Module MAU23203: Analysis in Several Real Variables
Michaelmas Term 2022
Section 8: Differentiation of Functions of Several Real Variables

D. R. Wilkins

Contents

8 Differentiation of Functions of Several Real Variables

8.1 Functions with First Order Partial Derivatives

If a real-valued function of a single real variable is differentiable, then it is guaranteed to be continuous. However, for a function of two or more real variables, the mere existence of first order partial derivatives throughout the domain of the function is not sufficient to ensure continuity.

Example Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined so that

$$f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0); \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

If $(x, y) \neq (0, 0)$ then the partial derivatives of f are well-defined at (x, y), and

$$\frac{\partial f}{\partial x} = \frac{2y(x^2 - y^2)}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y} = \frac{-2x(x^2 - y^2)}{(x^2 + y^2)^2}.$$

The partial derivatives of the function f at $(0, 0)$ are also well-defined, and are equal to zero, because the function f has the value zero along the lines $y = 0$ and $x = 0$. Thus the first order partial derivatives of the function f are well-defined throughout the domain $\mathbb{R}^2$ of the function.

Nevertheless $f(x, y) = 1$ at all points of the line $x = y$ with the exception of the origin $(0, 0)$, where the function takes the value zero. It follows from this that the function f is discontinuous at $(0, 0)$.

Example Let $g\colon \mathbb{R}^2 \to \mathbb{R}$ be defined so that

$$g(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y^2}\, e^{\frac{1}{x^2+y^2}} & \text{if } (x, y) \neq (0, 0); \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

This function g also have well-defined first order partial derivatives throughout $\mathbb{R}^2$. But $|g(x, y)|$ increases faster than any negative power of the distance from the origin as the point (x, y) approaches the origin along any straight line other than the lines $y = 0$ and $x = 0$.

Example Let $h\colon \mathbb{R}^2 \to \mathbb{R}$ be defined so that

$$h(x, y) = \begin{cases} \dfrac{2x^2 y}{x^4 + y^2} & \text{if } (x, y) \neq (0, 0); \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

The function h takes the value zero along the lines $y = 0$ and $x = 0$. It therefore has well-defined first order partial derivatives at the origin that have the value zero. It also has well-defined first order partial derivatives at all other points of $\mathbb{R}^2$.

Now if (u, v) is a point of $\mathbb{R}^2$, and if $v \neq 0$ and $t \neq 0$, then

$$h(tu, tv) = \frac{2tu^2v}{t^2u^4 + v^2}.$$

It follows that $\lim_{t \to 0} h(tu, tv) = 0$ whenever $v \neq 0$. This limit is also zero when $v = 0$ because the function takes the value zero along the line $y = 0$. Nevertheless $h(t, t^2) = 1$ for all non-zero real numbers t. The point (t, t^2) approaches the origin $(0, 0)$ as t tends to zero, and $h(0, 0) = 0$. It follows that the function h is not continuous at the origin.

8.2 Growth of Functions with Bounded Partial Derivatives

An open set X in $\mathbb{R}^m$ is a product of open intervals $J_1, J_2, \ldots, J_m$ if

$$\begin{aligned} X &= J_1 \times J_2 \times \cdots \times J_m \\ &= \{(x_1, x_2, \ldots, x_m) \in \mathbb{R}^m : x_i \in J_i \text{ for } i = 1, 2, \ldots, m\}. \end{aligned}$$

Suppose that $\mathbf{u}$ and $\mathbf{v}$ are points of X, where X is an open set in $\mathbb{R}^m$ that is a product of open intervals J_i for $i = 1, 2, \ldots, n$. Then there exist real numbers a_i and b_i in the open interval J_i for $i = 1, 2, \ldots, m$ such that $a_i < u_i < b_i$ and $a_i < v_i < b_i$ for $i = 1, 2, \ldots, m$. Let H be the closed subset of $\mathbb{R}^m$ consisting of those points $(x_1, x_2, \ldots, x_m)$ whose ith coordinate x_i satisfies $\min(u_i, v_i) \leq x_i \leq \max(u_i, v_i)$ for $i = 1, 2, \ldots, m$. Then $H \subset X$.

Lemma 8.1 *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be points of* $\mathbb{R}^m$*. Then*

$$\sum_{i=1}^{m} |u_i - v_i| \leq \sqrt{m}\, |\mathbf{u} - \mathbf{v}|.$$

Proof Consider the scalar product $(\mathbf{u}-\mathbf{v}) . \mathbf{s}$ of the m-dimensional vector $\mathbf{u}-\mathbf{v}$ with the vector $\mathbf{s}$ whose ith component s_i is determined for $i = 1, 2, \ldots, m$ so that $s_i = 1$ if $u_i \geq v_i$ and $s_i = -1$ if $u_i < v_i$. Now $(\mathbf{u} - \mathbf{v}) . \mathbf{s} = \sum_{i=1}^{m} |u_i - v_i|$ and $|\mathbf{s}| = \sqrt{m}$. Schwarz's Inequality (Lemma 2.1) ensures that $(\mathbf{u} - \mathbf{v}) . \mathbf{s} \leq |\mathbf{u} - \mathbf{v}|\, |\mathbf{s}|$. The required inequality follows immediately. ∎

Proposition 8.2 *Let X be an open set in $\mathbb{R}^m$ that is a product of open intervals, let $f\colon X \to \mathbb{R}$ be a real-valued function on X, and let M be a positive constant. Suppose that*

$$\left|\frac{\partial f}{\partial x_i}\right| \leq M$$

throughout the open set X for $i = 1, 2, \ldots, m$. Then

$$|f(\mathbf{u}) - f(\mathbf{v})| \leq \sqrt{m}\, M\, |\mathbf{u} - \mathbf{v}|$$

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

Proof Let $\mathbf{u} = (u_1, u_2, \ldots, u_m)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_m)$. Then real numbers a_i and b_i can be found such that $a_i < u_i < b_i$, $a_i < v_i < b_i$ and a_i and b_i both belong to or are endpoints of the interval J_i for $i = 1, 2, \ldots, m$. For each integer k between 0 and m, let

$$\mathbf{w}_k = (w_{k,1}, w_{k,2}, \ldots, w_{k,m})$$

where

$$w_{k,i} = \begin{cases} v_i & \text{if } i \leq k; \\ u_i & \text{if } i > k. \end{cases}$$

