Chapter 2

Differentiation in higher dimensions

2.1 The Total Derivative

Recall that if $f : \mathbb{R} \to \mathbb{R}$ is a 1-variable function, and $a \in \mathbb{R}$, we say that f is differentiable at $x = a$ if and only if the ratio $\frac{f(a+h)-f(a)}{h}$ tends to a finite limit, denoted $f'(a)$, as h tends to 0.

There are two possible ways to generalize this for vector fields

$$f : \mathcal{D} \to \mathbb{R}^m,\ \mathcal{D} \subseteq \mathbb{R}^n,$$

for points a in the *interior* $\mathcal{D}^0$ of $\mathcal{D}$. (The interior of a set X is defined to be the subset X^0 obtained by removing all the boundary points. Since every point of X^0 is an interior point, it is open.) The reader seeing this material for the first time will be well advised to stick to vector fields f with domain all of $\mathbb{R}^n$ in the beginning. Even in the one dimensional case, if a function is defined on a closed interval $[a, b]$, say, then one can properly speak of differentiability only at points in the open interval (a, b).

The first thing one might do is to fix a vector v in $\mathbb{R}^n$ and saythat f is **differentiable along** v iff the following limit makes sense:

$$\lim_{h \to 0} \frac{1}{h} \left(f(a + hv) - f(a) \right).$$

When it does, we write $f'(a; v)$ for the limit. Note that this definition makes sense because a is an interior point. Indeed, under this hypothesis, $\mathcal{D}$ contains a basic open set U containing a, and so $a + hv$ will, for small enough h, fall into U, allowing us to speak of $f(a + hv)$. This derivative behaves exactly like the one variable derivative and has analogous properties. For example, we have the following

Theorem 1 *(Mean Value Theorem for scalar fields) Suppose f is a scalar field. Assume $f'(a+tv;v)$ exists for all $0 \leq t \leq 1$. Then there is a t_o with $0 \leq t_o \leq 1$ for which $f(a+v) - f(a) = f'(a+t_0 v; v)$.*

Proof. Put $\phi(t) = f(a+tv)$. By hypothesis, ϕ is differentiable at every t in $[0,1]$, and $\phi'(t) = f'(a+tv;v)$. By the one variable mean value theorem, there exists a t_0 such that $\phi'(t_0)$ is $\phi(1) - \phi(0)$, which equals $f(a+v) - f(a)$. Done.

When v is a **unit vector**, $f'(a;v)$ is called the **directional derivative** of f at a in the direction of v.

The disadvantage of this construction is that it forces us to study the change of f in one direction at a time. So we revisit the one-dimensional definition and note that the condition for differentiability there is equivalent to requiring that there exists a constant $c\,(= f'(a))$, such that $\lim_{h\to 0}\left(\frac{f(a+h) - f(a) - ch}{h}\right) = 0$. If we put $L(h) = f'(a)h$, then $L : \mathbb{R} \to \mathbb{R}$ is clearly a linear map. We generalize this idea in higher dimensions as follows:

Definition. Let $f : \mathcal{D} \to \mathbb{R}^m$ ($\mathcal{D} \subseteq \mathbb{R}^n$) be a vector field and a an interior point of $\mathcal{D}$. Then f is differentiable at $x = a$ if and only if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$(*) \qquad \lim_{u\to 0} \frac{||f(a+u) - f(a) - L(u)||}{||u||} = 0.$$

Note that the norm $||\cdot||$ denotes the length of vectors in $\mathbb{R}^m$ in the numerator and in $\mathbb{R}^n$ in the denominator. This should not lead to any confusion, however.

Lemma 1 *Such an L, if it exists, is unique.*

Proof. Suppose we have $L, M : \mathbb{R}^n \to \mathbb{R}^m$ satisfying (*) at $x = a$. Then

$$\begin{aligned}
\lim_{u\to 0} \frac{||L(u) - M(u)||}{||u||} &= \lim_{u\to 0} \frac{||L(u) + f(a) - f(a+u) + (f(a+u) - f(a) - M(u))||}{||u||} \\
&\leq \lim_{u\to 0} \frac{||L(u) + f(a) - f(a+u)||}{||u||} \\
&\quad + \lim_{u\to 0} \frac{||f(a+u) - f(a) - M(u)||}{||u||} = 0.
\end{aligned}$$

Pick any non-zero $v \in \mathbb{R}^n$, and set $u = tv$, with $t \in \mathbb{R}$. Then, the linearity of L, M implies that $L(tv) = tL(v)$ and $M(tv) = tM(v)$. Consequently, we have

$$\begin{aligned}\lim_{t\to 0} \frac{||L(tv) - M(tv)||}{||tv||} &= 0 \\ &= \lim_{t\to 0} \frac{|t|\,||L(v) - M(v)||}{|t|\,||v||} \\ &= \frac{1}{||v||}||L(v) - M(v)||.\end{aligned}$$

Then $L(v) - M(v)$ must be zero.

