k exist and are continuous throughout V, as required. ∎

Lemma 8.3 *Let V be an open set in $\mathbb{R}^n$, and let $f\colon V \to \mathbb{R}$ and $g\colon V \to \mathbb{R}$ be real-functions on V, and let $f+g$, $f-g$ and $f\,.\,g$ denote the sum, difference and product of these functions, where*

$$(f+g)(\mathbf{x}) = f(\mathbf{x})+g(\mathbf{x}), \quad (f-g)(\mathbf{x}) = f(\mathbf{x})-g(\mathbf{x}), \quad (f\,.\,g)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$$

for all $\mathbf{x} \in V$. Suppose that the functions f and g are C^k for some positive integer k. Then so are the functions $f+g$, $f-g$ and $f\,.\,g$.

Proof The result can be proved by induction on k. It follows from Theorem 7.8 that the result is true when $k = 1$.

A real-valued function on V is C^k for some positive integer k if and only if all the partial derivatives of its components of degree less than or equal to k exist and are continuous throughout the open set V. It follows from this that a real-valued function f on V is C^k if and only if its first order partial derivatives $\partial_i f$ are C^{k-1}, where $\partial_i f = \dfrac{\partial f}{\partial x_i}$ for $i = 1, 2, \ldots, n$.

Thus suppose as our induction hypothesis that $k > 1$ and that all sums, differences and products of C^{k-1} functions are known to be C^{k-1}. Let f and g be C^k functions. Then

$$\partial_i(f+g) = \partial_i f + \partial_i g, \quad \partial_i(f-g) = \partial_i f - \partial_i g,$$

$$\partial_i(f\,.\,g) = f\,.\,(\partial_i g) + (\partial_i f)\,.\,g$$

for $i = 1, 2, \ldots, n$. Now the functions f, g, $\partial_i f$ and $\partial_i g$ are all C^{k-1}. The induction hypothesis then ensures that $\partial_i(f+g)$, $\partial_i(f-g)$ and $\partial_i(f\,.\,g)$ are all C^{k-1} for $i = 1, 2, \ldots, n$, and therefore the functions $f+g$, $f-g$ and $f\,.\,g$ are C^k.

The required result therefore follows by induction on the degree k of the derivatives required to be continuous. ∎

Lemma 8.4 *Let V and W be open sets in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, and let $\varphi\colon V \to \mathbb{R}^m$ and $\psi\colon W \to \mathbb{R}^l$ be functions mapping V and W into $\mathbb{R}^m$ and $\mathbb{R}^l$ respectively, where $\varphi(V) \subset W$. Suppose that the functions $\varphi\colon V \to \mathbb{R}^m$ and $\psi\colon W \to \mathbb{R}^l$ are C^k. Then the composition function $\psi \circ \varphi\colon V \to \mathbb{R}^l$ is also C^k.*

Proof We prove the result by induction on k. The Chain Rule for functions of several real variables (Theorem 7.9) ensures that the result is true for $k = 1$.

We have shown that sums, differences and products of C^k functions are C^k (see Lemma 8.3). We suppose as our induction hypothesis that all compositions of C^{k-1} functions of several real variables are C^{k-1} for some positive integer k, and show that this implies that all compositions of C^k functions of several real variables are C^k.

Let $\varphi: V \to \mathbb{R}^m$ and $\psi: W \to \mathbb{R}^l$ be C^k functions, where V is an open set in $\mathbb{R}^n$, W is an open set in $\mathbb{R}^m$ and $\varphi(V) = W$. Let the components of φ be $f_1, f_2, \ldots, f_n$ and let the components of ψ be $g_1, g_2, \ldots, g_m$, where $f_1, f_2, \ldots, f_n$ are real-valued functions on V, $g_1, g_2, \ldots, g_m$ are real-valued functions on W,

$$\varphi(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_n(\mathbf{x}))$$

for all $\mathbf{x} \in V$ and

$$\psi(\mathbf{y}) = (g_1(\mathbf{y}), g_2(\mathbf{y}), \ldots, f_m(\mathbf{y}))$$

for all $\mathbf{y} \in W$. It then follows from the Chain Rule (Theorem 7.9) that

$$\frac{\partial}{\partial x_i}\Big(g_j(\varphi(x_1, x_2, \ldots, x_n))\Big) = \sum_{s=1}^{m} \left(\frac{\partial g_j}{\partial u_s} \circ \varphi\right) \frac{\partial f_s}{\partial x_i}.$$

Now the functions $\dfrac{\partial g_j}{\partial u_s} \circ \varphi$ are compositions of C^{k-1} functions. The induction hypothesis therefore ensures that these functions are C^{k-1}. This then ensures that the functions $\dfrac{\partial}{\partial x_i}\Big(g_j(\varphi(x_1, x_2, \ldots, x_n))\Big)$ are expressible as sums of products of C^{k-1} functions, and must therefore themselves be C^{k-1} functions (see Lemma 8.3). We have thus shown that the first order partial derivatives of the components of the composition function $\psi \circ \varphi$ are C^{k-1} functions. It follows that $\psi \circ \varphi$ must itself be a C^k function.

The required result therefore follows by induction on the degree k of the derivatives required to be continuous. ∎

It follows from Lemma 8.3 and Lemma 8.4 that functions that are constructed from smooth vector-valued functions defined over open sets in Euclidean spaces by means of the operations of additions, subtraction, multiplication and composition of functions must themselves be smooth functions over the open sets over which they are defined.

We now prove a lemma that guarantees the smoothness of matrix-valued functions obtained from smooth matrix-valued functions through the operation of matrix inversion. The lemma applies to functions $F: V \to \mathrm{GL}(m, \mathbb{R})$

defined over an open subset V of a Euclidean space $\mathbb{R}^n$ and taking values in the set $\mathrm{GL}(m, \mathbb{R})$ of invertible $m \times m$ matrices. The value $F(\mathbf{x})$ of such a function at a point $\mathbf{x}$ of V is thus an invertible $m \times m$ matrix, and thus the function $F: V \to \mathrm{GL}(m, \mathbb{R})$ determines a corresponding function $G: V \to \mathrm{GL}(m, \mathbb{R})$, where $G(\mathbf{x}) = F(\mathbf{x})^{-1}$ for all $\mathbf{x} \in V$. The coefficients of the matrices $F(\mathbf{x})$ and $G(\mathbf{x})$ are then functions of $\mathbf{x}$ as $\mathbf{x}$ varies over the open set V. Now the function F is C^k if and only if, for all i and j between 1 and m, the coefficient of the matrix $F(\mathbf{x})$ in the ith row and jth column is a C^k function of $\mathbf{x}$ throughout the open set V. We prove that if the function F is C^k for some positive integer k then the function G is also C^k. It follows that if the function F is smooth, then the function G is smooth.

Lemma 8.5 *Let m be a positive integer, let $M_m(\mathbb{R})$ denote the real vector space consisting of all $m \times m$ matrices with real coefficients, and let $\mathrm{GL}(m, \mathbb{R})$ be the open set in $M_m(\mathbb{R})$ whose elements are the invertible $m \times m$ matrices with real coefficients. Let V be an open set in $\mathbb{R}^n$ let $F: V \to \mathrm{GL}(m, \mathbb{R})$ be a function mapping V into $\mathrm{GL}(m, \mathbb{R})$, and let $G: V \to \mathrm{GL}(m, \mathbb{R})$ be defined such that $G(\mathbf{x}) = F(\mathbf{x})^{-1}$ for all $\mathbf{x} \in V$. Suppose that the function F is C^k. Then the function G is C^k.*

Proof For each $\mathbf{x} \in V$, the matrices $F(\mathbf{x})$ and $G(\mathbf{x})$ satisfy $F(\mathbf{x})G(\mathbf{x}) = I$, where I is the identity matrix. On differentiating this identity with respect to the ith coordinate function x_i on V, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, we find that

$$\frac{\partial F(\mathbf{x})}{\partial x_i} G(\mathbf{x}) + F(\mathbf{x}) \frac{\partial G(\mathbf{x})}{\partial x_i} = 0,$$

and therefore

$$\frac{\partial G(\mathbf{x})}{\partial x_i} = -F(\mathbf{x})^{-1} \frac{\partial F(\mathbf{x})}{\partial x_i} G(\mathbf{x}) = -G(\mathbf{x}) \frac{\partial F(\mathbf{x})}{\partial x_i} G(\mathbf{x}).$$

(In the above equation $F(\mathbf{x})$, $G(\mathbf{x})$ and their inverses and partial derivatives are $m \times m$ matrices that are multiplied using the standard operation of matrix multiplication.) Now sums and products of C^k real-valued functions are themselves C^k (see Lemma 8.3). It follows that if matrices are multiplied together, where the coefficients of those matrices are C^k real-valued functions defined over the open set V, the coefficients of the resultant matrix will also be C^k real-valued functions defined over V.

The equation above ensures that if the matrix-valued function F is C^k (so that the functions determining the coefficients of the matrix are real-valued C^k functions on V), then the first order partial derivatives of the function G are continuous, and therefore the function G itself is C^1, where

$G(\mathbf{x}) = F(\mathbf{x})^{-1}$ for all $\mathbf{x} \in V$. Moreover if G is C^j, where $1 \leq j < k$ then the coefficients of the first order partial derivatives of G are expressible as a sums of products of C^j real-valued functions and thus are themselves C^j functions. Thus the matrix-valued function G itself is C^{j+1}. Repeated applications of this result ensure that G is a C^k function as required. ∎

8.2 Lipschitz Conditions satisfied locally by Continuously Differentiable Functions

Let $\varphi: X \to \mathbb{R}^m$ be a function defined over a subset X of $\mathbb{R}^n$. The function V is said to satisfy a *Lipschitz condition* with *Lipschitz constant* M on X if the inequality

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

satisfied for all points $\mathbf{x}$ and $\mathbf{x}$ of X. A function on X that satisfies such a Lipschitz condition is said to be *Lipschitz continuous* on X.