Then $a_i < w_{k,i} < b_i$ for $k = 0, 1, 2, \ldots, m$ and $i = 1, 2, \ldots, m$. Moreover, for each integer k between 1 and m, the points $\mathbf{w}_{k-1}$ and and $\mathbf{w}_k$ differ only in the kth coordinate, and the line segment joining these points is wholly contained in the open set X. It follows that

$$\frac{d}{dt}\left(f((1-t)\mathbf{w}_{k-1} + t\mathbf{w}_k)\right) = (v_k - u_k)\left.\frac{\partial f}{\partial x_k}\right|_{(1-t)\mathbf{w}_{k-1}+t\mathbf{w}_k}.$$

Consequently

$$\left|\frac{d}{dt}\left(f((1-t)\mathbf{w}_{k-1} + t\mathbf{w}_k)\right)\right| \leq M\,|u_k - v_k|,$$

and therefore

$$|f(\mathbf{w}_{k-1}) - f(\mathbf{w}_k)| \leq M\,|u_k - v_k|$$

for $i = 1, 2, \ldots, m$ (see Corollary 7.9). Thus (applying the inequality stated in Lemma 8.1,) we conclude that

$$\begin{aligned} |f(\mathbf{u}) - f(\mathbf{v})| &\leq \sum_{k=1}^{m} |f(\mathbf{w}_{k-1}) - f(\mathbf{w}_k)| \leq M \sum_{k=1}^{m} |u_k - v_k| \\ &\leq \sqrt{m}\, M\, |\mathbf{u} - \mathbf{v}|, \end{aligned}$$

as required. ∎

Example Let $f: \mathbb{R}^m \to \mathbb{R}$ be the function defined so that $f(x_1, x_2, \ldots, x_m) = \sum_{i=1}^{m} x_i$ for all $(x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$. Now $\dfrac{\partial f}{\partial x_i} = 1$ throughout $\mathbb{R}^m$ for $i = 1, 2, \ldots, m$. Let $\mathbf{u} = (0, 0, \ldots, 0)$ and $\mathbf{v} = (1, 1, \ldots, 1)$. Then $|\mathbf{u} - \mathbf{v}| = \sqrt{m}$ and $|f(\mathbf{u}) - f(\mathbf{v})| = m$, and thus $|f(\mathbf{u}) - f(\mathbf{v})| = \sqrt{m}\,|\mathbf{u} - \mathbf{v}|$. This shows that the inequality proved in Proposition 8.2 is sharp, i.e., there exist instances where, with an appropriate choice of the function f and the points $\mathbf{u}$ and $\mathbf{v}$, the stated upper bound on $|f(\mathbf{u}) - f(\mathbf{v})|$ is attained.

Corollary 8.3 *Let X be an open set in $\mathbb{R}^m$ that is a product of open intervals, let $\varphi: X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, and let M be a positive constant. Suppose that*

$$\left|\frac{\partial f_j}{\partial x_i}\right| \leq M$$

throughout the open set X for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, where $f_j: X \to \mathbb{R}$ is the jth component of the map φ. Then

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v})| \leq \sqrt{mn}\, M\, |\mathbf{u} - \mathbf{v}|$$

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

Proof Let $\mathbf{u}$ and $\mathbf{v}$ be points of X. Then (on applying the inequality stated in Proposition 8.2,) we find that

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v})|^2 = \sum_{j=1}^{n} (f_j(\mathbf{u}) - f_j(\mathbf{v}))^2 \leq mnM^2\, |\mathbf{u} - \mathbf{v}|^2.$$

The result follows. ∎

Corollary 8.4 *Let X be an open set in $\mathbb{R}^m$, let $\varphi: X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, and let M be a positive constant. Suppose that*

$$\left|\frac{\partial f_j}{\partial x_i}\right| \leq M$$

throughout the open set X for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, where $f_j: X \to \mathbb{R}$ is the jth component of the map φ. Then the function φ is continuous on X.

Proof Let $\mathbf{p}$ be a point of X. The set X is open in $\mathbb{R}^m$, and therefore there exists an open set V that is a product of open intervals such that $\mathbf{p} \in V$ and $V \subset X$. It then follows from Corollary 8.3 that

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v})| \leq \sqrt{mn}\, M\, |\mathbf{u} - \mathbf{v}|$$

for all points $\mathbf{u}$ and $\mathbf{v}$ of V. This inequality ensures that the function φ is continuous around the point $\mathbf{p}$. The result follows. ∎

8.3 Functions with Continuous Partial Derivatives

We now investigate the behaviour of functions of several real variables whose first order partial derivatives are continuous.

Definition Let X be an open set in $\mathbb{R}^m$, and let $f: X \to \mathbb{R}^n$ be real-valued function on X, and let $\mathbf{p}$ be a point of X. Suppose that the first order partial derivatives of f are defined at the point $\mathbf{p}$. The *gradient* $(\nabla f)_{\mathbf{p}}$ of f at the point $\mathbf{p}$ is the m-dimensional vector whose components are the partial derivatives of the function f at the point $\mathbf{p}$. Thus

$$(\nabla f)_{\mathbf{p}} = \left(\left.\frac{\partial f}{\partial x_1}\right|_{\mathbf{x}=\mathbf{p}}, \left.\frac{\partial f}{\partial x_2}\right|_{\mathbf{x}=\mathbf{p}}, \ldots \left.\frac{\partial f}{\partial x_m}\right|_{\mathbf{x}=\mathbf{p}} \right).$$

The mere existence of first order partial derivatives of a real-valued function around a given point is not sufficient to enable the gradient of that function to provide a reasonable approximation to the function around that point. On the other hand, as we shall see, if the first order partial derivatives are not only defined around that point but are also continuous there, then the gradient of the function does determine a "first order" approximation to the function around that point.

Proposition 8.5 *Let X be an open set in $\mathbb{R}^m$, let $f: X \to \mathbb{R}$ be a real-valued function on X, and let $\mathbf{p}$ be a point of X. Suppose that the first order partial derivatives of the function f are defined throughout the set X and are continuous at the point $\mathbf{p}$. Then, given any positive real number ε, there exists some positive real number δ such that*

$$|f(\mathbf{u}) - f(\mathbf{v}) - (\nabla f)_{\mathbf{p}} \,.\, (\mathbf{u} - \mathbf{v})| \leq \varepsilon \, |\mathbf{u} - \mathbf{v}|$$

for all points $\mathbf{u}$ and $\mathbf{v}$ of X satisfying $|\mathbf{u} - \mathbf{p}| < \delta$ and $|\mathbf{v} - \mathbf{p}| < \delta$, where $(\nabla f)_{\mathbf{p}}$ denotes the gradient of the function f at the point $\mathbf{p}$.