Definition. If the limit condition $(*)$ holds for a linear map L, we call L the **total derivative** of f at a, and denote it by T_af.

It is mind boggling at first to think of the derivative as a linear map. A natural question which arises immediately is to know what the value of T_af is at any vector v in $\mathbb{R}^n$. We will show in section 2.3 that this value is precisely $f'(a; v)$, thus linking the two generalizations of the one-dimensional derivative.

Sometimes one can guess what the answer should be, and if (*) holds for this choice, then it must be the derivative by uniqueness. Here are **two examples** which illustrate this.

(1) Let f be a **constant vector field**, i.e., there exists a vector $w \in \mathbb{R}^m$ such that $f(x) = w$, for all x in the domain $\mathcal{D}$. Then we claim that f is differentiable at any $a \in \mathcal{D}^0$ with **derivative zero**. Indeed, if we put $L(u) = 0$, for any $u \in \mathbb{R}^n$, then (*) is satisfied, because $f(a + u) - f(a) = w - w = 0$.

(2) Let f be a **linear map**. Then we claim that f is differentiable everywhere with $T_af = f$. Indeed, if we put $L(u) = f(u)$, then by the linearity of f, $f(a + u) - f(a) = f(u)$, and so $f(a+u) - f(a) - L(u)$ is zero for any $u \in \mathbb{R}^n$. Hence (*) holds trivially for this choice of L.

Before we leave this section, it will be useful to take note of the following:

Lemma 2 *Let $f_1, \ldots, f_m$ be the component (scalar) fields of f. Then f is differentiable at a iff each f_i is differentiable at a. Moreover, $Tf(v) = (Tf_1(v), Tf_2(v), \ldots, Tf_n(g))$.*

An easy consequence of this lemma is that, when $\mathbf{n = 1}$, f is differentiable at a iff the following familiar looking limit exists in $\mathbb{R}^m$:

$$\lim_{h\to 0} \frac{f(a + h) - f(a)}{h},$$

allowing us to suggestively write $f'(a)$ instead of $T_a f$. Clearly, $f'(a)$ is given by the vector $(f_1'(a), \ldots, f_m'(a))$, so that $(T_a f)(h) = f'(a)h$, for any $h \in \mathbb{R}$.

Proof. Let f be differentiable at a. For each $v \in \mathbb{R}^n$, write $L_i(v)$ for the i-th component of $(T_a f)(v)$. Then L_i is clearly linear. Since $f_i(a+u) - f_i(u) - L_i(u)$ is the i-th component of $f(a+u) - f(a) - L(u)$, the norm of the former is less than or equal to that of the latter. This shows that (*) holds with f replaced by f_i and L replaced by L_i. So f_i is differentiable for any i. Conversely, suppose each f_i differentiable. Put $L(v) = ((T_a f_1)(v), \ldots, (T_a f_m)(v))$. Then L is a linear map, and by the triangle inequality,

$$||f(a+u) - f(a) - L(u)|| \leq \sum_{i=1}^{m} |f_i(a+u) - f_i(a) - (T_a f_i)(u)|.$$

It follows easily that (*) exists and so f is differentiable at a.

2.2 Partial Derivatives

Let $\{e_1, \ldots, e_n\}$ denote the standard basis of $\mathbb{R}^n$. The directional derivatives along the unit vectors e_j are of special importance.

Definition. Let $j \leq n$. The jth partial derivative of f at $x = a$ is $f'(a; e_j)$, denoted by $\frac{\partial f}{\partial x_j}(a)$ or $D_j f(a)$.

Just as in the case of the total derivative, it can be shown that $\frac{\partial f}{\partial x_j}(a)$ exists iff $\frac{\partial f_i}{\partial x_j}(a)$ exists for each coordinate field f_i.