A standard theorem (often referred to as *Picard's Theorem*) in the theory of ordinary diffential equations guaranteeing the existence and uniqueness of solutions of initial value problems is only applicable when the function determining the differential equation satisfies an appropriate Lipschitz condition.

We use the result of Proposition 7.10 to show that continuously differentiable functions satisfy Lipschitz conditions with arbitrarily small Lipschitz constants in the neighbourhood around points where their derivative is zero.

Proposition 8.6 *Let $\varphi: V \to \mathbb{R}^m$ be a continuously differentiable function defined over an open set V in $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of V at which $(D\varphi)_{\mathbf{p}} = 0$. Then, given any positive real number λ, there exists some positive real number δ such that*

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

for all points $\mathbf{x}$ and $\mathbf{x}'$ of V that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$ and $|\mathbf{x}' - \mathbf{p}| < \delta$.

Proof Let $\mathbf{p} = (p_1, p_2, \ldots, p_n)$, and let

$$\varphi(\mathbf{x}) = (f_1(\mathbf{x}),\ f_2(\mathbf{x}), \ldots\ f_m(\mathbf{x}))$$

for all $\mathbf{x} \in V$. Then the derivative $(D\varphi)_{\mathbf{x}}$ of φ at a point $\mathbf{x}$ of V is represented by value at $\mathbf{x}$ of the Jacobian matrix whose coefficients are the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ for $i, j = 1, 2, \ldots, n$. Now the first order partial derivatives of the functions $f_1, f_2, \ldots, f_m$ are continuous, because φ is a continuously differentiable function. It follows that there exists some positive real number δ such that

$$\left|\frac{\partial f_i}{\partial x_j}\right| < \frac{\lambda}{\sqrt{mn}}$$

at all points $(x_1, x_2, \ldots, x_n)$ that satisfy $|x_j - p_j| < \delta$ for $i = 1, 2, \ldots, n$. It then follows from Proposition 7.10 that if the points $\mathbf{x}$ and $\mathbf{x}'$ of V satisfy $|\mathbf{x} - \mathbf{p}| < \delta$ and $|\mathbf{x}' - \mathbf{p}| < \delta$ then

$$|f_i(\mathbf{x}') - f_i(\mathbf{x})| \leq \frac{\lambda}{\sqrt{m}} |\mathbf{x}' - \mathbf{x}|$$

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

$$|\varphi(\mathbf{x}') - \varphi(\mathbf{x})|^2 = \sum_{i=1}^{m} |f_i(\mathbf{x}') - f_i(\mathbf{x})|^2 \leq \lambda^2 |\mathbf{x}' - \mathbf{x}|^2.$$

and therefore $|\varphi(\mathbf{x}') - \varphi(\mathbf{x})| \leq \lambda |\mathbf{x}' - \mathbf{x}|$, as required. ∎

We shall apply Proposition 8.6 in order to prove a result that yields a Lipschitz condition satisfied by continuously differentiable functions. The statement of the result will make reference to the *operator norm* of a linear transformation. We therefore proceed by giving the definition of the operator norm of a linear transformation between (finite-dimensional) Euclidean spaces.

Definition Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation between Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ The *operator norm* $\|T\|_{\mathrm{op}}$ of the linear transformation T is defined so that

$$\|T\|_{\mathrm{op}} = \sup\{|T\mathbf{v}| : \mathbf{v} \in \mathbb{R}^n \text{ and } |\mathbf{v}| = 1\}.$$

Let T, T_1 and T_2 be linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ and let c be a real number. Let $\mathbf{v}$ be a non-zero vector in $\mathbb{R}^\ltimes$, and let $\hat{\mathbf{v}} = |\mathbf{v}|^{-1}\mathbf{v}$. Then

$$|T\mathbf{v}| = ||\mathbf{v}| \, (T\hat{\mathbf{v}})| = |\mathbf{v}| \, |T\hat{\mathbf{v}}| \leq \|T\|_{\mathrm{op}} |\mathbf{v}|.$$

Also $|T\mathbf{v}| = 0$ when $\mathbf{v} = \mathbf{0}$. It follows that $|T\mathbf{v}| \leq \|T\|_{\mathrm{op}} |\mathbf{v}|$ for all $\mathbf{v} \in \mathbb{R}^n$. Also

$$|(T_1 + T_2)\mathbf{v}| = |T_1\mathbf{v} + T_2\mathbf{v}| \leq |T_1\mathbf{v}| + |T_2\mathbf{v}| \leq (\|T_1\|_{\mathrm{op}} + \|T_2\|_{\mathbf{op}}) |\mathbf{v}|$$

for all $\mathbf{v} \in \mathbb{R}^n$. It follows that $\|T_1 + T_2\|_{\mathrm{op}} \leq \|T_1\|_{\mathrm{op}} + \|T_2\|_{\mathbf{op}}$. Also $|(cT)\mathbf{v}| = |c| \, |T\mathbf{v}|$ for all $\mathbf{v} \in \mathbb{R}^n$, and therefore $\|cT\|_{\mathrm{op}} = |c| \, \|T\|_{\mathrm{op}}$. The linear transformation T satisfies $\|T\|_{\mathrm{op}} = 0$ if and only if $T = 0$.

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ and $S: \mathbb{R}^m \to \mathbb{R}^l$ be linear transformations. Then

$$|ST\mathbf{v}| \leq \|S\|_{\mathrm{op}} |T\mathbf{v}| \leq \|S\|_{\mathrm{op}} \, \|T\|_{\mathrm{op}} |\mathbf{v}|,$$

and therefore $\|ST\|_{\mathrm{op}} \leq \|S\|_{\mathrm{op}} \, \|T\|_{\mathrm{op}}$.

It was shown in Lemma 7.1 that $|T\mathbf{v}| \leq \|T\|_{\mathrm{HS}} |\mathbf{v}|$ for all $\mathbf{v} \in \mathbb{R}^n$, where $\|T\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm of the linear operator T. It follows that $\|T\|_{\mathrm{op}} \leq \|T\|_{\mathrm{HS}}$.

Corollary 8.7 *Let $\varphi: V \to \mathbb{R}^m$ be a continuously differentiable function defined over an open set V in $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of V. Let M be a positive real number satisfying $M > \|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}}$, where*

$$\|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}} = \sup\{|(D\varphi)_{\mathbf{p}}\mathbf{v}| : \mathbf{v} \in \mathbb{R}^n \text{ and } |\mathbf{v}| = 1\}.$$

Then there exists a positive real number δ such that

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

for all points $\mathbf{x}$ and $\mathbf{x}'$ of V that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$ and $|\mathbf{x}' - \mathbf{p}| < \delta$.

Proof Let $T = (D\varphi)_{\mathbf{p}}$, and let

$$M_0 = \|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}} = \sup\{|T\mathbf{v}| : \mathbf{v} \in \mathbb{R}^n \text{ and } |\mathbf{v}| = 1\},$$

and let $\lambda = M - M_0$. Let $\varphi: V \to \mathbb{R}^m$ be defined such that

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

for all $\mathbf{x} \in V$. Then $(D\psi)_{\mathbf{p}} = (D\varphi)_{\mathbf{p}} - T = 0$. It follows from Proposition 8.6 that there exists a positive real number δ such that

$$|\psi(\mathbf{x}) - \psi(\mathbf{x}')| \leq \lambda|\mathbf{x} - \mathbf{x}'|$$

for all points $\mathbf{x}$ and $\mathbf{x}'$ of V that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$ and $|\mathbf{x}' - \mathbf{p}| < \delta$. Then

$$\begin{aligned} |\varphi(\mathbf{x}) - \varphi(\mathbf{x}')| &= |\psi(\mathbf{x}) - \psi(\mathbf{x}') + T(\mathbf{x} - \mathbf{x}')| \\ &\leq |\psi(\mathbf{x}) - \psi(\mathbf{x}')| + |T(\mathbf{x} - \mathbf{x}')| \\ &\leq \lambda|\mathbf{x} - \mathbf{x}'| + M_0|\mathbf{x} - \mathbf{x}'| = M|\mathbf{x} - \mathbf{x}'| \end{aligned}$$

for all points $\mathbf{x}$ and $\mathbf{x}'$ of V that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$ and $|\mathbf{x}' - \mathbf{p}| < \delta$, as required. ∎

Corollary 8.7 ensures that continuously differentiable functions of several real variables are *locally Lipschitz continuous*. This means that they satisfy a Lipschitz condition in some sufficiently small neighbourhood of any given point. This in turn ensures that standard theorems concerning the existence and uniqueness of ordinary differential equations can be applied to systems of ordinary differential equations specified in terms of continuously differentiable functions.

8.3 Local Invertibility of Differentiable Functions

Definition Let $\varphi: V \to \mathbb{R}^n$ be a continuous function defined over an open set V in $\mathbb{R}^n$ and mapping that open set into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of V. A *local inverse* of the map $\varphi: V \to \mathbb{R}^n$ around the point $\mathbf{p}$ is a continuous function $\mu: W \to V$ defined over an open set W in $\mathbb{R}^n$ that satisfies the following conditions:

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

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

If there exists a function $\mu: W \to V$ satisfying these conditions, then the function φ is said to be *locally invertible* around the point $\mathbf{p}$.

