1. **More on differentiability, differentials and linear approximation.**

Let \( m \) and \( n \) be positive integers.

2. **Standard basis vectors.**

**Definition 2.1.** For each \( j \in \{1, \ldots, n\} \) let

\[ e_j \]

be the vector in \( \mathbb{R}^n \) all of whose components are zero except the \( j \)-th which is one. Thus if \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \) then

\[ x = \sum_{j=1}^n x_j e_j. \]

3. **Linear functions.**

**Definition 3.1.** We say \( L : \mathbb{R}^n \to \mathbb{R}^m \) is **linear** if whenever \( c \in \mathbb{R} \) and \( x, y \in \mathbb{R}^n \) we have

(i) \( L(cx) = cL(x) \);

(ii) \( L(x + y) = L(x) + L(y) \).

Suppose \( L : \mathbb{R}^n \to \mathbb{R}^m \) is linear and \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \). Then

\[
L(x) = L \left( \sum_{j=1}^n x_j e_j \right) \\
= \sum_{j=1}^n L(x_j e_j) \\
= \sum_{j=1}^n x_j L(e_j).
\]

(1)

Thus \( L \) is determined by its values on the vectors \( e_j, j \in \{1, \ldots, n\} \). Conversely, if \( w_j \in \mathbb{R}^m, j \in \{1, \ldots, n\} \), and we define

\[
L(x) = \sum_{j=1}^n x_j w_j \quad \text{for} \ x \in \mathbb{R}^n
\]

one easily verifies that \( L \) is linear and \( L(e_j) = w_j \).

4. **The differential; the general case.**

Suppose \( m \) and \( n \) are positive integers , \( A \subset \mathbb{R}^n \),

\[ f : A \to \mathbb{R}^m \]

and \( a \in \text{int} \ A \). (Previously \( m = 1 \).)

**Definition 4.1.** (Partial derivatives.) For each \( j \in \{1, \ldots, n\} \) we let

\[
\partial_j f(a) = \lim_{t \to 0} \frac{f(a + te_j) - f(a)}{t} \in \mathbb{R}^m.
\]
Remark 4.1. Note that if \( m = 1, n = 2 \) and \( \mathbf{a} = (a, b) \) then
\[
\partial_1 f(a, b) = \frac{\partial f}{\partial x}(a, b), \quad \partial_2 f(a, b) = \frac{\partial f}{\partial y}(a, b).
\]

Note that if \( m = 1, n = 3 \) and \( \mathbf{a} = (a, b, c) \) then
\[
\partial_1 f(a, b, c) = \frac{\partial f}{\partial x}(a, b, c), \quad \partial_2 f(a, b, c) = \frac{\partial f}{\partial y}(a, b, c), \quad \partial_3 f(a, b, c) = \frac{\partial f}{\partial z}(a, b, c).
\]

Definition 4.2. We say \( f \) is **differentiable at** \( \mathbf{a} \) if there exists a linear map \( L : \mathbb{R}^n \to \mathbb{R}^m \) such that
\[
\lim_{x \to \mathbf{a}} \frac{|f(x) - f(a) - L(x - a)|}{|x - \mathbf{a}|} = 0.
\]
The linear map \( L \) is immediately seen to be unique; it is called the **differential of** \( f \) **at** \( \mathbf{a} \) and is written
\[
\partial f(\mathbf{a}).
\]

Suppose \( \mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n \) and \( f \) is differentiable at \( \mathbf{a} \). Then
\[
\partial f(\mathbf{a})(\mathbf{v}) = \sum_{j=1}^{n} v_j \partial_j f(\mathbf{a}).
\]

In case \( m = 1 \)
\[
\partial f(\mathbf{a})(\mathbf{v}) = df(\mathbf{a})(\mathbf{v}) = \nabla f(\mathbf{a}) \cdot \mathbf{v}.
\]

If we let
\[
A(\mathbf{x}) = f(\mathbf{a}) + \partial f(\mathbf{a})(\mathbf{x} - \mathbf{a}) \quad \text{and} \quad e(\mathbf{x}) = f(\mathbf{x}) - A(\mathbf{x})
\]
for \( \mathbf{x} \in A \) we find that (2) is equivalent to
\[
\lim_{\mathbf{x} \to \mathbf{a}} \frac{|e(\mathbf{x})|}{|\mathbf{x} - \mathbf{a}|} = 0.
\]

One calls the function \( A \) the **standard affine approximation to** \( f \) **at** \( \mathbf{a} \). The difference \( e = f - A \) is the error in using \( A \) to approximate \( f \).