Example: Define $f : \mathbb{R}^3 \to \mathbb{R}^2$ by

$$f(x, y, z) = (e^{x\sin(y)}, z\cos(y)).$$

All the partial derivatives exist at any $a = (x_0, y_0, z_0)$. We will show this for $\frac{\partial f}{\partial y}$ and leave it to the reader to check the remaining cases. Note that

$$\frac{1}{h}(f(a + he_2) - f(a)) = (\frac{e^{x_0\sin(y_0+h)} - e^{x_0\sin(y_0)}}{h}, z_0\frac{\cos(y_0+h) - \cos(y_0)}{h}).$$

We have to understand the limit as h goes to 0. Then the methods of one variable calculus show that the right hand side tends to the finite limit $(x_0\cos(y_0)e^{x_0\sin(y_0)}, -z_0\sin(y_0))$, which is $\frac{\partial f}{\partial y}(a)$. In effect, the partial derivative with respect to y is calculated like a one variable derivative, keeping x and z fixed. Let us note without proof that $\frac{\partial f}{\partial x}(a)$ is $(\sin(y_0)e^{x_0 \sin(y_0)}, 0)$ and $\frac{\partial f}{\partial z}(a)$ is $(0, \cos(y_0))$.

It is easy to see from the definition that $f'(a; tv)$ equals $tf'(a; v)$, for any $t \in \mathbb{R}$. This follows as $\frac{1}{h}(f(a + h(tv)) - f(a)) = t\frac{1}{th}(f(a + (ht)v) - f(a))$. In particular the Mean Value Theorem for scalar fields gives $f_i(a + hv) - f(a) = hf_i'(a + t_0 hv) = hf_i(a + \tau v)$ for some $0 \le \tau \le h$.

We also have the following

Lemma 3 *Suppose the derivatives of f along any $v \in \mathbb{R}^n$ exist near a and are continuous at a. Then*

$$f'(a; v + v') = f'(a; v) + f'(a; v'),$$

for all v, v' in $\mathbb{R}^n$. In particular, the directional derivatives of f are all determined by the n partial derivatives.

We will do this for the scalar fields f_i. Notice

$$\begin{aligned} f_i(a + hv + hv') - f_i(a) &= f_i(a + hv + hv') - f_i(a + hv) + f_i(a + hv) - f(a) \\ &= hf_i(a + hv + \tau v') + hf_i(a + \tau' v) \end{aligned}$$

where here $0 \le \tau \le h$ and $0 \le \tau' \le h$. Now dividing by h and taking the limit and $h \to 0$ gives $f_i'(a; v + v')$ for the first expression. The last expression gives a sum of two limits

$$\lim_{h \to 0} f_i'(a + hv + \tau v') + \lim_{h \to 0} f_i'(a + \tau' v; v').$$

But this is $f_i'(a; v) + f_i'(a; v')$.Recall both τ and τ' are between 0 and h and so as h goes to 0 so do τ and τ'. Here we have used the continuity of the derivatives of f along any line in a neighborhood of a.

Now pick $e_1, e_2, \ldots, e_n$ the usual orthogonal basis and recall $v = \sum \alpha_i e_i$. Then $f'(a; v) = f'(a; \sum \alpha_i e_i) = \sum \alpha_i f'(a; e_i)$. Also the $f'(a; e_i)$ are the partial derivatives. The Lemma now follows easily.

In the next section (Theorem 1a) we will show that the conclusion of this lemma remains valid without the continuity hypothesis **if** we assume instead that f has a total derivative at a.

The **gradient** of a scalar field g at an interior point a of its domain in $\mathbb{R}^n$ is defined to be the following vector in $\mathbb{R}^n$:

$$\nabla g(a) \;=\; grad\, g(a) \;=\; \left(\frac{\partial g}{\partial x_1}(a), \ldots, \frac{\partial g}{\partial x_n}(a)\right),$$

assuming that the partial derivatives exist at a.