Lemma 8.8 *Let $\varphi: V \to \mathbb{R}^n$ be a continuous function defined over an open set V in $\mathbb{R}^n$ and mapping that open set into $\mathbb{R}^n$, let $\mathbf{p}$ be a point of V. and let $\mu: W \to V$ be a local inverse for the map ϕ around the point $\mathbf{p}$. Then $\varphi(\mathbf{x}) \in W$ and $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in \mu(W)$.*

Proof The definition of local inverses ensures that $\mu(W)$ is an open subset of V, $\mathbf{p} \in \mu(W)$ and $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$. Let $\mathbf{x} \in \mu(W)$. Then $\mathbf{x} = \mu(\mathbf{y})$ for some $\mathbf{y} \in W$. But then $\varphi(\mathbf{x}) = \varphi(\mu(\mathbf{y})) = \mathbf{y}$, and therefore $\varphi(\mathbf{x}) \in W$. Moreover $\mu(\varphi(\mathbf{x})) = \mu(\mathbf{y}) = \mathbf{x}$, as required. ∎

Let $\varphi: V \to \mathbb{R}^n$ be a continuous function defined over an open set V in $\mathbb{R}^n$ and mapping that open set into $\mathbb{R}^n$, let $\mathbf{p}$ be a point of V. and let $\mu: W \to V$ be a local inverse for the map ϕ around the point $\mathbf{p}$. Then the function from the open set $\mu(W)$ to the open set W that sends each point $\mathbf{x}$ of $\mu(W)$ to $\varphi(x)$ is invertible, and its inverse is the continuous function from W to $\varphi(W)$ that sends each point $\mathbf{y}$ of W to $\mu(\mathbf{y})$. A function between sets is *bijective* if it has a well-defined inverse. A continuous bijective function whose inverse is also continuous is said to be a *homeomorphism*. We see therefore that the restriction of the map φ to the image $\mu(W)$ of the local inverse $\mu: W \to V$ determines a homeomorphism from the open set $\mu(W)$ to the open set W.

Example The function $\varphi: \mathbb{R}^2 \to \mathbb{R}^2 \setminus \{(0,0)\}$ defined such that

$$\varphi(u, v) = (e^u \cos v, e^u \sin v)$$

for all $u, v \in \mathbb{R}^2$ is locally invertible, though it is not bijective. Indeed, given $(u_0, v_0) \in \mathbb{R}$, let

$$W = \{(r \cos(v_0 + \theta), r \sin(v_0 + \theta)) : r, \theta \in \mathbb{R}, \;\; r > 0 \text{ and } -\pi < \theta < \pi\},$$

and let

$$\mu(r\cos(v_0+\theta), r\sin(v_0+\theta)) = (\log r, v_0+\theta)$$

whenever $r > 0$ and $-\pi < \theta < 1$. Then W is an open set in $\mathbb{R}^2$,

$$\mu(W) = \{(u,v) \in \mathbb{R}^2 : v_0 - \pi < v < v_0 + \pi\},$$

and $\mu(\varphi(u,v)) = (u,v)$ for all $(u,v) \in \mu(W)$. Note that the smoothness of the logarithm and inverse trigonometrical functions guarantees that the local inverse $\mu: W \to \mathbb{R}^2$ is itself smooth.

A smooth function may have a continuous inverse, but that inverse is not guaranteed to be differentiable, as the following example demonstrates.

Example Let $f: \mathbb{R} \to \mathbb{R}$ be defined so that $f(x) = x^3$ for all real numbers x. The function f is smooth and has a continuous inverse $f^{-1}: \mathbb{R} \to \mathbb{R}$, where $f^{-1}(x) = \sqrt[3]{x}$ when $x \geq 0$ and $f^{-1}(x) = -\sqrt[3]{-x}$ when $x < 0$. This inverse function is not differentiable at zero.

Lemma 8.9 *Let $\varphi: V \to \mathbb{R}^n$ be a continuously differentiable function defined over an open set V in $\mathbb{R}^n$. Suppose that φ is locally invertible around some point $\mathbf{p}$ of V. Suppose also that a local inverse to φ around $\mathbf{p}$ is differentiable at the point $\varphi(\mathbf{p})$. Then 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 $\mathbb{R}^n$. Thus if*

$$\varphi(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n),$$

for all $(x_1, x_2, \ldots, x_n) \in V$, where $y_1, y_2, \ldots, y_n$ are differentiable functions of $x_1, x_2, \ldots, x_n$, and if φ has a differentiable local inverse around the point $\mathbf{p}$, then the Jacobian matrix

$$\begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_1}{\partial x_2} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \dfrac{\partial y_2}{\partial x_1} & \dfrac{\partial y_2}{\partial x_2} & \cdots & \dfrac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial y_n}{\partial x_1} & \dfrac{\partial y_n}{\partial x_2} & \cdots & \dfrac{\partial y_n}{\partial x_n} \end{pmatrix}$$

is invertible at the point $\mathbf{p}$.

Proof Let $\mu: W \to V$ be a local inverse of φ around $\mathbf{p}$, where W is an open set in $\mathbb{R}^n$, $\mathbf{p} \in \mu(W)$, $\mu(W) \subset V$ and $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in \mu(W)$. Suppose that $\mu: W \to V$ is differentiable at $\varphi(\mathbf{p})$. The identity $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ holds throughout the open neighbourhood $\mu(W)$ of point $\mathbf{p}$. Applying the Chain Rule (Theorem 7.9), we find that $(D\mu)_{\varphi(\mathbf{p})}(D\varphi)_{\mathbf{p}}$ is the identity operator on $\mathbb{R}^n$. It follows that the linear operators $(D\mu)_{\varphi(\mathbf{p})}$ and $(D\varphi)_{\mathbf{p}}$ on $\mathbb{R}^n$ are inverses of one another, and therefore $(D\varphi)_{\mathbf{p}}$ is an invertible linear operator on $\mathbb{R}^n$. The result follows. ∎

Lemma 8.10 *Let $\varphi: V \to \mathbb{R}^n$ be a continuously differentiable function defined over an open set V in $\mathbb{R}^n$ that is locally invertible around some point of V and let $\mu: W \to \mathbb{R}^n$ be a local inverse for φ. Suppose that $\varphi: V \to \mathbb{R}^n$ is continuously differentiable and that the local inverse $\mu: W \to \mathbb{R}^n$ is Lipschitz continuous throughout W. Then $\mu: W \to \mathbb{R}^n$ is continuously differentiable throughout W.*

Proof The function $\mu: W \to \mathbb{R}^n$ is Lipschitz continuous, and therefore there exists a positive constant C such that

$$|\mu(\mathbf{y}) - \mu(\mathbf{y}')| \leq C\,|\mathbf{y} - \mathbf{y}'|$$

for all $\mathbf{q}, \mathbf{y} \in W$. Let $\mathbf{q} \in W$, let $\mathbf{p} = \mu(\mathbf{q})$, and let S be the derivative of φ at $\mathbf{p}$. Then

$$S\mathbf{v} = \lim_{t \to 0} \frac{1}{t}(\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p}))$$

for all $\mathbf{v} \in \mathbb{R}^n$ (see Lemma 7.7). Now it follows from an inequality previously established that

$$|t||\mathbf{v}| \leq C\,|\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{v})|.$$

for all values of t sufficiently close to zero. It follows that

$$|S\mathbf{v}| = \lim_{t \to 0} \frac{1}{|t|}|\varphi(\mathbf{p} + t\mathbf{v}) - \varphi(\mathbf{p})| \geq \frac{1}{C}|\mathbf{v}|$$

for all $\mathbf{v} \in \mathbb{R}^n$, and therefore $S\mathbf{v} \neq \mathbf{0}$ for all non-zero vectors $\mathbf{v}$. It follows from basic linear algebra that the linear operator S on $\mathbb{R}^n$ is invertible. Moreover $|S^{-1}\mathbf{v}| \leq C|\mathbf{v}|$ for all $\mathbf{v} \in \mathbb{R}^n$.

Now

$$\lim_{\mathbf{x} \to \mathbf{p}} \frac{1}{|\mathbf{x} - \mathbf{p}|}\,|\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - S(\mathbf{x} - \mathbf{p})| = 0,$$

because the function φ is differentiable at $\mathbf{p}$. Now $\mu(\mathbf{y}) \neq \mathbf{p}$ when $\mathbf{y} \neq \mathbf{q}$, because $\mathbf{q} = \varphi(\mathbf{p})$ and $\mathbf{y} = \varphi(\mu(\mathbf{y}))$. Also the continuity of μ ensures that $\mu(\mathbf{y})$ tends to $\mathbf{p}$ as $\mathbf{y}$ tends to $\mathbf{q}$. It follows that

$$\lim_{\mathbf{y} \to \mathbf{q}} \frac{1}{|\mu(\mathbf{y}) - \mathbf{p}|}\,|\mathbf{y} - \mathbf{q} - S(\mu(\mathbf{y}) - \mathbf{p})| = 0.$$

Now

$$|S^{-1}(\mathbf{y}-\mathbf{q}) - (\mu(\mathbf{y}) - \mathbf{p})| \leq C|\mathbf{y}-\mathbf{q} - S(\mu(\mathbf{y}) - \mathbf{p})|$$

for all $\mathbf{y} \in W$. Also

$$\frac{1}{|\mathbf{y}-\mathbf{q}|} \leq \frac{C}{|\mathbf{p} - \mu(\mathbf{y})|}$$

for all $\mathbf{y} \in W$ satisfying $\mathbf{y} \neq \mathbf{q}$. It follows that

$$\frac{1}{|\mathbf{y}-\mathbf{q}|}|\mu(\mathbf{y}) - \mathbf{p} - S^{-1}(\mathbf{y}-\mathbf{q})| \leq \frac{C^2}{|\mu(\mathbf{y}) - \mathbf{p}|}\,|\mathbf{y}-\mathbf{q} - S(\mu(\mathbf{y}) - \mathbf{p})|.$$

It follows that

$$\lim_{\mathbf{y}\to\mathbf{q}} \frac{1}{|\mathbf{y}-\mathbf{q}|}|\mu(\mathbf{y}) - \mathbf{p} - S^{-1}(\mathbf{y}-\mathbf{q})| = 0,$$

and therefore the function μ is differentiable at $\mathbf{q}$ with derivative S^{-1}. Thus $(D\mu)_{\mathbf{q}} = (D\varphi)_{\mathbf{p}}^{-1}$ for all $\mathbf{q} \in W$. It follows from this that $(D\mu)_{\mathbf{q}}$ depends continuously on $\mathbf{q}$, and thus the function μ is continuously differentiable on W, as required. ∎