Theorem 4.1. Suppose for some \( r > 0 \) the function \( f \) has partial derivatives on \( A \cap \mathbf{U}(\mathbf{a}, r) \) which are continuous at \( \mathbf{a} \). Then \( f \) is differentiable at \( \mathbf{a} \) and
\[
\nabla f(\mathbf{a}) = \left( \frac{\partial f}{\partial x_1}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right).
\]

Proof. Apply the corresponding result in the scalar case which we have already obtained to each component of \( f \). \( \square \)

Theorem 4.2. (**The Chain Rule.**) Suppose
(i) \( f \) is differentiable at \( \mathbf{a} \);
(ii) \( l \) is a positive integer, \( B \) is a subset of \( \mathbb{R}^m \) and \( g : B \to \mathbb{R}^l \);
(iii) \( \mathbf{f}(\mathbf{a}) \in \text{int} \ B \) and \( g \) is differentiable at \( \mathbf{f}(\mathbf{a}) \).

Then \( \mathbf{a} \) is an interior point of the domain of \( g \circ f \), \( g \circ f \) is differentiable at \( \mathbf{a} \) and
\[
\partial(g \circ f)(\mathbf{a}) = \partial g(\mathbf{f}(\mathbf{a})) \circ \partial f(\mathbf{a}).
\]

Proof. Use the affine approximation approximations of \( f \) at \( \mathbf{a} \) and \( g \) near \( \mathbf{f}(\mathbf{a}) \) to obtain an affine approximation of \( g \circ f \) near \( f(\mathbf{a}) \). \( \square \)
Remark 4.2. (What to remember about (4). Suppose 
\[ f = (f_1, \ldots, f_m) \quad \text{and} \quad g = (g_1, \ldots, g_l). \]
Then (4) is equivalent to
\[ \partial_j (g_i \circ f)(a) = \sum_{k=1}^m \partial_k g_i(f(a)) \partial_j f_k(a) = \nabla g_i(f(a)) \cdot \partial_j f(a) \]
for \( i = 1, \ldots, l \) and \( j = 1, \ldots n. \)

Example 4.1. Here is an example of the Chain Rule when \( n = 1, m = 2 \) and \( l = 1. \) Suppose \( U \subset \mathbb{R}^2, U \) is open and \( u : U \to \mathbb{R} \) is continuously differentiable on \( u; u \) will correspond to \( g \) above. Suppose \( P : \mathbb{R}^2 \to \mathbb{R}^2 \) is defined by
\[ P(r, \theta) = (r \cos \theta, r \sin \theta), \quad (r, \theta) \in \mathbb{R}^2; \]
\( P \) will correspond to \( f \) above.

We have
\[ \frac{\partial}{\partial r} u(r \cos \theta, r \sin \theta) = \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) \frac{\partial}{\partial r} (r \cos \theta) + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) \frac{\partial}{\partial r} r \sin \theta \]
\[ = \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) \cos \theta + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) r \sin \theta \]
as well as
\[ \frac{\partial}{\partial \theta} u(r \cos \theta, r \sin \theta) = \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) \frac{\partial}{\partial \theta} (r \cos \theta) + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) \frac{\partial}{\partial \theta} r \sin \theta \]
\[ = -\frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) r \sin \theta + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) r \cos \theta. \]