Proof Let $\mathbf{p} = (p_1, p_2, \ldots, p_m)$, and let $g: X \to \mathbb{R}$ be the real-valued function on X defined so that

$$g(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{p}) - (\nabla f)_{\mathbf{p}} \,.\, (\mathbf{x} - \mathbf{p})$$

for all $\mathbf{x} \in X$. Then the function g has first order partial derivatives, defined throughout the open set X, which are continuous at the point $\mathbf{p}$. Moreover $g(\mathbf{p}) = 0$ and

$$\left.\frac{\partial g}{\partial x_i}\right|_{\mathbf{x}=\mathbf{p}} = 0$$

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

$$g(\mathbf{u}) - g(\mathbf{v}) = f(\mathbf{u}) - f(\mathbf{v}) - (\nabla f)_{\mathbf{p}} \,.\, (\mathbf{u} - \mathbf{v})$$

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

Let some positive real number ε be given. for $i = 1, 2, \ldots, n$. Now the domain X of the functions f and g is an open subset of $\mathbb{R}^m$. This, together with the continuity of the first order partial derivatives of the function g at the point $\mathbf{p}$, ensures that some positive real number δ can then be chosen small enough to ensure both that

$$\{(x_1, x_2, \ldots, x_m) \in \mathbb{R}^m : |p_i - x_i| \leq \delta \text{ for } i = 1, 2, \ldots, m\} \subset X$$

and also that

$$\left| \frac{\partial g}{\partial x_i} \right| \leq \frac{\varepsilon}{\sqrt{m}}$$

for $i = 1, 2, \ldots, m$ at all points $(x_1, x_2, \ldots, x_m)$ of $\mathbb{R}^m$ that satisfy $|x_i - p_i| < \delta$ for $i = 1, 2, \ldots, m$.

It now follows (on applying Proposition 8.2) that

$$|g(\mathbf{u}) - g(\mathbf{v})| \leq \varepsilon \, |\mathbf{u} - \mathbf{v}|$$

at all points $\mathbf{x}$ of $\mathbb{R}^m$ whose components $x_1, x_2, \ldots, x_m$ satisfy $|x_i - p_i| < \delta$ for $i = 1, 2, \ldots, m$. The result follows. ∎

Corollary 8.6 *Let X be an open set in $\mathbb{R}^m$, let $f\colon X \to \mathbb{R}^n$ be a real-valued function on X, and let $\mathbf{p}$ be a point of X. Suppose that the first order partial derivatives of the function f are defined throughout the set X and are continuous at the point $\mathbf{p}$. Then*

$$\lim_{\mathbf{x} \to \mathbf{p}} \frac{1}{|\mathbf{x} - \mathbf{p}|} \, |f(\mathbf{x}) - f(\mathbf{p}) - (\nabla f)_{\mathbf{p}} \,.\, (\mathbf{x} - \mathbf{p}))| = 0,$$

where $(\nabla f)_{\mathbf{p}}$ denotes the gradient of the function f at the point $\mathbf{p}$.

Proof Proposition 8.5 ensures that, given any positive real number ε, there exists a positive real number δ such that

$$\frac{1}{|\mathbf{x} - \mathbf{p}|} \, |f(\mathbf{x}) - f(\mathbf{p}) - (\nabla f)_{\mathbf{p}} \,.\, (\mathbf{x} - \mathbf{p})| \leq \varepsilon$$

for all points $\mathbf{x}$ of X satisfying $0 < |\mathbf{x} - \mathbf{p}| < \delta$. The result therefore follows directly from the formal definition of limits of functions of several real variables. ∎

Corollary 8.7 *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function on X taking values in $\mathbb{R}^n$, let $f_1, f_2, \ldots, f_n$ be the components of the map φ, and let $\mathbf{p}$ be a point of X. Suppose that the first order partial derivatives of the components of the map φ are defined throughout the set X and are continuous at the point $\mathbf{p}$. Then, given any positive real number ε, there exists some positive real number δ such that*

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v}) - (D\varphi)_{\mathbf{p}}\,(\mathbf{u} - \mathbf{v})| \leq \varepsilon\,|\mathbf{u} - \mathbf{v}|$$

for all points $\mathbf{u}$ and $\mathbf{v}$ of X satisfying $|\mathbf{u} - \mathbf{p}| < \delta$ and $|\mathbf{v} - \mathbf{p}| < \delta$, where

$$(D\varphi)_{\mathbf{p}}\,\mathbf{w} = ((\nabla f_1)_{\mathbf{p}} \,.\, \mathbf{w}, (\nabla f_2)_{\mathbf{p}} \,.\, \mathbf{w}, \ldots, (\nabla f_n)_{\mathbf{p}} \,.\, \mathbf{w})$$

for all $\mathbf{w} \in \mathbb{R}^m$.

Proof It follows from Proposition 8.5 that, given any positive real number ε, there exists some positive real number δ such that

$$|f_j(\mathbf{u}) - f_j(\mathbf{v}) - (\nabla f_j)_{\mathbf{p}} \,.\, (\mathbf{u} - \mathbf{v})| \leq \frac{\varepsilon}{\sqrt{n}}\,|\mathbf{u} - \mathbf{v}|$$

for $j = 1, 2, \ldots, n$, and for all points $\mathbf{u}$ and $\mathbf{v}$ of X satisfying $|\mathbf{u} - \mathbf{p}| < \delta$ and $|\mathbf{v} - \mathbf{p}| < \delta$. Then

$$\begin{aligned}
&|\varphi(\mathbf{u}) - \varphi(\mathbf{v}) - (D\varphi)_{\mathbf{p}}\,(\mathbf{u} - \mathbf{v})|^2 \\
&\qquad = \sum_{j=1}^{n} (f_j(\mathbf{u}) - f_j(\mathbf{v}) - (\nabla f_j)_{\mathbf{p}} \,.\, (\mathbf{u} - \mathbf{v}))^2 \\
&\qquad \leq \varepsilon^2\,|\mathbf{u} - \mathbf{v}|^2
\end{aligned}$$

for all points $\mathbf{u}$ and $\mathbf{v}$ of X satisfying $|\mathbf{u} - \mathbf{p}| < \delta$ and $|\mathbf{v} - \mathbf{p}| < \delta$. The result follows. ∎

Corollary 8.8 *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function on X taking values in $\mathbb{R}^n$, let $f_1, f_2, \ldots, f_n$ be the components of the map φ, and let $\mathbf{p}$ be a point of X. Suppose that the first order partial derivatives of the components of the map φ are defined throughout the set X and are continuous at the point $\mathbf{p}$. Then*

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

where

$$(D\varphi)_{\mathbf{p}}\,\mathbf{w} = ((\nabla f_1)_{\mathbf{p}} \,.\, \mathbf{w}, (\nabla f_2)_{\mathbf{p}} \,.\, \mathbf{w}, \ldots, (\nabla f_n)_{\mathbf{p}} \,.\, \mathbf{w})$$

for all $\mathbf{w} \in \mathbb{R}^m$.