Lemma 8.11 *Let $\varphi: V \to \mathbb{R}^n$ be a continuously differentiable function defined over an open set V in $\mathbb{R}^n$ that is locally invertible around some point of V and let $\mu: W \to \mathbb{R}^n$ be a local inverse for φ. Suppose that $\varphi: V \to \mathbb{R}^n$ is C^k and that the local inverse $\mu: W \to \mathbb{R}^n$ is differentiable throughout W. Then $\mu: W \to \mathbb{R}^n$ is C^k throughout W.*

Proof The functions φ and μ are differentiable, and $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in \mu(W)$. The Chain Rule (Theorem 7.9) then ensures that $(D\mu)_{\varphi(\mathbf{x})} \circ (D\varphi)_{\mathbf{x}}$ is the identity operator. Let $F(\mathbf{x})$ denote the Jacobian matrix representing the derivative $(D\varphi)_{\mathbf{x}}$ of φ at each point $\mathbf{x}$ of $\mu(W)$, and let $G(\mathbf{x})$ denote the Jacobian matrix representing the derivative $(D\mu)_{\varphi(\mathbf{x})}$ of μ at $\varphi(\mathbf{x})$. Then the Chain Rule ensures that $G(\mathbf{x})F(\mathbf{x})$ is the identity matrix. It follows that $F(\mathbf{x})$ and $G(\mathbf{x})$ are invertible matrices and $G(\mathbf{x}) = F(\mathbf{x})^{-1}$ for all $\mathbf{x} \in \mu(W)$. Now the function φ is C^k on V and therefore the matrix-valued function $F: \mu(W) \to \mathrm{GL}(n, \mathbb{R})$ is is C^k on $\mu(W)$. It follows from Lemma 8.5 that the matrix-valued function $G: \mu(W) \to \mathrm{GL}(n, \mathbb{R})$ is also C^k on $\mu(W)$.

Now the $(D\mu)_{\mathbf{y}}$ is represented by the matrix $G(\mu(\mathbf{y}))$ for all $\mathbf{y} \in W$. It follows from the continuity of μ and G that the derivative $D\mu$ of μ is continuous on W. It follows that μ is C^1. Moreover if $\mu: W \to V$ is C^j for any integer j satisfying $1 \leq j < k$ then $G \circ \mu$ is a composition of C^j functions and is therefore C^j (Lemma 8.4). But the coefficients of the matrix $G(\mu(\mathbf{y}))$ are the first order partial derivatives of the components of μ at $\mathbf{y}$ at each point $\mathbf{y}$ of W. It follows therefore that the first order partial derivatives of μ are C^j and therefore the function μ itself is C^{j+1}. It follows by repeated application of this process that the function μ is C^k on W, as required. ∎

8.4 The Inverse Function Theorem

The *Inverse Function Theorem* ensures that, for a C^k function of several real variables, mapping an open set in one Euclidean space into a Euclidean space of the same dimension, the invertibility of the derivative of the function at a given point is sufficient to ensure the local invertibility of that function around the given point, and moreover ensures that the inverse function is also locally a C^k function.

The proof uses the method of successive approximations, using a convergence criterion for infinite sequences of points in Euclidean space that we establish in the following lemma.

Lemma 8.12 *Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ be an infinite sequence of points in n-dimensional Euclidean space $\mathbb{R}^n$, and let λ be a real number satisfying $0 < \lambda < 1$. Suppose that*

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

for all integers j satisfying $j > 1$. Then the infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is convergent.

Proof We show that an infinite sequence of points in Euclidean space satisfying the stated criterion is a Cauchy sequence and is therefore convergent. Now the infinite sequence satisfies

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

for all positive integers j, where $C = |\mathbf{x}_2 - \mathbf{x}_1|/\lambda$. Let j and k be positive integers satisfying $j < k$. Then

$$\begin{aligned} |\mathbf{x}_k - \mathbf{x}_j| &= \left| \sum_{s=j}^{k-1} (\mathbf{x}_{s+1} - \mathbf{x}_s) \right| \leq \sum_{s=j}^{k-1} |\mathbf{x}_{s+1} - \mathbf{x}_s| \\ &\leq C \sum_{s=j}^{k-1} \lambda^s = C\lambda^j \frac{1 - \lambda^{k-j}}{1 - \lambda} < \frac{C\lambda^j}{1 - \lambda}. \end{aligned}$$

We now show that the infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is a Cauchy sequence. Let some positive real number ε be given. Then a positive integer N can be chosen large enough to ensure that $C\lambda^N < (1 - \lambda)\varepsilon$. Then

$|\mathbf{x}_k - \mathbf{x}_j| < \varepsilon$ whenever $j \geq N$ and $k \geq N$. Therefore the given infinite sequence is a Cauchy sequence. Now all Cauchy sequences in $\mathbb{R}^n$ are convergent (see Lemma 6.4). Therefore the given infinite sequence is convergent, as required. ∎

Theorem 8.13 (Inverse Function Theorem) *Let $\varphi: V \to \mathbb{R}^n$ be a continuously differentiable function defined over an open set V in n-dimensional Euclidean space $\mathbb{R}^n$ and mapping V into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of V. Suppose that $k \geq 1$ and that the derivative $(D\varphi)_{\mathbf{p}}: \mathbb{R}^n \to \mathbb{R}^n$ 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 $\mu: W \to V$ that satisfies the following conditions:—*

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

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

Moreover if the function $\varphi: V \to \mathbb{R}^n$ is C^k for some positive integer k, then so is the function $\mu: W \to V$.

Proof We may assume, without loss of generality, that $\mathbf{p} = \mathbf{0}$ and $\varphi(\mathbf{p}) = \mathbf{0}$. Indeed the result in the general case can then be deduced by applying the result in this special case to the function that sends $\mathbf{z}$ to $\varphi(\mathbf{p} + \mathbf{z}) - \varphi(\mathbf{p})$ for all $\mathbf{z} \in \mathbb{R}^n$ for which $\mathbf{p} + \mathbf{z} \in V$.

Now $(D\varphi)_{\mathbf{0}}: \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear transformation, by assumption. Let $T = (D\varphi)_{\mathbf{0}}^{-1}$, and let $\psi: V \to \mathbb{R}^n$ be defined such that

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

for all $\mathbf{x} \in V$. Now the derivative of any linear transformation at any point is equal to that linear transformation (see Lemma 7.4). It follows from the Chain Rule that the derivative of the composition function $T \circ \varphi$ at any point $\mathbf{x}$ of V is equal to $T(D\varphi)_{\mathbf{x}}$. It follows that $(D\psi)_{\mathbf{x}} = I - T(D\varphi)_{\mathbf{x}}$ for all $\mathbf{x} \in V$, where I denotes the identity operator on $\mathbb{R}^n$. In particular $(D\psi)_{\mathbf{0}} = I - T(D\varphi)_{\mathbf{0}} = 0$. It then follows from Proposition 8.6 that there exists a positive real number δ such that

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

whenever $|\mathbf{x}| < \delta$ and $|\mathbf{x}'| < \delta$.

Now $\psi(\mathbf{0}) = \mathbf{0}$. It follows from the inequality just proved that $|\psi(\mathbf{x})| \leq \frac{1}{2}|\mathbf{x}|$ whenever $|\mathbf{x}| < \delta$.

Let W be the open set in $\mathbb{R}^n$ defined so that

$$W = \{\mathbf{y} \in \mathbb{R}^n : |T(\mathbf{y})| < \tfrac{1}{2}\delta\},$$

and let $\mu_0, \mu_1, \mu_2, \ldots$ be the infinite sequence of functions from W to $\mathbb{R}^n$ defined so that $\mu_0(\mathbf{y}) = 0$ for all $\mathbf{y} \in W$ and

$$\mu_j(\mathbf{y}) = \mu_{j-1}(\mathbf{y}) + T(\mathbf{y} - \varphi(\mu_{j-1}(\mathbf{y})))$$

for all positive integers j. We shall prove that there is a well-defined function $\mu: W \to \mathbb{R}^n$ defined such that $\mu(\mathbf{y}) = \lim\limits_{j\to+\infty} \mu_j(\mathbf{y})$ and that this function μ is a local inverse for φ defined on the open set W that satisfies the required properties.

Let $\mathbf{y} \in W$ and let $\mathbf{x}_j = \mu_j(\mathbf{y})$ for all non-negative integers j. Then $\mathbf{x}_0 = \mathbf{0}$ and

$$\begin{aligned} \mathbf{x}_j &= \mathbf{x}_{j-1} + T(\mathbf{y} - \varphi(\mathbf{x}_{j-1})) \\ &= \psi(\mathbf{x}_{j-1}) + T\mathbf{y} \end{aligned}$$

for all positive integers j. Now we have already shown that $|\psi(\mathbf{x})| \leq \frac{1}{2}|\mathbf{x}|$ whenever $|\mathbf{x}| < \delta$. Also the definition of the open set W ensures that $|T\mathbf{y}| < \frac{1}{2}\delta$. It follows that if $|\mathbf{x}_{j-1}| < \delta$ then

$$|\mathbf{x}_j| \leq |\psi(\mathbf{x}_{j-1})| + |T\mathbf{y}| \leq \tfrac{1}{2}|\mathbf{x}_{j-1}| + |T\mathbf{y}| < \tfrac{1}{2}\delta + |T\mathbf{y}| < \delta.$$

It follows by induction on j that $|\mathbf{x}_j| < \frac{1}{2}\delta + |T\mathbf{y}|$ for all non-negative integers j. Also

$$\begin{aligned} \mathbf{x}_{j+1} - \mathbf{x}_j &= \mathbf{x}_j - \mathbf{x}_{j-1} - T(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_{j-1})) \\ &= \psi(\mathbf{x}_j) - \psi(\mathbf{x}_{j-1}) \end{aligned}$$

for all positive integers j. But $|\mathbf{x}_j| < \delta$ and $|\mathbf{x}_{j-1}| < \delta$ and therefore

$$|\mathbf{x}_{j+1} - \mathbf{x}_j| = |\psi(\mathbf{x}_j) - \psi(\mathbf{x}_{j-1})| \leq \tfrac{1}{2}|\mathbf{x}_j - \mathbf{x}_{j-1}|$$

for all positive integers j. It then follows from Lemma 8.12 that the infinite sequence $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ is convergent. Now $\mathbf{x}_j = \mu_j(\mathbf{y})$ for all non-negative integers j, where $\mathbf{y}$ is an arbitrary element of the open set W. The convergence result just obtained therefore guarantees that there is a well-defined function $\mu: W \to \mathbb{R}^n$ which satisfies

$$\mu(\mathbf{y}) = \lim_{j\to+\infty} \mu_j(\mathbf{y})$$

for all $\mathbf{y} \in W$. Moreover $|\mu_j(\mathbf{y})| < \frac{1}{2}\delta + |T\mathbf{y}|$ for all positive integers j and for all $\mathbf{y} \in W$, and therefore

$$|\mu(\mathbf{y})| \leq \tfrac{1}{2}\delta + |T\mathbf{y}| < \delta$$

for all $\mathbf{y} \in W$.