Given a vector field f as above, we can then put together the gradients of its component fields f_i, $1 \leq i \leq m$, and form the following important matrix, called the **Jacobian matrix** at a:

$$Df(a) \;=\; \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{1\leq i\leq m, 1\leq j\leq n} \;\in\; M_{mn}(\mathbb{R}).$$

The i-th row is given by $\nabla f_i(a)$, while the j-th column is given by $\frac{\partial f}{\partial x_j}(a)$. Here we are using the notation $M_{mn}(\mathbb{R})$ for the collection of all $m \times n$-matrices with real coefficients. When $m = n$, we will simply write $M_n(\mathbb{R})$.

2.3 The main theorem

In this section we collect the main properties of the total and partial derivatives.

Theorem 2 *Let $f : \mathcal{D} \to \mathbb{R}^m$ be a vector field, and a an interior point of its domain $\mathcal{D} \subseteq \mathbb{R}^n$.*

(a) *If f is differentiable at a, then for any vector v in $\mathbb{R}^n$,*

$$(T_a f)(v) = f'(a, v).$$

In particular, since $T_a f$ is linear, we have

$$f'(a; \alpha v + \beta v') \;=\; \alpha f'(a; v) + \beta f'(a; v'),$$

for all v, v' in $\mathbb{R}^n$ and α, β in $\mathbb{R}$.

(b) *Again assume that f is differentiable. Then the matrix of the linear map $T_a f$ relative to the standard bases of $\mathbb{R}^n$, $\mathbb{R}^m$ is simply the Jacobian matrix of f at a.*

(c) *f differentiable at $a \;\Rightarrow\; f$ continuous at a.*

(d) *Suppose all the partial derivatives of f exist near a and are continuous at a. Then $T_a f$ exists.*

(e) *(chain rule) Consider*

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^m & \xrightarrow{g} & \mathbb{R}^h. \\ a & \mapsto & b = f(a) & & \end{array}$$

Suppose f is differentiable at a and g is diffentiable at $b = f(a)$. Then the composite function $h = g \circ f$ is differentiable at a, and moreover,

$$T_a h \;=\; T_b g \circ T_a f.$$

In terms of the Jacobian matrices, this reads as

$$Dh(a) \;=\; Dg(b) Df(a) \;\in\; M_{kn}.$$

(f) $(m = 1)$ *Let f, g be scalar fields, differentiable at a. Then*

(i) $T_a(f + g) = T_a f + T_a g$ *(additivity)*

(ii) $T_a(fg) = f(a)T_a g + g(a)T_a f$ *(product rule)*

(iii) $T_a(\frac{f}{g}) = \dfrac{g(a)T_a f - f(a)T_a g}{g(a)^2}$ *if* $g(a) \neq 0$ *(quotient rule)*

The following corollary is an immediate consequence of the theorem, which we will make use of, in the next chapter on normal vectors and extrema.

Corollary 1 *Let g be a scalar field, differentiable at an interior point b of its domain $\mathcal{D}$ in $\mathbb{R}^n$, and let v be any vector in $\mathbb{R}^n$. Then we have*

$$\nabla g(b) \cdot v \;=\; g'(b; v).$$

Furthermore, let ϕ be a function from a subset of $\mathbb{R}$ into $\mathcal{D} \subseteq \mathbb{R}^n$, differentiable at an interior point a mapping to b. Put $h = g \circ \phi$. Then h is differentiable at a with

$$h'(a) \;=\; \nabla g(b) \cdot \phi'(a).$$

Proof of main theorem. (a) It suffices to show that $(T_a f_i)(v) = f_i(a; v)$ for each $i \leq n$. By definition,

$$\lim_{u \to 0} \frac{||f_i(a + u) - f_i(a) - (T_a f_i)(u)||}{||u||} = 0$$

This means that we can write for $u = hv$, $h \in \mathbb{R}$,

$$\lim_{h \to 0} \frac{f_i(a+hv) - f_i(a) - h(T_a f_i)(v)}{|h|||v||} = 0.$$

In other words, the limit $\lim\limits_{h \to 0} \frac{f_i(a+hv)-f_i(a)}{h}$ exists and equals $(T_a f_i)(v)$. Done.

(b) By part (a), each partial derivative exists at a (since f is assumed to be differentiable at a). The matrix of the linear map $T_a f$ is determined by the effect on the standard basis vectors. Let $\{e'_i | 1 \leq i \leq m\}$ denote the standard basis in $\mathbb{R}^m$. Then we have, by definition,

$$(T_a f)(e_j) = \sum_{i=1}^{m} (T_a f_i)(e_j) e'_i = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(a) e'_i.$$

The matrix obtained is easily seen to be $Df(a)$.