Proof Proposition 8.7 ensures that, given any positive real number ε, there exists a positive real number δ such that

$$\frac{1}{|\mathbf{x}-\mathbf{p}|}\,|\varphi(\mathbf{x})-\varphi(\mathbf{p})-(Df)_{\mathbf{p}}\,(\mathbf{x}-\mathbf{p})|\leq\varepsilon$$

for all points $\mathbf{x}$ of X satisfying $0<|\mathbf{x}-\mathbf{p}|<\delta$. The result therefore follows directly from the formal definition of limits of functions of several real variables. ∎

8.4 Derivatives of Functions of Several Variables

Definition Let X be an open subset of $\mathbb{R}^m$ let $\varphi\colon X\to\mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, let $T\colon\mathbb{R}^m\to\mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X. The function φ is said to be *differentiable* at $\mathbf{p}$, with *derivative* $T\colon\mathbb{R}^m\to\mathbb{R}^n$ if and only if

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

Henceforth we shall usually denote the derivative of a differentiable map $\varphi\colon X\to\mathbb{R}^n$ at a point $\mathbf{p}$ of its domain X by $(D\varphi)_{\mathbf{p}}$.

The derivative $(D\varphi)_{\mathbf{p}}$ of φ at $\mathbf{p}$ is sometimes referred to as the *total derivative* of φ at $\mathbf{p}$. If φ is differentiable at every point of X then we say that φ is differentiable on X.

Lemma 8.9 *Let $T\colon\mathbb{R}^m\to\mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ into $\mathbb{R}^n$. Then T is differentiable at each point $\mathbf{p}$ of $\mathbb{R}^m$, and $(DT)_{\mathbf{p}}=T$.*

Proof This follows immediately from definition of differentiability, given that $T\mathbf{x}-T\mathbf{p}-T(\mathbf{x}-\mathbf{p})=\mathbf{0}$ for all $\mathbf{x}\in\mathbb{R}^m$. ∎

Lemma 8.10 *Let X be an open subset of $\mathbb{R}^m$ let $\varphi\colon X\to\mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, let $T\colon\mathbb{R}^m\to\mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X. Then the function φ is differentiable at $\mathbf{p}$, with derivative T, if and only if, given any positive real number ε, there exists some positive real number δ such that*

$$|\varphi(\mathbf{x})-\varphi(\mathbf{p})-T(\mathbf{x}-\mathbf{p})|\leq\varepsilon|\mathbf{x}-\mathbf{p}|$$

at all points $\mathbf{x}$ of X that satisfy $|\mathbf{x}-\mathbf{p}|<\delta$.

Proof First suppose that the function $\varphi: X \to \mathbb{R}^n$ has the property that, given any positive real number ε_0, there exists some positive real number δ such that

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - T(\mathbf{x} - \mathbf{p})| \leq \varepsilon_0 |\mathbf{x} - \mathbf{p}|$$

at all points $\mathbf{x}$ of X that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$. Let some positive number ε be given, and let ε_0 be chosen so that $0 < \varepsilon_0 < \varepsilon$. Then there exists some positive real number δ such that the above inequality holds at all points $\mathbf{x}$ of X that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$. But then

$$\frac{1}{|\mathbf{x} - \mathbf{p}|} |\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - T(\mathbf{x} - \mathbf{p})| < \varepsilon$$

at all points $\mathbf{x}$ of X that satisfy $0 < |\mathbf{x} - \mathbf{p}| < \delta$, and therefore

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

Thus the function φ is differentiable at the point $\mathbf{p}$.

Conversely suppose that the function φ is differentiable at the point $\mathbf{p}$. Let some positive real number ε be given. Then there exists some positive real number δ such that

$$\frac{1}{|\mathbf{x} - \mathbf{p}|} |\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - T(\mathbf{x} - \mathbf{p})| < \varepsilon$$

at all points $\mathbf{x}$ of X that satisfy $0 < |\mathbf{x} - \mathbf{p}| < \delta$. Considering separately the cases when $\mathbf{x} = \mathbf{p}$ and when $0 < |\mathbf{x} - \mathbf{p}| < \delta$, it then follows that

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - T(\mathbf{x} - \mathbf{p})| \leq \varepsilon |\mathbf{x} - \mathbf{p}|$$

at all points $\mathbf{x}$ of X that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$. The result follows. ∎

Lemma 8.11 *Let X be an open subset of $\mathbb{R}^m$ let $\varphi: X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X. Suppose that the function φ is differentiable at the point $\mathbf{p}$. Then φ is continuous at $\mathbf{p}$.*

Proof Suppose that the function φ is differentiable at $\mathbf{p}$ with derivative $(D\varphi)_{\mathbf{p}}$. It then follows from the definition of differentiability that

$$\lim_{\mathbf{x} \to \mathbf{p}} \frac{1}{|\mathbf{x} - \mathbf{p}|} |\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| = 0.$$

It then follows from basic properties of limits that

$$\begin{aligned}
&\lim_{\mathbf{x}\to\mathbf{p}} |\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})| \\
&\quad = \left(\lim_{\mathbf{x}\to\mathbf{p}} |\mathbf{x}-\mathbf{p}|\right)\left(\lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|} |\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})|\right) \\
&\quad = 0.
\end{aligned}$$

Therefore

$$\lim_{\mathbf{x}\to\mathbf{p}} (\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})) = \mathbf{0}.$$

But then

$$\lim_{\mathbf{x}\to\mathbf{p}} \varphi(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{p}} (\varphi(\mathbf{p}) + (D\varphi)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})) = \varphi(\mathbf{p}).$$

Consequently the function φ is continuous at $\mathbf{p}$. The result follows. ∎

Proposition 8.12 *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function on X taking values in $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X. Suppose that the first order partial derivatives of the components of the map φ are defined throughout some open set to which the point $\mathbf{p}$ belongs and are also continuous at the point $\mathbf{p}$ itself. Then the function φ is differentiable at the point $\mathbf{p}$.*

Proof Let $f_1, f_2, \ldots, f_n$ be the components of the map φ. It follows from Corollary 8.8 that

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

where

$$(D\varphi)_{\mathbf{p}}\,\mathbf{w} = ((\nabla f_1)_{\mathbf{p}} \,.\, \mathbf{w}, (\nabla f_2)_{\mathbf{p}} \,.\, \mathbf{w}, \ldots, (\nabla f_n)_{\mathbf{p}} \,.\, \mathbf{w})$$

for all $\mathbf{w} \in \mathbb{R}^m$. The function φ therefore satisfies the definition of differentiability at $\mathbf{p}$, as required. ∎

8.5 The Jacobian Matrix of a Differentiable Function

Proposition 8.13 *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X at which the function φ is differentiable. Let $\mathbf{w}$ be an element of $\mathbb{R}^m$. Then*

$$(D\varphi)_{\mathbf{p}}\mathbf{w} = \lim_{t\to 0} \frac{1}{t}\left(\varphi(\mathbf{p}+t\mathbf{w}) - \varphi(\mathbf{p})\right).$$

Thus the derivative $(D\varphi)_{\mathbf{p}}$ of φ at $\mathbf{p}$ is uniquely determined by the map φ.