Next we prove that $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$. Now

$$\begin{aligned} \mu(\mathbf{y}) &= \lim_{j \to +\infty} \mu_j(\mathbf{y}) = \lim_{j \to +\infty} (\mu_{j-1}(\mathbf{y}) + T(\mathbf{y} - \varphi(\mu_{j-1}(\mathbf{y})))) \\ &= \mu(\mathbf{y}) + T(\mathbf{y} - \varphi(\mu(\mathbf{y}))) \end{aligned}$$

It follows that $T(\mathbf{y} - \varphi(\mu(\mathbf{y}))) = \mathbf{0}$, and therefore

$$\mathbf{y} - \varphi(\mu(\mathbf{y})) = (D\varphi)_{\mathbf{0}}(T(\mathbf{y} - \varphi(\mu(\mathbf{y})))) = (D\varphi)_{\mathbf{0}}(\mathbf{0}) = \mathbf{0}.$$

Thus $\mathbf{y} = \varphi(\mu(\mathbf{y}))$ for all $\mathbf{y} \in W$. Also $\mu_j(\mathbf{0}) = \mathbf{0}$ for all non-negative integers j, and therefore $\mu(\mathbf{0}) = \mathbf{0}$.

Next we show that if $\mathbf{x} \in \mathbb{R}^n$ satisfies $|\mathbf{x}| < \delta$ and if $\varphi(x) \in W$ then $\mathbf{x} = \mu(\varphi(\mathbf{x}))$. Now $\mathbf{x} = \psi(\mathbf{x}) + T\varphi(\mathbf{x})$ for all $\mathbf{x} \in V$. Also

$$|T\varphi(\mathbf{x})| \leq \|T\|_{\mathrm{op}} \, |\varphi(\mathbf{x})|$$

for all $\mathbf{x} \in V$, where

$$\|T\|_{\mathrm{op}} = \sup\{|T\mathbf{v}| : \mathbf{v} \in \mathbb{R}^n \text{ and } |\mathbf{v}| = 1\}.$$

It follows that

$$\begin{aligned} |\mathbf{x} - \mathbf{x}'| &= |\psi(\mathbf{x}) - \psi(\mathbf{x}') + T(\varphi(\mathbf{x}) - \varphi(\mathbf{x}'))| \\ &\leq |\psi(\mathbf{x}) - \psi(\mathbf{x}')| + |T^{-1}(\varphi(\mathbf{x}) - \varphi(\mathbf{x}'))| \\ &\leq \tfrac{1}{2}|\mathbf{x} - \mathbf{x}'| + \|T\|_{\mathrm{op}} \, |\varphi(\mathbf{x}) - \varphi(\mathbf{x}')| \end{aligned}$$

whenever $|\mathbf{x}| < \delta$ and $|\mathbf{x}'| < \delta$. Subtracting $\frac{1}{2}|\mathbf{x} - \mathbf{x}'|$ from both sides of the above inequality, and then multiplying by two, we find that

$$|\mathbf{x} - \mathbf{x}'| \leq 2\|T\|_{\mathrm{op}} \, |\varphi(\mathbf{x}) - \varphi(\mathbf{x}')|.$$

whenever $|\mathbf{x}| < \delta$ and $|\mathbf{x}'| < \delta$. Substituting $\mathbf{x}' = \mu(\mathbf{y})$, we find that

$$|\mathbf{x} - \mu(\mathbf{y})| \leq 2\|T\|_{\mathrm{op}} \, |\varphi(\mathbf{x}) - \mathbf{y}|$$

for all $\mathbf{x} \in V$ satisfying $|\mathbf{x}| < \delta$ and for all $\mathbf{y} \in W$. It follows that if $\mathbf{x} \in V$ satisfies $|\mathbf{x}| < \delta$ and if $\varphi(\mathbf{x}) = \mathbf{y}$ for some $\mathbf{y} \in W$ then $\mathbf{x} = \mu(\mathbf{y})$. The inequality also ensures that

$$|\mu(\mathbf{y}) - \mu(\mathbf{y}')| \leq 2\|T\|_{\mathrm{op}} \, |\mathbf{y} - \mathbf{y}'|$$

for all $\mathbf{y}, \mathbf{y}' \in W$. Thus the function $\mu: W \to V$ is Lipschitz continuous. It then follows from Lemma 8.10 that the function μ is continuously differentiable.

Next we prove that $\mu(W)$ is an open subset of V. Now $\mu(W) \subset \varphi^{-1}(W)$ because $\mathbf{y} = \varphi(\mu(\mathbf{y}))$ for all $\mathbf{y} \in W$. We have also proved that $|\mu(\mathbf{y})| < \delta$ for all $\mathbf{y} \in W$. It follows that

$$\mu(W) \subset \varphi^{-1}(W) \cap \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < \delta\}.$$

But we have also shown that if $\mathbf{x} \in V$ satisfies $|\mathbf{x}| < \delta$, and if $\varphi(\mathbf{x}) \in W$ then $\mathbf{x} = \mu(\varphi(\mathbf{x}))$, and therefore $\mathbf{x} \in \mu(W)$. It follows that

$$\mu(W) = \varphi^{-1}(W) \cap \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < \delta\}.$$

Now $\varphi^{-1}(W)$ is an open subset in V, because $\varphi: V \to \mathbb{R}^n$ is continuous and W is an open set in $\mathbb{R}^n$ (see Proposition 6.19). It follows that $\mu(W)$ is an intersection of two open sets, and is thus itself an open set. Moreover $\mathbf{0} \in \mu(W)$, because $\mu(\mathbf{0}) = \mathbf{0}$. We have now completed the proof that $\mu: W \to V$ satisfies properties (i) and (ii) in the statement of the theorem, and is thus a continuously differentiable local inverse for the map $\varphi: V \to \mathbb{R}^n$.

The result that this local inverse is C^k when φ is C^k then follows from Lemma 8.11. This completes the proof of the Inverse Function Theorem. ∎

Corollary 8.14 *Let $\varphi: V \to \mathbb{R}^n$ be a smooth function defined over an open set V in n-dimensional Euclidean space $\mathbb{R}^n$ and mapping V into $\mathbb{R}^n$. Then φ has a smooth local inverse around any point $\mathbf{p}$ at which the derivative $(D\varphi)_\mathbf{p}$ is invertible.*

Proof This result follows directly from the Inverse Function Theorem (Theorem 8.13), in view of the fact that a function $\varphi: V \to \mathbb{R}^n$ is smooth if and only if it is C^k for all positive integers k.

Definition Let V and W be open sets in n-dimensional Euclidean space $\mathbb{R}^n$, and let $\varphi: V \to W$ be a function from V to W. The function φ is said to be a *diffeomorphism* if it has a well-defined inverse $\varphi^{-1}: W \to V$ and both the function $\varphi: V \to W$ and its inverse $\varphi^{-1}: W \to V$ are smooth functions.

Definition Let V be an open set in n-dimensional Euclidean space $\mathbb{R}^n$, and let $\varphi: V \to \mathbb{R}^n$ be a smooth function from V to $\mathbb{R}^n$. Let U be an open subset of V. We say that φ maps U *diffeomorphically* onto an open set of $\mathbb{R}^n$ if $\varphi(U)$ is an open set in $\mathbb{R}^n$ and the restriction of the function φ to U is a diffeomorphism from U to $\varphi(U)$.

Corollary 8.15 *Let V be an open set in n-dimensional Euclidean space $\mathbb{R}^n$, and let $\varphi: V \to \mathbb{R}^n$ be a smooth function from V to $\mathbb{R}^n$, and let $\mathbf{p} \in V$. Suppose that the derivative $(D\varphi)_{\mathbf{p}}$ of φ is invertible at the point $\mathbf{p}$. Then there exists an open subset U of V, where $\mathbf{p} \in U$, that is mapped diffeomorphically by φ onto an open set in $\mathbb{R}^n$.*

Proof The derivative $(D\varphi)_{\mathbf{p}}$ of φ is invertible at the point $\mathbf{p}$. It follows from the Inverse Function Theorem (Theorem 8.13) that there exists an open set W in $\mathbb{R}^n$ and a smooth map $\mu: W \to V$ such that $\mu(W)$ is an open subset of V, $\mathbf{p} \in \mu(W)$ and $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$. Let $U = \mu(W)$. Then $\varphi(U) = W$, because $\varphi(\mu(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in W$. Moreover if $\mathbf{x} \in U$ then $\mathbf{x} = \mu(\mathbf{y})$ for some point $\mathbf{y}$ of W. But then

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

Thus $\mu(\varphi(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in U$. It follows that φ maps the open set U diffeomorphically onto W, and the inverse of this diffeomorphism from U to W is the smooth map μ. The result follows. ∎

8.5 Smooth Curvilinear Coordinate Systems

Definition Let U be an open set in n-dimensional Euclidean space $\mathbb{R}^n$, and let $u_1, u_2, \ldots, u_n$ be smooth real-valued functions defined throughout U, and let

$$\tilde{U} = \{(u_1(\mathbf{x}), u_2(\mathbf{x}), \ldots, u_n(\mathbf{x})) : \mathbf{x} \in U\}.$$

Then the smooth real-valued functions $u_1, u_2, \ldots, u_n$ are said to constitute a *smooth curvilinear coordinate system* on U if $\tilde{U}$ is an open set in $\mathbb{R}^n$ and there exist smooth real-valued functions $\xi_1, \xi_2, \ldots, \xi_n$ defined over $\tilde{U}$ such that

$$x_i = \xi_i(u_1(x_1, x_2, \ldots, x_n), u_2(x_1, x_2, \ldots, x_n), \ldots, u_n(x_1, x_2, \ldots, x_n))$$

for all $\mathbf{x} \in U$.