(c) First we need the following simple

Lemma 4 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then, $\exists c > 0$ such that $||Tv|| \leq c||v||$ for any $v \in \mathbb{R}^n$.*

Proof of Lemma. Let A be the matrix of T relative to the standard bases. Put $C = \max_j\{||T(e_j)||\}$. If $v = \sum\limits_{j=1}^{n} \alpha_j e_j$, then

$$\begin{aligned} ||T(v)|| &= ||\sum_j \alpha_j T(e_j)|| \leq C \sum_{j=1}^{n} |\alpha_j| \cdot 1 \\ &\leq C(\sum_{j=1}^{n} |\alpha_j|^2)^{1/2} (\sum_{j=1}^{n} 1)^{1/2} \leq C\sqrt{n}||v||, \end{aligned}$$

by the Cauchy–Schwarz inequality. We are done by setting $c = C\sqrt{n}$.

This shows that a linear map is continuous as if $||v - w|| < \delta$ then $||T(v) - T(w)|| = ||T(v - w)|| < c||v - w|| < c\delta$.

(c) Suppose f is differentiable at a. This certainly implies that the limit of the function $f(a+u) - f(a) - (T_a f)(u)$, as u tends to $0 \in \mathbb{R}^n$, is $0 \in \mathbb{R}^m$ (from the very definition of $T_a f$, $||f(a+u) - f(a) - (T_a f)(u)||$ tends to zero "faster" than $||u||$, in particular it tends to zero). Since $T_a f$ is linear, $T_a f$ is continuous (everywhere), so that $\lim_{u \to 0} (T_a f)(u) = 0$. Hence $\lim_{u \to 0} f(a+u) = f(a)$ which means that f is continuous at a.

(d) By hypothesis, all the partial derivatives exist near $a = (a_1, \ldots, a_n)$ and are continuous there. It suffices to show that each f_i is differentiable at a by lemma 2. So we have only to show that (*) holds with f replaced by f_i and $L(u) = f_i'(a; u)$. Write $u = (h_1, \ldots, h_n)$. By Lemma 3, we know that $f_i'(a; -)$ is linear. So

$$L(u) \;=\; \sum_{j=1}^{n} h_j \frac{\partial f_i}{\partial x_j}(a),$$

and we can write

$$f_i(a+u) - f_i(a) \;=\; \sum_{j=1}^{n} (\phi_j(a_j + h_j) - \phi_j(a_j)),$$

where each ϕ_j is a one variable function defined by

$$\phi_j(t) \;=\; f_i(a_1 + h_1, \ldots, a_{j-1} + h_{j-1}, t, a_{j+1}, \ldots, a_n).$$

By the mean value theorem,

$$\phi_j(a_j + h_j) - \phi_j(a_j) \;=\; h_j \phi_j'(t_j) \;=\; h_j \frac{\partial f_i}{\partial x_j}(y(j)),$$

for some $t_j \in [a_j, a_j + h_j]$, with

$$y(j) \;=\; (a_1 + h_1, \ldots, a_{j-1} + h_{j-1}, t_j, a_{j+1}, \ldots, a_n).$$

Putting these together, we see that it suffices to show that the following limit is zero:

$$\lim_{u \to 0} \frac{1}{||u||} \left| \sum_{j=1}^{n} h_j \left(\frac{\partial f_i}{\partial x_j}(a) - \frac{\partial f_i}{\partial x_j}(y(j)) \right) \right|.$$

Clearly, $|h_j| \leq ||u||$, for each j. So it follows, by the triangle inequality, that this limit is bounded above by the sum over j of $\lim_{h_j \to 0} |\frac{\partial f_i}{\partial x_j}(a) - \frac{\partial f_i}{\partial x_j}(y(j))|$, which is zero by the continuity of the partial derivatives at a. Here we are using the fact that each $y(j)$ approaches a as h_j goes to 0. Done.

Proof of (e) Write $L = T_a f$, $M = T_b g$, $N = M \circ L$. *To show:* $T_a h = N$.