Proof Let $\mathbf{w}$ be some vector in $\mathbb{R}^m$, and let some positive real number ε be given. Then let some positive real number ε_0 be chosen so that $\varepsilon_0|\mathbf{w}| \leq \varepsilon$. The differentiability of φ at $\mathbf{p}$ then ensures that there exists some positive real number δ_0 which is small enough to ensure that

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| \leq \varepsilon_0|\mathbf{x} - \mathbf{p}|$$

at all points $\mathbf{x}$ of X that satisfy $|\mathbf{x} - \mathbf{p}| < \delta_0$. (see Lemma 8.10). Setting $\mathbf{x} = \mathbf{p} + t\mathbf{w}$, and choosing a positive real number δ for which $|\mathbf{w}|\delta \leq \delta_0$, we find that

$$\frac{1}{|t|}\,|\varphi(\mathbf{p} + t\mathbf{w}) - \varphi(\mathbf{p}) - t(D\varphi)_{\mathbf{p}}\mathbf{w}| \leq \varepsilon_0\,|\mathbf{w}| \leq \varepsilon$$

whenever $0 < |t| < \delta$.

Considering separately the cases as t tends to zero through positive and negative values, we can then conclude that

$$\lim_{t \to 0^+} \frac{1}{t}\,(\varphi(\mathbf{p} + t\mathbf{w}) - \varphi(\mathbf{p}) - t(D\varphi)_{\mathbf{p}}\mathbf{w}) = \mathbf{0}$$

and

$$\lim_{t \to 0^-} \frac{1}{t}\,(\varphi(\mathbf{p} + t\mathbf{w}) - \varphi(\mathbf{p}) - t(D\varphi)_{\mathbf{p}}\mathbf{w}) = \mathbf{0},$$

It follows that

$$\lim_{t \to 0} \frac{1}{t}\,(\varphi(\mathbf{p} + t\mathbf{w}) - \varphi(\mathbf{p})) = (D\varphi)_{\mathbf{p}}\mathbf{w},$$

as required. ∎

Corollary 8.14 *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X at which the function φ is differentiable. Then the derivative $(D\varphi)_{\mathbf{p}}$ of φ at the point $\mathbf{p}$ is uniquely determined by the map φ.*

Proof The result of Proposition 8.13 shows that, for all $\mathbf{w} \in \mathbb{R}^m$, the value of $(D\varphi)_{\mathbf{p}}\,\mathbf{w}$ is expressible as a limit involving the function φ itself and is thus uniquely determined by the function φ itself. Thus there cannot be more than one linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ that can represent the derivative of the function φ at the point $\mathbf{p}$. ∎

Corollary 8.15 *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, and let $\mathbf{p}$ be a point of X at which the function φ is differentiable. Let $f_1, f_2, \ldots, f_n$ denote the components of the function let $\varphi\colon X \to \mathbb{R}^n$. Then the first order partial derivatives of the components of φ are all defined at the point $\mathbf{p}$, and the derivative $(D\varphi)_{\mathbf{p}}$ of the map φ at the* *point* $\mathbf{p}$ *is represented by the* $n \times m$ *matrix whose coefficient in the* i*th row and* j*th column is equal to the value at* $\mathbf{p}$ *of the partial derivative* $\dfrac{\partial f_i}{\partial x_j}$ *of* f_i *with respect to the* j*th coordinate function* x_j *on* X*.*

Proof Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ denote the standard basis of $\mathbb{R}^m$, where the ith component of the vector $\mathbf{e}_j$ is equal to 1 when $i = j$, but is equal to zero otherwise. Basic linear algebra ensures that the linear transformation $(D\varphi)_{\mathbf{p}}: \mathbb{R}^m \to \mathbb{R}^n$ is represented by the matrix $J(\mathbf{p})$ whose coefficient $J_{i,j}(\mathbf{p})$ in the ith row and jth column is equal to the ith component of the vector $(D\varphi)_{\mathbf{p}}\, \mathbf{e}_j$. It then follows (applying Proposition 8.14) that

$$J_{i,j}(\mathbf{p}) = \lim_{t \to 0} \frac{1}{t} \left(f_i(\mathbf{p} + t\mathbf{e}_j) - f_i(\mathbf{e}_j) \right) = \left. \frac{\partial f_i}{\partial x_j} \right|_{\mathbf{x}=\mathbf{p}},$$

as required. ∎

Let X be an open set in $\mathbb{R}^m$, let $\varphi: X \to \mathbb{R}^n$ be a differentiable function mapping X into $\mathbb{R}^n$. Corollary 8.15 ensures that the derivative of φ at any point $\mathbf{p}$ of X is the linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ that sends $\mathbf{w} \in \mathbb{R}^m$ to $J(\mathbf{p})\mathbf{w}$, where J is the $n \times m$ matrix

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

of functions on X whose coefficients are the first order partial derivatives of the components $f_1, f_2, \ldots, f_n$ of the map φ. This matrix of partial derivatives is known as the *Jacobian matrix* of the map φ.

Example Let $\varphi: \mathbb{R}^2 \to \mathbb{R}^2$ be defined so that

$$\varphi\left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}$$

for all real numbers x and y. Let p and q be fixed real numbers. Then

$$\varphi\left(\begin{pmatrix} x \\ y \end{pmatrix} \right) - \varphi\left(\begin{pmatrix} p \\ q \end{pmatrix} \right)$$

$$
\begin{aligned}
&= \begin{pmatrix} x^2-y^2 \\ 2xy \end{pmatrix} - \begin{pmatrix} p^2-q^2 \\ 2pq \end{pmatrix} \\
&= \begin{pmatrix} (x+p)(x-p)-(y+q)(y-q) \\ 2q(x-p)+2p(y-q)+2(x-p)(y-q) \end{pmatrix} \\
&= \begin{pmatrix} 2p(x-p)-2q(y-q)+(x-p)^2-(y-q)^2 \\ 2q(x-p)+2p(y-q)+2(x-p)(y-q) \end{pmatrix} \\
&= \begin{pmatrix} 2p & -2q \\ 2q & 2p \end{pmatrix}\begin{pmatrix} x-p \\ y-q \end{pmatrix} + \begin{pmatrix} (x-p)^2-(y-q)^2 \\ 2(x-p)(y-q) \end{pmatrix}.
\end{aligned}
$$