Let U be an open set in n-dimensional Euclidean space $\mathbb{R}^n$, and let $u_1, u_2, \ldots, u_n$ be smooth real-valued functions defined throughout U, and let

$$\tilde{U} = \{(u_1(\mathbf{x}), u_2(\mathbf{x}), \ldots, u_n(\mathbf{x})) : \mathbf{x} \in U\}.$$

Let $\varphi: U \to \tilde{U}$ be defined so that

$$\varphi(\mathbf{x}) = (u_1(\mathbf{x}), u_2(\mathbf{x}), \ldots, u_n(\mathbf{x})).$$

Then the smooth real-valued functions $u_1, u_2, \ldots, u_n$ constitute a smooth curvilinear coordinate system on U if and only if $\varphi: U \to \tilde{U}$ is a diffeomorphism.

Suppose that $u_1, u_2, \ldots, u_n$ constitute a smooth curvilinear coordinate system on the open set U. Let the open set $\tilde{U}$ and the diffeomorphism $\varphi U \to \tilde{U}$ be defined as described above. A differentiable function $f: U$ determines a corresponding differentiable function $f \circ \varphi^{-1}: \tilde{U} \to U$. The partial derivatives of the function f with respect to the curvilinear coordinates $u_1, u_2, \ldots, u_n$ are then defined so that

$$\begin{aligned} \left.\frac{\partial f}{\partial u_j}\right|_{\mathbf{p}} &= (\partial_j (f \circ \varphi^{-1}))(\varphi(\mathbf{p})) \\ &= \left.\frac{\partial f(\xi_1(y_1, \ldots, y_n), \ldots, \xi_n(y_1, \ldots, y_n))}{\partial y_j}\right|_{(y_1, \ldots, y_n) = \varphi(\mathbf{p})} \end{aligned}$$

for all $\mathbf{p} \in U$, where $\partial_j (f \circ \varphi^{-1})$ denotes the partial derivative of $f \circ \varphi^{-1}$ with respect to the jth Cartesian coordinate on the open set $\tilde{U}$. The Chain Rule (Theorem 7.9) ensures that

$$(Df)_{\mathbf{p}} = D(f \circ \varphi^{-1})_{\varphi(\mathbf{p})} (D\varphi)_{\mathbf{p}}.$$

It follows that

$$\begin{aligned} \left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{p}} &= (\partial_i f)(\mathbf{p}) = \sum_{j=1}^{n} (\partial_j (f \circ \varphi^{-1}))(\varphi(\mathbf{p}))\, (\partial_i u_j)(\mathbf{p}) \\ &= \sum_{j=1}^{n} (\partial_j (f \circ \varphi^{-1}))(\varphi(\mathbf{p}))\, (\partial_i u_j)(\mathbf{p}) \\ &= \sum_{j=1}^{n} \left.\frac{\partial f}{\partial u_j}\right|_{\mathbf{p}} \left.\frac{\partial u_j}{\partial x_i}\right|_{\mathbf{p}}. \end{aligned}$$

This establishes the Chain Rule

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial u_j} \frac{\partial u_j}{\partial x_i}$$

used to compute the partial derivatives of a smooth real-valued function f on the domain U of a smooth curvilinear coordinate system $u_1, u_2, \ldots, u_n$.

We can apply the Chain Rule when the functions to be differentiated are the Cartesian coordinate functions on U itself. We find that

$$\sum_{j=1}^{n} \frac{\partial x_k}{\partial u_j} \frac{\partial u_j}{\partial x_i} = \begin{cases} 1 & \text{if } i = k; \\ 0 & \text{if } i \neq k. \end{cases}$$

It follows that the Jacobian matrices associated with the change of coordinates satisfy

$$\begin{pmatrix} \dfrac{\partial x_1}{\partial u_1} & \dfrac{\partial x_1}{\partial u_2} & \cdots & \dfrac{\partial x_1}{\partial u_n} \\ \dfrac{\partial x_2}{\partial u_1} & \dfrac{\partial x_2}{\partial u_2} & \cdots & \dfrac{\partial x_2}{\partial u_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial x_n}{\partial u_1} & \dfrac{\partial x_n}{\partial u_2} & \cdots & \dfrac{\partial x_n}{\partial u_n} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{pmatrix}^{-1}.$$

Let $v_1, v_2, ldots, v_n$ be another smooth local coordinate system defined over an open set V, where $U \cap V$ is non-empty. Then

$$\sum_{j=1}^{n} \frac{\partial f}{\partial u_j}\frac{\partial u_j}{\partial x_i} = \frac{\partial f}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial f}{\partial v_k}\frac{\partial v_k}{\partial x_i} = \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial f}{\partial v_k}\frac{\partial v_k}{\partial u_j}\frac{\partial u_j}{\partial x_i}$$

throughout $U \cap V$ for $i = 1, 2, \ldots, n$. It then follows from the invertibility of the Jacobian matrix of partial derivatives of $u_1, u_2, \ldots, u_n$ with respect to $x_1, x_2, \ldots, x_n$ that

$$\frac{\partial f}{\partial u_j} = \sum_{j=1}^{n} \frac{\partial f}{\partial v_k}\frac{\partial v_k}{\partial u_j}$$

throughout $U \cap V$ for $j = 1, 2, \ldots, n$.

The fact that compositions of smooth functions are smooth ensures that the smooth curvilinear coordinates $v_1, v_2, \ldots, v_n$ can be expressed as smooth functions of $u_1, u_2, \ldots, u_n$ and vice versa throughout the open set $U \cap V$ where the domains of the smooth curvilinear coordinate systems overlap.

Proposition 8.16 *Let $u_1, u_2, \ldots, u_n$ be smooth real-valued functions defined throughout some open neighbourhood of a point $\mathbf{p}$ of n-dimensional Euclidean space $\mathbb{R}^n$. Suppose that the Jacobian matrix*

$$\begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{pmatrix}$$

of partial derivatives of $u_1, u_2, \ldots, u_n$ with respect to $x_1, x_2, \ldots, x_n$ is invertible at the point $\mathbf{p}$. Then there exists an open set U containing the point $\mathbf{p}$

such that the restrictions of the functions $u_1, u_2, \ldots, u_n$ to the open set U constitute a smooth curvilinear coordinate system over the open set U.

Proof This result is essentially a restatement of the Inverse Function Theorem (Theorem 8.13), and follows directly from Corollary 8.15 and the definition of smooth curvilinear coordinate systems. ∎

8.6 The Implicit Function Theorem

Theorem 8.17 *Let $\mathbf{p}$ be a point of $\mathbb{R}^n$, where $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ and let $u_1, u_2, \ldots, u_m$ be a smooth real-valued functions defined over an open neighbourhood V of the point $\mathbf{p}$ in $\mathbb{R}^n$, where $m < n$, and let*

$$M = \{\mathbf{x} \in V : u_j(\mathbf{x}) = 0 \text{ for } j = 1, 2, \ldots, m\}.$$

Suppose that $u_1, u_2, \ldots, u_n$ are zero at $\mathbf{p}$ and that the matrix

$$\begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_m} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial u_m}{\partial x_1} & \dfrac{\partial u_m}{\partial x_2} & \cdots & \dfrac{\partial u_m}{\partial x_m} \end{pmatrix}$$

is invertible at the point $\mathbf{p}$. Then there exists an open neighbourhood U of $\mathbf{p}$ and a smooth functions $f_1, f_2, \ldots, f_m$ of $n - m$ real variables, defined around $(p_{m+1}, \ldots, p_n)$ in $\mathbb{R}^{n-m}$, such that

$$M \cap U = \{(x_1, x_2, \ldots, x_n) \in U : x_j = f_j(x_{m+1}, \ldots, x_n) \text{ for } j = 1, 2, \ldots, m\}.$$

Proof Let $u_j = x_j$ for $j = m + 1, \ldots, n$, and let

$$J_0 = \begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_m} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial u_m}{\partial x_1} & \dfrac{\partial u_m}{\partial x_2} & \cdots & \dfrac{\partial u_m}{\partial x_m} \end{pmatrix}, \quad J = \begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{pmatrix}.$$

(The matrix J_0 is thus the leading $m \times m$ minor of the $n \times n$ matrix J.) Now

$$J = \begin{pmatrix} \dfrac{\partial u_1}{\partial x_1} & \cdots & \dfrac{\partial u_1}{\partial x_m} & \dfrac{\partial u_1}{\partial x_{m+1}} & \dfrac{\partial u_1}{\partial x_{m+2}} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \\ \dfrac{\partial u_m}{\partial x_1} & \cdots & \dfrac{\partial u_m}{\partial x_m} & \dfrac{\partial u_m}{\partial x_{m+1}} & \dfrac{\partial u_m}{\partial x_{m+2}} & \cdots & \dfrac{\partial u_m}{\partial x_n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