Define $F(x) = f(x) - f(a) - L(x-a)$, $G(y) = g(y) - g(b) - M(y-b)$ and $H(x) = h(x) - h(a) - N(x-a)$. Then we have

$$\lim_{x \to a} \frac{||F(x)||}{||x-a||} = 0 = \lim_{y \to b} \frac{||G(y)||}{||y-b||}.$$

So we need to show:

$$\lim_{x\to a} \frac{||H(x)||}{||x-a||} = 0.$$

But

$$H(x) = g(f(x)) - g(b) - M(L(x-a))$$

Since $L(x-a) = f(x) - f(a) - F(x)$, we get

$$H(x) = [g(f(x)) - g(b) - M(f(x) - f(a))] + M(F(x)) = G(f(x)) + M(F(x)).$$

Therefore it suffices to prove:

(i) $\lim\limits_{x\to a} \dfrac{||G(f(x))||}{||x-a||} = 0$ and

(ii) $\lim\limits_{x\to a} \dfrac{||M(F(x))||}{||x-a||} = 0.$

By Lemma 4, we have $||M(F(x))|| \leq c||F(x)||$, for some $c > 0$. Then $\dfrac{||M(F(x))||}{||x-a||} \leq c \lim\limits_{x\to a} \dfrac{||F(x)||}{||x-a||} = 0$, yielding (ii).

On the other hand, we know $\lim\limits_{y\to b} \dfrac{||G(y)||}{||y-b||} = 0$. So we can find, for every $\epsilon > 0$, a $\delta > 0$ such that $||G(f(x))|| < \epsilon||f(x) - b||$ if $||f(x) - b|| < \delta$. But since f is continuous, $||f(x) - b|| < \delta$ whenever $||x - a|| < \delta_1$, for a small enough $\delta_1 > 0$. Hence

$$\begin{aligned} ||G(f(x))|| &< \epsilon||f(x) - b|| = \epsilon||F(x) + L(x-a)|| \\ &\leq \epsilon||F(x)|| + \epsilon||L(x-a)||, \end{aligned}$$

by the triangle inequality. Since $\lim\limits_{x\to a} \frac{||F(x)||}{||x-a||}$ is zero, we get

$$\lim_{x\to a} \frac{||G(f(x))||}{||x-a||} \leq \epsilon \lim_{x\to a} \frac{||L(x-a)||}{||x-a||}.$$

Applying Lemma 4 again, we get $||L(x-a)|| \leq c'||x-a||$, for some $c' > 0$. Now (i) follows easily.

(f) (i) We can think of $f+g$ as the composite $h = s(f,g)$ where $(f,g)(x) = (f(x), g(x))$ and $s(u,v) = u+v$ ("sum"). Set $b = (f(a), g(a))$. Applying (e), we get

$$T_a(f+g) \ = \ T_b(s) \circ T_a(f,g) \ = \ T_a(f) + T_a(g).$$

Done. The proofs of (ii) and (iii) are similar and will be left to the reader.

QED.

Remark. It is important to take note of the fact that a vector field f may be differentiable at a without the partial derivatives being continuous. We have a counterexample already when $n = m = 1$ as seen by taking

$$f(x) \;=\; x^2 \sin\left(\frac{1}{x}\right) \quad \text{if} \quad x \neq 0,$$

and $f(0) = 0$. This is differentiable everywhere. The only question is at $x = 0$, where the relevant limit $\lim\limits_{h\to 0} \frac{f(h)}{h}$ is clearly zero, so that $f'(0) = 0$. But for $x \neq 0$, we have by the product rule,

$$f'(x) \;=\; 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right),$$

which does not tend to $f'(0) = 0$ as x goes to 0. So f' is not continuous at 0.

2.4 Mixed partial derivatives

Let f be a scalar field, and a an interior point in its domain $\mathcal{D} \subseteq \mathbb{R}^n$. For $j, k \leq n$, we may consider the second partial derivative

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(a) \;=\; \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right)(a),$$

when it exists. It is called the *mixed partial derivative* when $j \neq k$, in which case it is of interest to know whether we have the equality

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(a) \;=\; \frac{\partial^2 f}{\partial x_k \partial x_j}(a). \tag{3.4.1}$$

Proposition 1 *Suppose* $\dfrac{\partial^2 f}{\partial x_j \partial x_k}$ *and* $\dfrac{\partial^2 f}{\partial x_k \partial x_j}$ *both exist near a and are continuous there. Then the equality (3.4.1) holds.*

The proof is similar to the proof of part (d) of Theorem 1.