Now, given $(x, y) \in \mathbb{R}^2$, let $r = \sqrt{(x-p)^2 + (y-q)^2}$. Then $|x-p| < r$ and $|y-q| < r$, and therefore

$$|(x-p)^2 - (y-q)^2| \leq |x-p|^2 + |y-q|^2 \leq 2r^2$$

and $2(x-p)(y-q) \leq 2r^2$, and thus

$$\frac{|(x-p)^2-(y-q)^2|}{\sqrt{(x-p)^2+(y-q)^2}} \leq 2r \quad \text{and} \quad \frac{|2(x-p)(y-q)|}{\sqrt{(x-p)^2+(y-q)^2}} \leq 2r.$$

Thus, given any positive real number ε, let $\delta = \frac{1}{2}\varepsilon$. Then

$$\left|\frac{(x-p)^2-(y-q)^2}{\sqrt{(x-p)^2+(y-q)^2}}\right| < \varepsilon \quad \text{and} \quad \left|\frac{2(x-p)(y-q)}{\sqrt{(x-p)^2+(y-q)^2}}\right| < \varepsilon$$

whenever $0 < |(x, y) - (p, q)| < \delta$. It follows therefore that

$$\lim_{(x,y)\to(0,0)} \frac{1}{\sqrt{(x-p)^2+(y-q)^2}} \begin{pmatrix} (x-p)^2-(y-q)^2 \\ 2(x-p)(y-q) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Thus the function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable at (p, q), and the derivative of this function at (p, q) is the linear transformation represented by the matrix

$$\begin{pmatrix} 2p & -2q \\ 2q & 2p \end{pmatrix}.$$

8.6 Sums, Differences and Multiples of Differentiable Functions

Proposition 8.16 *Let X be an open set in $\mathbb{R}^m$, and let $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ be functions mapping X into $\mathbb{R}$. Let $\mathbf{p}$ be a point of X. Suppose* *that f and g are differentiable at $\mathbf{p}$. Then the functions $f+g$ and $f-g$ are differentiable at $\mathbf{p}$, and*

$$D(f+g)_{\mathbf{p}} = (Df)_{\mathbf{p}} + (Dg)_{\mathbf{p}}$$

and

$$D(f-g)_{\mathbf{p}} = (Df)_{\mathbf{p}} - (Dg)_{\mathbf{p}}.$$

Moreover, given any real number c, the function cf is differentiable at $\mathbf{p}$ and

$$D(cf)_{\mathbf{p}} = c(Df)_{\mathbf{p}}.$$

Proof The limit of a sum of functions is the sum of the limits of those functions, provided that these limits exist. Applying the definition of differentiability, it therefore follows that

$$\begin{aligned}
&\lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|}\Bigl(f(\mathbf{x}) + g(\mathbf{x}) - (f(\mathbf{p}) + g(\mathbf{p})) - ((Df)_{\mathbf{p}} + (Dg)_{\mathbf{p}})(\mathbf{x}-\mathbf{p})\Bigr) \\
&\qquad = \lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|}\Bigl(f(\mathbf{x}) - f(\mathbf{p}) - (Df)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})\Bigr) \\
&\qquad\qquad + \lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|}\Bigl(g(\mathbf{x}) - g(\mathbf{p}) - (Dg)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})\Bigr) \\
&\qquad = 0.
\end{aligned}$$

Therefore

$$D(f+g)_{\mathbf{p}} = (Df)_{\mathbf{p}} + (Dg)_{\mathbf{p}}.$$

Also the function $-g$ is differentiable, with derivative $-(Dg)_{\mathbf{p}}$. It follows that $f-g$ is differentiable, with derivative $(Df)_{\mathbf{p}} - (Dg)_{\mathbf{p}}$.

Let c be a real number. Then

$$\begin{aligned}
&\lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|}\Bigl(cf(\mathbf{x}) - cf(\mathbf{p}) - c(Df)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})\Bigr) \\
&\qquad = c \lim_{\mathbf{x}\to\mathbf{p}} \frac{1}{|\mathbf{x}-\mathbf{p}|}\Bigl(f(\mathbf{x}) - f(\mathbf{p}) - (Df)_{\mathbf{p}}(\mathbf{x}-\mathbf{p})\Bigr) \\
&\qquad = 0.
\end{aligned}$$

It follows that the function cf is differentiable at p, and $D(cf)_{\mathbf{p}} = c(Df)_{\mathbf{p}}$, as required. ∎

8.7 An Inequality limiting the Growth of a Differentiable Function

We shall derive an inequality bounding the growth of a function of of several real variables around a point where it is differentiable. The statement of the result makes reference to the *operator norm* of a linear transformation. Accordingly we now give the definition of operator norms.

Definition Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. The *operator norm* $\|T\|_{\mathrm{op}}$ of T is the smallest non-negative real number with the property that $|T\mathbf{w}| \leq \|T\|_{\mathrm{op}} |\mathbf{w}|$ for all $\mathbf{w} \in \mathbb{R}^m$.

The operator norm $\|T\|_{\mathrm{op}}$ of a linear transformation $T\colon \mathbb{R}^m \to \mathbb{R}^n$ may be characterized as the maximum value attained by $|T\mathbf{w}|$ as $\mathbf{w}$ ranges over all vectors in $\mathbb{R}^m$ that satisfy $|\mathbf{w}| = 1$.

Lemma 8.17 *Let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, and let $\|T\|_{\mathrm{op}}$ denote the operator norm of T. Also let $A_{i,j} = T\mathbf{e}_i \,.\, T\mathbf{e}_j$ for all integers i and j between 1 and m, where, for each integer j between 1 and m, the vector $\mathbf{e}_j$ is the jth standard vector whose jth component is equal to one and whose other components are all zero. Then let A denote the matrix whose coefficient in the ith row and jth column is $A_{i,j}$, and let $\lambda_{\max}$ denote the maximum eigenvalue of the real symmetric matrix A. Then $\|T\|_{\mathrm{op}} = \sqrt{\lambda_{\max}}$.*

Proof The matrix A is a real symmetric matrix, and, given any such matrix, there exists an orthogonal matrix R, with transpose R^T for which RAR^T is diagonal. The inverse of the matrix R is then equal to its transpose R^T. Let λ_i denote the coefficient in the ith row and column of the diagonal matrix RAR^T for $i = 1, 2, \ldots, m$. Then $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the eigenvalues of the matrix A.