It follows from basic properties of determinants that $\det J = \det J_0$, and therefore $\det J$ is non-zero at the point $\mathbf{p}$. It follows that matrix J whose coefficients are the first order partial derivatives of $u_1, u_2, \ldots, u_n$ with respect to $x_1, x_2, \ldots, x_n$ is invertible at the point p. It then follows from Proposition 8.16 that $u_1, u_2, \ldots, u_n$ is a smooth curvilinear coordinate system defined over some open set U that contains the point $\mathbf{p}$ and satisfies $u_j = x_j$ for $j > m$. It then follows that there exist smooth real-valued functions ξ_1, ξ_2, ξ_n such that

$$x_j = \xi_j(u_1, u_2, \ldots, u_m, x_{m+1}, \ldots, x_n)$$

for $j = 1, 2, \ldots, n$. Let

$$f_j(x_{m+1}, \ldots, x_n) = \xi_j(0, 0, \ldots, 0, x_{m+1}, \ldots, x_n)$$

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

$$M \cap U = \{(x_1, x_2, \ldots, x_n) \in U : x_j = f_j(x_{m+1}, \ldots, x_n) \text{ for } j = 1, 2, \ldots, m\},$$

as required. ∎

Corollary 8.18 *Let $u\colon V \to \mathbb{R}$ be a smooth real-valued function defined over an open subset V of $\mathbb{R}^n$. Suppose that $\dfrac{\partial u}{\partial x_n} \neq 0$ at some point $\mathbf{p}$ of V, where $p = (p_1, p_2, \ldots, p_n)$. Then there exist an open neighbourhood U of $\mathbf{p}$ and a smooth real-valued function f, defined throughout some open neighbourhood of $(p_1, p_2, \ldots, p_{n-1})$ in $\mathbb{R}^{n-1}$, such that*

$$\{\mathbf{x} \in U : u(\mathbf{x}) = 0\} = \{(x_1, x_2, \ldots, x_n) \in U : x_n = f(x_1, x_2, \ldots, x_{n-1})\}.$$

Proof This result comes directly on applying the Implicit Function Theorem (Theorem 8.17), after reordering Cartesian coordinates so that x_n precedes $x_1, x_2, \ldots, x_{n-1}$. ∎

8.7 Submanifolds of Euclidean Spaces

A function is said to be *injective* (or *one-to-one*) if distinct points of the domain get mapped to distinct points of the codomain.

Let M be a subset of n-dimensional Euclidean space $\mathbb{R}^n$, and let Let $\alpha: U \to M$ be a smooth function mapping some open subset U of a Euclidean space $\mathbb{R}^k$ into M, where $0 < k < n$. The function α is injective if and only if $\alpha(\mathbf{u}) \neq \alpha(\mathbf{u}')$ for all $\mathbf{u}, \mathbf{u}' \in U$ satisfying $\mathbf{u} \neq \mathbf{u}'$. If $\alpha: U \to M$ is injective, then there is a well-defined function $\rho: \alpha(U) \to U$ defined such that $\rho(\alpha(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in U$.

The range $\alpha(U)$ of the map α is open in M if and only if, given any point $\mathbf{p}$ of $\alpha(U)$, there exists some $\delta > 0$ such that all points of M that lie within a distance δ of the point $\mathbf{p}$ belong to $\alpha(U)$.

The derivative $(D\alpha)_{\mathbf{u}}$ of α at each point $\mathbf{u}$ has a *rank* which is by definition the dimension of the image $(D\alpha)_{\mathbf{u}}(\mathbb{R}^k)$ of the linear transformation $(D\alpha)_{\mathbf{u}}: \mathbb{R}^k \to \mathbb{R}^n$. The rank of $(D\alpha)_{\mathbf{u}}$ is some integer between 0 and k. We consider the local properties of the image of a smooth injective function $\alpha: U \to \mathbb{R}^n$ defined over an open subset U of $\mathbb{R}^k$ in the case where the rank of the derivative of α at each point of U has its maximum possible value, which is k.

Proposition 8.19 *Let k and n be positive integers satisfying $k < n$, let let U be an open set in $\mathbb{R}^k$, let $\alpha: U \to \mathbb{R}^n$ be a smooth injective function from U into $\mathbb{R}^n$. Suppose that the following conditions are satisfied:—*

(i) the function $\alpha: U \to \mathbb{R}^n$ is injective;

(ii) the inverse of α on the set $\alpha(U)$ is a continuous map from $\alpha(U)$ to U;

(iii) the derivative $(D\alpha)_{\mathbf{u}}$ of α at each point $\mathbf{u}$ of U has rank k.

Then, given any point $\mathbf{p}$ of $\alpha(U)$, there exists an open set W in $\mathbb{R}^n$, where $\mathbf{p} \in W$, and a smooth curvilinear coordinate system $w_1, w_2, \ldots, w_n$ defined over W such that

$$\alpha(U) \cap W = \{\mathbf{x} \in W : w_j(\mathbf{x}) = 0 \text{ for } k < j \leq n\}$$

and $w_j(\alpha(u_1, u_2, \ldots, u_k)) = u_j$ for all $(u_1, u_2, \ldots, u_k) \in U$.

Proof Let $\mathbf{p}$ be a point of $\alpha(U)$. Then there exists $\mathbf{u} \in U$ such that $\mathbf{p} = \alpha(\mathbf{u})$. We may assume, without loss of generality, that $\mathbf{p} = \alpha(\mathbf{0})$. Then $(D\alpha)_{\mathbf{0}}(\mathbb{R}^k)$ is a vector subspace of $\mathbb{R}^n$ of dimension k. Let

$$\mathbf{v}_j = \left.\frac{\partial\alpha(u_1, u_2, \ldots, u_k)}{\partial u_j}\right|_{(u_1, u_2, \ldots, u_k) = \mathbf{0}}$$

for $j = 1, 2, \ldots, k$. Then the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent vectors in $(D\alpha)_{\mathbf{p}}(\mathbb{R}^k)$ that span this vector space. It follows from standard linear algebra that there exist vectors $\mathbf{v}_j$ in $\mathbb{R}^n$ for $j = k+1, \ldots, n$ such that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ consistute a basis of the vector space $\mathbb{R}^n$. The smooth function $\alpha\colon U \to \mathbb{R}^n$ then extends to a smooth function $\beta\colon V \to \mathbb{R}^n$, where

$$V = \{(u_1, u_2, \ldots, u_k, 0, \ldots, 0) \in \mathbb{R}^n : (u_1, u_2, \ldots, u_k) \in U\}$$

and

$$\beta(u_1, u_2, \ldots, u_n) = \alpha(u_1, u_2, \ldots, u_k) + \sum_{j=k+1}^{n} u_j \mathbf{v}_j.$$

Let $\lambda\colon U \to V$ be defined so that

$$\lambda(u_1, u_2, \ldots, u_k) = (u_1, u_2, \ldots, u_k, 0, \ldots, 0)$$

for all $(u_1, u_2, \ldots, u_k) \in U$. Then $\alpha(\mathbf{u}) = \beta(\lambda(\mathbf{u}))$ for all $\mathbf{u} \in U$.

Let $\mathbf{e}_j$ be the unit vector in $\mathbb{R}^n$ whose jth component has the value 1 and whose other components are zero for $j = 1, 2, \ldots, n$. Then $(D\beta)_{\mathbf{0}}\, \mathbf{e}_j = \mathbf{v}_j$ for $j = 1, 2, \ldots, n$. It follows that $(D\beta)_{\mathbf{0}}$ is an invertible linear transformation whose inverse sends $\mathbf{v}_j$ to $\mathbf{e}_j$ for $j = 1, 2, \ldots, n$. It then follows from the Inverse Function Theorem (Theorem 8.13) that there exists a smooth local inverse $\mu\colon W_0 \to V$ for the map α around the point $\mathbf{p}$ defined over some open set W_0 in $\mathbb{R}^n$. Then $\mathbf{p} \in W_0$, $\mu\colon W_0 \to V$ is a smooth function from W_0 to V, $\mu(W_0)$ is an open subset of V and $\beta(\mu(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in W_0$.

Now $\lambda^{-1}(\mu(W_0))$ is an open set in $\mathbb{R}^k$ and $\mathbf{0} \in \lambda^{-1}(\mu(W_0))$. It follows that there exists some positive number η such that $\mathbf{u} \in \lambda^{-1}(\mu(W_0))$ for all $\mathbf{u} \in \mathbb{R}^k$ satisfying $|\mathbf{u}| < \eta$. The continuity of the inverse of α on $\alpha(U)$ then ensures the existence of a positive real number δ such that $|\mathbf{u}| < \eta$ for all $\mathbf{u} \in U$ satisfying $|\alpha(\mathbf{u}) - \mathbf{p}| < \delta$. Let

$$W = \{\mathbf{x} \in W_0 : |\mathbf{x} - \mathbf{p}| < \delta\}.$$

If $\mathbf{u} \in U$ and if $\alpha(\mathbf{u}) \in W$ then $|\alpha(\mathbf{u}) - \mathbf{p}| < \delta$, and therefore $|\mathbf{u}| < \eta$. But then $\mathbf{u} \in \lambda^{-1}(\mu(W_0))$, and therefore $\lambda(\mathbf{u}) = \mu(\mathbf{x})$ for some $\mathbf{x} \in W_0$. But then

$$\alpha(\mathbf{u}) = \beta(\lambda(\mathbf{u})) = \beta(\mu(\mathbf{x})) = \mathbf{x}.$$

It follows that

$$\mu(\alpha(\mathbf{u})) = \mu(\mathbf{x}) = \lambda(\mathbf{u}).$$

We have thus shown that $\mu(\alpha(\mathbf{u})) = \lambda(\mathbf{u})$ for all $\mathbf{u} \in U$ for which $\alpha(\mathbf{u}) \in W$.

Let the smooth real-valued functions $w_1, w_2, \ldots, w_n$ be defined throughout the open subset W of $\mathbb{R}^n$ so that

$$\mu(\mathbf{x}) = (w_1(\mathbf{x}), w_2(\mathbf{x}), \ldots, w_n(\mathbf{x}))$$

for all $\mathbf{x} \in W$. Then $w_1, w_2, \ldots, w_n$ is a smooth curvilinear coordinate system defined over the open set W. Let $\mathbf{u} \in U$. Suppose that $\alpha(\mathbf{u}) \in W$. Then $\mu(\alpha(\mathbf{u})) = \lambda(\mathbf{u})$. It follows that $w_j(\alpha(\mathbf{u})) = u_j$ for $j = 1, 2, \ldots, k$, and $w_j(\alpha(\mathbf{u})) = 0$ when $j > k$. It follows from this that

$$\alpha(U) \cap W = \{\mathbf{x} \in W : w_j(\mathbf{x}) = 0 \text{ for } k < j \leq n\},$$

as required. ∎

The proof of Proposition 8.19 involves some technicalities that depend on the requirement that the inverse of the map $\alpha: U \to \mathbb{R}^n$ in the statement of the proposition be continuous on $\alpha(U)$. The following example demonstrates that the conclusions of the proposition may fail to hold in situations where the other requirements of the proposition are satisfied by the continuity requirement (ii) in the statements of the proposition is not satisfied.