If you thought that was tough wait till you see what comes next! We have
\[ \frac{\partial^2}{\partial r^2} u(r \cos \theta, r \sin \theta) = \frac{\partial}{\partial r} \left( \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) \cos \theta + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) \sin \theta \right) \]
\[ = \frac{\partial}{\partial r} \left( \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) \right) \cos \theta + \frac{\partial}{\partial r} \left( \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) \right) \sin \theta \]
\[ = \left( \frac{\partial^2 u}{\partial x^2}(r \cos \theta, r \sin \theta) \cos \theta + \frac{\partial^2 u}{\partial y \partial x}(r \cos \theta, r \sin \theta) \sin \theta \right) \cos \theta \]
\[ + \left( \frac{\partial^2 u}{\partial x \partial y}(r \cos \theta, r \sin \theta) \cos \theta + \frac{\partial^2 u}{\partial y^2}(r \cos \theta, r \sin \theta) \sin \theta \right) \sin \theta \]
\[ = \frac{\partial^2 u}{\partial x^2}(r \cos \theta, r \sin \theta) \cos^2 \theta \]
\[ + 2 \frac{\partial^2 u}{\partial x \partial y}(r \cos \theta, r \sin \theta) \cos \theta \sin \theta \]
\[ + \frac{\partial^2 u}{\partial y^2}(r \cos \theta, r \sin \theta) \sin^2 \theta. \]
as well as
\[ \frac{\partial^2 u}{\partial \theta^2}(r \cos \theta, r \sin \theta) = \frac{\partial}{\partial \theta} \left( - \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta)r \sin \theta + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta)r \cos \theta \right) \]
\[ = - \frac{\partial}{\partial \theta} \left( \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) \right) r \sin \theta - \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) \frac{\partial}{\partial \theta} r \sin \theta \]
\[ + \frac{\partial}{\partial \theta} \left( \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) \right) r \cos \theta + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) \frac{\partial}{\partial \theta} r \cos \theta \]
\[ = - \left( - \frac{\partial^2 u}{\partial x^2}(r \cos \theta, r \sin \theta)r \sin \theta + \frac{\partial^2 u}{\partial y \partial x}(r \cos \theta, r \sin \theta)r \cos \theta \right) r \sin \theta \]
\[ - \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta)r \cos \theta \]
\[ + \left( - \frac{\partial^2 u}{\partial x \partial y}(r \cos \theta, r \sin \theta) r \sin \theta + \frac{\partial^2 u}{\partial y^2}(r \cos \theta, r \sin \theta) r \cos \theta \right) r \cos \theta \]
\[ - \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) r \sin \theta \]
\[ = - \left( \frac{\partial u}{\partial x}(r \cos \theta, r \sin \theta) r \cos \theta + \frac{\partial u}{\partial y}(r \cos \theta, r \sin \theta) r \sin \theta \right) \]
\[ + \frac{\partial^2 u}{\partial x^2}(r \cos \theta, r \sin \theta) r^2 \sin^2 \theta - 2 \frac{\partial^2 u}{\partial x \partial y}(r \cos \theta, r \sin \theta) r^2 \sin \theta \cos \theta \]
\[ + \frac{\partial^2 u}{\partial y^2}(r \cos \theta, r \sin \theta) r^2 \cos^2 \theta. \]

Putting it all together we get
\[ \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right)(r \cos \theta, r \sin \theta) = \left( \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial \theta^2} \right) u(r \cos \theta, r \sin \theta). \]

So, for example, if
\[ u(x, y) = \log \sqrt{x^2 + y^2} \quad \text{for} \quad (x, y) \neq (0, 0) \]
in which case
\[ u(r \cos \theta, r \sin \theta) = \log r \quad \text{for} \quad r \neq 0 \]
or if
\[ u(x, y) = \arcsin \frac{y}{\sqrt{x^2 + y^2}} \quad \text{for} \quad (x, y) \in \mathbb{R}^2 \text{ such that } x > 0 \text{ if } y = 0 \]
in which case
\[ u(r \cos \theta, r \sin \theta) = \theta \quad \text{if} \quad r > 0 \text{ and } -\pi < \theta < \pi \]
then
\[ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \]

The operator
\[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \]
is called the Laplacian: it is, by far, the most important partial differential operator in mathematics and physics. (Typically, an operator is a function whose domain is a set of functions and whose range is a set of functions.) To define it the way we
did above turns out to be not so good an idea; there are much better definitions but they require a bit of machinery which we will develop later.