Let $\mathbf{w} \in \mathbb{R}^m$, and let $\mathbf{w} = (w_1, w_2, \ldots, w_m)$. Then

$$|T\mathbf{w}|^2 = \left(\sum_{i=1}^m w_i T\mathbf{e}_i\right) . \left(\sum_{j=1}^m w_j T\mathbf{e}_j\right) = \sum_{i=1}^m \sum_{j=1}^m A_{i,j} w_i w_j.$$

Thus if we represent $\mathbf{w}$ in matrix algebra as a column vector with coefficients $w_1, w_2, \ldots, w_n$ then $|T\mathbf{w}|^2 = \mathbf{w}^T A \mathbf{w}$, where $\mathbf{w}^T$ denotes the row vector that is the transpose of the column vector $\mathbf{w}$.

Let $\mathbf{u} = R\mathbf{w}$, so that $u_k = \sum\limits_{j=1}^m R_{k,j} w_j$ for $k = 1, 2, \ldots, m$. Then $\mathbf{w} = R^T\mathbf{u}$, and therefore

$$|T\mathbf{w}|^2 = \mathbf{w}^T A \mathbf{w} = \mathbf{u}^T RAR^T \mathbf{u} = \sum_{k=1}^m \lambda_k u_k^2 \leq \lambda_{\max} \sum_{k=1}^m u_k^2.$$

Moreover

$$\sum_{k=1}^n u_k^2 = \mathbf{u}^T\mathbf{u} = \mathbf{w}^T R^T R \mathbf{w} = \mathbf{w}^T\mathbf{w} = |\mathbf{w}|^2.$$

We conclude therefore that

$$|T\mathbf{w}|^2 \leq \lambda_{\max}|\mathbf{w}|^2$$

for all $\mathbf{w} \in \mathbb{R}^m$. Moreover if $\mathbf{w} = R^T\mathbf{e}_k$, where k is an integer between 1 and m for which $\lambda_k = \lambda_{\max}$, then

$$|T\mathbf{w}|^2 = \lambda_{\max}|\mathbf{w}|^2.$$

The result follows. ∎

Proposition 8.18 *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, let $\mathbf{p}$ be a point of X at which the function φ is differentiable, and let M be a real number satisfying $M > \|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}}$, where $\|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}}$ denotes the operator norm of the derivative $(D\varphi)_{\mathbf{p}}$ of φ at $\mathbf{p}$. Then there exists some positive real number δ such that*

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

for all points $\mathbf{x}$ of X satisfying $|\mathbf{x} - \mathbf{p}| < \delta$.

Proof Let $\varepsilon = M - \|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}}$. Then $\varepsilon > 0$. Now

$$|(D\varphi)_{\mathbf{p}}\mathbf{w}| \leq \|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}}\,|\mathbf{w}|$$

for all $\mathbf{w} \in \mathbb{R}^m$. Also the differentiability of the function φ at the point $\mathbf{p}$ ensures that there exists some positive real number δ that is small enough to ensure that

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| \leq \varepsilon\,|\mathbf{x} - \mathbf{p}|$$

for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$ (see Lemma 8.10).

It then follows from the Triangle Inequality satisfied by the Euclidean distance function that

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{p})| \leq |(D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| + \varepsilon|\mathbf{x} - \mathbf{p}|$$

for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$. But the definition of the operator norm ensures that

$$|(D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| \leq \|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}}\,|\mathbf{x} - \mathbf{p}|$$

for all $\mathbf{x} \in X$. It follows that

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{p})| \leq (\|(D\varphi)_{\mathbf{p}}\|_{\mathrm{op}} + \varepsilon)|\mathbf{x} - \mathbf{p}| = M\,|\mathbf{x} - \mathbf{p}|$$

for all $\mathbf{x} \in X$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$, as required. ∎

8.8 The Product Rule for Functions of Several Variables

Let $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ be real-valued functions defined over an open subset X of $\mathbb{R}^m$. We denote by $f \cdot g$ the product of the functions f and g, defined so that $(f \cdot g)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ for all $\mathbf{x} \in X$.

Proposition 8.19 (Product Rule) *Let X be an open set in $\mathbb{R}^m$, let $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ be real-valued functions on X, and let $\mathbf{p}$ be a point of X. Suppose that f and g are differentiable at $\mathbf{p}$. Then the function $f \cdot g$ is differentiable at $\mathbf{p}$, and*

$$D(f \cdot g)_{\mathbf{p}} = g(\mathbf{p})(Df)_{\mathbf{p}} + f(\mathbf{p})(Dg)_{\mathbf{p}}.$$

Proof The differentiability of the functions f and g at $\mathbf{p}$ ensures the existence of positive real numbers ensures that there exist positive real numbers M, N and δ that ensure that

$$|f(\mathbf{x}) - f(\mathbf{p})| \leq M\,|\mathbf{x} - \mathbf{p}|$$

and

$$|g(\mathbf{x}) - g(\mathbf{p})| \leq N\,|\mathbf{x} - \mathbf{p}|$$

at all points $\mathbf{x}$ of $\mathbb{R}^m$ that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$. (This follows on applying Proposition 8.18.) Let $h\colon X \to \mathbb{R}$ be the real-valued function on X defined so that

$$\begin{aligned} h(\mathbf{x}) &= (f(\mathbf{x}) - f(\mathbf{p}))(g(\mathbf{x}) - g(\mathbf{p})) \\ &= f(\mathbf{x})g(\mathbf{x}) + f(\mathbf{p})g(\mathbf{p}) - f(\mathbf{p})g(\mathbf{x}) - f(\mathbf{x})g(\mathbf{p}) \end{aligned}$$

for all $\mathbf{x} \in X$. Then $h(\mathbf{p}) = 0$ and

$$\begin{aligned} \frac{|h(\mathbf{x})|}{|\mathbf{x} - \mathbf{p}|} &= \frac{1}{|\mathbf{x} - \mathbf{p}|} |f(\mathbf{x}) - f(\mathbf{p})|\,|g(\mathbf{x}) - g(\mathbf{p})| \\ &\leq MN\,|\mathbf{x} - \mathbf{p}| \end{aligned}$$

at all points $\mathbf{x}$ of $\mathbb{R}^m$ that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$. It follows that

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

It then follows from the definition of differentiability that the function $h\colon X \to \mathbb{R}$ is differentiable at the point $\mathbf{p}$, and $(Dh)_{\mathbf{p}} = 0$.