Example Let ν be an irrational number and let $\gamma: \mathbb{R} \to \mathbb{R}^3$ be the smooth curve in $\mathbb{R}^3$ defined such that

$$\gamma(t) = ((2 + \cos 2\pi t)\, \cos 2\pi\nu t, (2 + \cos 2\pi t)\, \sin 2\pi\nu t \sin 2\pi t)$$

for all $t \in \mathbb{R}$. Then γ is a smooth curve which winds around the torus

$$\{(x, y, z) \in \mathbb{R}^3 : (\sqrt{x^2 + y^2} - 2)^2 + z^2 = 1\}.$$

Morover the velocity vector $\dfrac{d\gamma(t)}{dt}$ is everywhere non-zero. The map γ is injective. Indeed suppose that t_1 and t_2 are real numbers satisfying $\gamma(t_1) = \gamma(t_2)$. Then both $t_1 - t_2$ and $\nu(t_1 - t_2)$ are integers, and the fact that ν is irrational ensures that this can only happen when $t_1 = t_2$.

Now if p, p', q, q' are integers and if and if $p - \nu q = p' - \nu q'$ then $p = p'$ and $q = q'$. We use this fact to construct infinite sequences $p_1, p_2, p_3, \ldots$ and $q_1, q_2, q_3, \ldots$ of integers such that $p_n - q_n\nu > 0$ and

$$0 < p_{n+1} - \nu q_{n+1} < \tfrac{1}{2}(p_n - \nu q_n)$$

for all positive integers n. Choose integers p_1 and q_1 for which $0 < p_1 - \nu q_1 < 1$. Then suppose that integers $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ have been determined so as to satisfy the required inequalities. Then $0 < p_n - \nu q_n < 1$. We show

how to determine integers p_{n+1} and q_{n+1} for which $0 < p_{n+1} - \nu q_{n+1} < \frac{1}{2}(p_n - \nu q_n)$.

Let $p = p_n + 1$ and $q = q_n$. Then $p - \nu q > p_n - \nu q_n$ and $p - \nu q \neq k(p_n - \nu q_n)$ for all integers k. Let k_0 be the largest integer for which $k_0(p_n - \nu q_n) < p - \nu q$, and let $p' = p - k_0 p_n$ and $q' = q - k_0 q_n$. Then $p' - \nu q' < p_n - \nu q_n$. Then let $p_{n+1} = (1 - k_1)p_n - k_1 p'$ and $q_{n+1} = (1 - k_1)q_n - k_1 q'$, where k_1 be the largest positive integer for which $(p_n - \nu q_n) + k_1(p' - p_n - \nu(q' - q_n)) > 0$. Then $0 < p_{n+1} - \nu q_{n+1} < \frac{1}{2}(p_n - \nu q_n)$. The infinite sequences of integers constructed in this fashion have the property that

$$0 < p_n - q_n \omega < \frac{1}{2^{n-1}}(p_1 - q_1 \omega)$$

for all positive integers n. It follows that if $u_n = p_n - q_n \omega$ then the real numbers $u_1, u_2, u_3, \ldots$ constitute a decreasing sequence of real numbers converging to zero. Moreover the real numbers u_j are all distinct, and each u_j is uniquely determined by the value of q_j. It follows that the integers $q_1, q_2, q_3, \ldots$ are distinct.

Now $\cos 2\pi q_n = 1$, $\sin 2\pi q_n = 0$ for all positive integers n. Also

$$\cos 2\pi \nu q_n = \cos 2\pi(p_n + u_n) = \cos 2\pi u_n$$

and similarly $\sin 2\pi \nu q_n = \sin 2\pi u_n$ for all positive integers n. It follows that

$$\gamma(q_n) = (3\cos 2\pi u_n, 3\sin 2\pi u_n, 0)$$

for all positive integers n, and therefore $\gamma(q_n) \to (3, 0, 0)$ as $n \to +\infty$. But the infinite sequence $q_1, q_2, q_3, \ldots$ of distinct integers is not convergent. It follows that the inverse of the function γ is not continuous on $\gamma(\mathbb{R})$. Also, given any open neighbourhood of $(3, 0, 0)$, no matter how small, the curve γ passes infinitely often through that open neighbourhood. It is not therefore possible to find a smooth curvilinear coordinate system around $(3, 0, 0)$ satisfying the requirements in the statement of Proposition 8.19.

Definition Let M be a subset of n-dimensional Euclidean space $\mathbb{R}^n$. Then M is said to be a *smooth submanifold* of $\mathbb{R}^n$ of dimension k if and only if, given any point $\mathbf{p}$ of M, there exists an open set W, where $\mathbf{p} \in W$, and a smooth curvilinear coordinate system $w_1, w_2, \ldots, w_n$ defined over the open set W such that

$$M \cap W = \{\mathbf{x} \in W : w_j(\mathbf{x}) = 0 \text{ for } k < j \leq n\}$$

and $w_j(\alpha(u_1, u_2, \ldots, u_k)) = u_j$ for all $(u_1, u_2, \ldots, u_k) \in U$.

Let M be a k-dimensional smooth submanifold of $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of M. Then there exists s smooth curvilinear coordinate system $w_1, w_2, \ldots, w_n$ defined over an open set W in $\mathbb{R}^n$ such that

$$M \cap W = \{\mathbf{x} \in W : w_j(\mathbf{x}) = 0 \text{ for } k < j \leq n\}.$$

Let U be the open set in $\mathbb{R}^k$ defined so that

$$U = \{(w_1(\mathbf{x}), w_2(\mathbf{x}), \ldots, w_k(\mathbf{x})) : \mathbf{x} \in M \cap W\},$$

and let $\alpha\colon U \to \mathbb{R}^n$ be the smooth map from U to W defined such that

$$w_j(\alpha(u_1, u_2, \ldots, u_k) = u_j \quad \text{for} \quad j = 1, 2, \ldots, k$$

and

$$w_j(\alpha(u_1, u_2, \ldots, u_k)) = u_j \quad \text{for} \quad j > k.$$

Then $\alpha(U) = M \cap W$, and therefore $\alpha(U)$ is open in M. Also the smooth map $\alpha\colon U \to \mathbb{R}^n$ is injective, and the components of its inverse on $\alpha(U)$ are the restrictions of the smooth real-valued functions $w_1, w_2, \ldots, w_k$ to $\alpha(U)$. It follows that the inverse of α is continuous on $\alpha(U)$. Finally the derivative $(D\alpha)_{\mathbf{q}}$ of α at any point $\mathbf{q}$ of $\mathbf{U}$ is represented by a matrix product JA where the components of the matrices $n \times n$ matrix J and the $n \times k$ matrix A satisfy

$$J_{i,j} = \left.\frac{\partial x_i}{\partial w_j}\right|_{\alpha(\mathbf{q})} \qquad \text{for } i, j = 1, 2, \ldots, n$$

and

$$A_{j,l} = \left.\frac{\partial w_j(\alpha(u_1, \ldots, u_k))}{\partial u_l}\right|_{\mathbf{q}} \qquad \text{for } j = 1, 2, \ldots, n \text{ and } l = 1, 2, \ldots, k.$$

Now $A_{i,m} = 1$ when $m = i$, and $A_{i,m} = 0$ when $m \neq i$. It follows that the matrix A has rank k. Also the matrix J is invertible. It follows that the derivative $(D\alpha)_{\mathbf{q}}$ of the function α has rank k at each point of U.

Corollary 8.20 *Let M be a subset of n-dimensional Euclidean space $\mathbb{R}^n$. Then M is a k-dimensional smooth submanifold of $\mathbb{R}^n$ if and only if, given any point $\mathbf{p}$ of M, there exists a smooth map $\alpha\colon U \to M$, defined over some open set in $\mathbb{R}^k$, which satisfies the following conditions:*

(i) $\alpha(U)$ is open in M and $\mathbf{p} \in \alpha(U)$;

(ii) the function $\alpha\colon U \to M$ is injective;

(iii) the inverse of α on the set $\alpha(U)$ is a continuous map from $\alpha(U)$ to U;

(iv) the derivative $(D\alpha)_{\mathbf{u}}$ of α at each point $\mathbf{u}$ of U has rank k.

Proof The preceding remarks show that if M is a smooth k-dimensional submanifold of $\mathbb{R}^k$, so that, given any point $\mathbf{p}$ of M, there exists an open set W in $\mathbb{R}^n$, where $\mathbf{p} \in W$, and a smooth curvilinear coordinate system $w_1, w_2, \ldots, w_n$ on W such that

$$M \cap W = \{\mathbf{x} \in W : w_j(\mathbf{x}) = 0 \text{ for } k < j \leq n\},$$

if

$$U = \{(w_1(\mathbf{x}), w_2(\mathbf{x}), \ldots, w_k(\mathbf{x})) : \mathbf{x} \in M \cap W\},$$

and if $\alpha\colon U \to \mathbb{R}^n$ is the smooth map from U to W defined such that

$$w_j(\alpha(u_1, u_2, \ldots, u_k) = u_j \quad \text{for} \quad j = 1, 2, \ldots, k,$$

then conditions (i), (ii), (iii) and (iv) of the corollary are satisfied by the map α.

Conversely if, given any point $\mathbf{p}$ there exist a smooth map α satisfying conditions (i), (ii), (iii) and (iv) of the corollary, then Proposition 8.19 ensures that M is a smooth submanifold of $\mathbb{R}^n$, as required. ∎