Now

$$f(\mathbf{x})g(\mathbf{x}) = f(\mathbf{p})g(\mathbf{x}) + g(\mathbf{p})f(\mathbf{x}) - f(\mathbf{p})g(\mathbf{p}) + h(\mathbf{x})$$

for all $\mathbf{x} \in X$. Differentiating, and using the fact that $(Dh)_{\mathbf{p}} = 0$, we find that $f \cdot g$ is differentiable at $\mathbf{p}$, and

$$(D(f \cdot g))_{\mathbf{p}} = f(\mathbf{p})\,(Dg)_{\mathbf{p}} + g(\mathbf{p})\,(Df)_{\mathbf{p}},$$

as required. ∎

8.9 The Chain Rule for Functions of Several Variables

Proposition 8.20 (Chain Rule) *Let X and Y be open sets in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively, let $\varphi: X \to \mathbb{R}^n$ and $\psi: Y \to \mathbb{R}^k$ be functions mapping X and Y into $\mathbb{R}^n$ and $\mathbb{R}^k$ respectively, where $\varphi(X) \subset Y$, and let $\mathbf{p}$ be a point of X. Suppose that φ is differentiable at $\mathbf{p}$ and that ψ is differentiable at $\varphi(\mathbf{p})$. Then the composition $\psi \circ \varphi: X \to \mathbb{R}^k$ is differentiable at $\mathbf{p}$, and*

$$D(\psi \circ \varphi)_{\mathbf{p}} = (D\psi)_{\varphi(\mathbf{p})} \circ (D\varphi)_{\mathbf{p}}.$$

Thus the derivative of the composition $\psi \circ \varphi$ of the functions at the point $\mathbf{p}$ is the composition of the derivatives of the functions φ and ψ at $\mathbf{p}$ and $\varphi(\mathbf{p})$ respectively.

Proof The differentiability of the functions φ and ψ at $\mathbf{p}$ and $\varphi(\mathbf{p})$ respectively ensures that there exist positive real numbers M, N, δ_1 and η_1 such that the following conditions hold: $\mathbf{x} \in X$ and $|\varphi(\mathbf{x}) - \varphi(\mathbf{p})| \leq M|\mathbf{x} - \mathbf{p}|$ for all $\mathbf{x} \in \mathbb{R}^m$ satisfying $|\mathbf{x}-\mathbf{p}| < \delta_1$; $\mathbf{y} \in Y$ and $|\psi(\mathbf{y})-\psi(\varphi(\mathbf{p}))| \leq N\,|\mathbf{y}-\varphi(\mathbf{p})|$ for all $\mathbf{y} \in \mathbb{R}^n$ satisfying $|\mathbf{y}-\varphi(\mathbf{p})| < \eta_1$; $|(D\psi)_{\varphi(\mathbf{p})}\mathbf{w}| \leq N\,|\mathbf{w}|$ for all $\mathbf{w} \in \mathbb{R}^n$. (This follows on applying Proposition 8.18.)

Let some positive real number ε be given. It follows from the differentiability of ψ at $\varphi(\mathbf{p})$ that there exists some real number η_2, where $0 < \eta_2 \leq \eta_1$, such that

$$\bigl|\psi(\mathbf{y}) - \psi(\varphi(\mathbf{p})) - (D\psi)_{\varphi(\mathbf{p})}(\mathbf{y} - \varphi(\mathbf{p}))\bigr| \leq \frac{\varepsilon}{2M}|\mathbf{y} - \varphi(\mathbf{p})|$$

for all $\mathbf{y} \in Y$ satisfying $|\mathbf{y}-\varphi(\mathbf{p})| < \eta_2$. (This follows from a direct application of Lemma 8.10.) Let some real number δ_2 be chosen so that $0 < \delta_2 \leq \delta_1$ and $M\delta_2 \leq \eta_2$. If $\mathbf{x} \in \mathbb{R}^m$ satisfies $|\mathbf{x} - \mathbf{p}| < \delta_2$ then $\mathbf{x} \in X$ and $|\varphi(\mathbf{x}) - \varphi(\mathbf{p})| \leq M|\mathbf{x} - \mathbf{p}| < \eta_2$. Consequently if $|\mathbf{x} - \mathbf{p}| < \delta_2$ then

$$\begin{aligned}\bigl|\psi(\varphi(\mathbf{x})) - \psi(\varphi(\mathbf{p})) - (D\psi)_{\varphi(\mathbf{p})}(\varphi(\mathbf{x}) - \varphi(\mathbf{p}))\bigr| &\leq \frac{\varepsilon}{2M}|\varphi(\mathbf{x}) - \varphi(\mathbf{p})| \\ &\leq \tfrac{1}{2}\,\varepsilon|\mathbf{x} - \mathbf{p}|.\end{aligned}$$

Now it follows from the differentiability of φ at $\mathbf{p}$ that there exists some real number δ satisfying the inequalities $0 < \delta \leq \delta_2$ that is small enough to ensure that

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

for all $\mathbf{x} \in \mathbb{R}^m$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$. Now $|(D\psi)_{\varphi(\mathbf{p})}\mathbf{w}| \leq N\,|\mathbf{w}|$ for all $\mathbf{w} \in \mathbb{R}^n$. It follows that

$$\begin{aligned} &\big|(D\psi)_{\varphi(\mathbf{p})}(\varphi(\mathbf{x}) - \varphi(\mathbf{p})) - (D\psi)_{\varphi(\mathbf{p})}(D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})\big| \\ &\qquad\qquad \leq \ N\,|\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})| \\ &\qquad\qquad \leq \ \tfrac{1}{2}\varepsilon|\mathbf{x} - \mathbf{p}| \end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^m$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta$.

The inequalities obtained above ensure that $\mathbf{x} \in X$ and

$$\begin{aligned} &\big|\psi(\varphi(\mathbf{x})) - \psi(\varphi(\mathbf{p})) - (D\psi)_{\varphi(\mathbf{p})}(D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})\big| \\ &\quad \leq \ \big|\psi(\varphi(\mathbf{x})) - \psi(\varphi(\mathbf{p})) - (D\psi)_{\varphi(\mathbf{p})}(\varphi(\mathbf{x}) - \varphi(\mathbf{p}))\big| \\ &\qquad\quad + \big|(D\psi)_{\varphi(\mathbf{p})}(\varphi(\mathbf{x}) - \varphi(\mathbf{p})) - (D\psi)_{\varphi(\mathbf{p})}(D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p})\big| \\ &\quad \leq \ \varepsilon|\mathbf{x} - \mathbf{p}| \end{aligned}$$

at all points $\mathbf{x}$ of $\mathbb{R}^m$ that satisfy $|\mathbf{x} - \mathbf{p}| < \delta$. It follows from this that the composition function $\psi \circ \varphi$ is differentiable at $\mathbf{p}$, and that $(D(\psi \circ \varphi))_{\mathbf{p}} = (D\psi)_{\varphi(\mathbf{p})} \circ (D\varphi)_{\mathbf{p}}$, as required. ∎