1. Differentiation

1.1. **Differentiability.** Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$ be a function. f is differentiable at $x \in U$ if there exists a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{y \to x} \frac{||f(y) - f(x) - A(y - x)||}{||y - x||} = 0 \ .$$

Intuitively, if we denote $\Delta x = y - x$ (change in x) and $\Delta f(x) = f(y) - f(x)$ (change in f), then $A\Delta x$ is a good approximation of $\Delta f(x)$ (the best possible by a linear function).

$$\Delta f(x) \approx A\Delta x$$

The linear transformation A is unique and we denote it by $\mathrm{D}\, f(x)$. This is the derivative of f at x.

Note that:

- Unlike the single variable case the derivative can not be expressed as a limit. It is a linear transformation which approximates change in the function in terms of change in the argument.
- If f is differentiable at x, then it is continuous at x.

1.2. **Partial derivatives.** Let $f : U \to \mathbb{R}$ where $U \subset \mathbb{R}^n$ open. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$ (e_i is the vector whose i-th coordinate is 1 and all other coordinates are 0). Then the i-th partial derivative of f at x is defined as

$$\partial_i f(x) := \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t} \ .$$

The partial derivative exists if the above limit exists.

If $x = (x_1, \ldots, x_n)$, then there is another commonly used notation for partial derivatives.

$$\frac{\partial f}{\partial x_i}(x) = \partial_i f(x)$$

Here we are converting the function f to a single variable function by keeping all the coordinates of the argument fixed except the i-th coordinate and then taking the derivative with respect to the i-th coordinate.

If $f : U \to \mathbb{R}^m$ is differentiable at $a \in U$ and $f(x) = (f_1(x), \ldots, f_m(x))$, then all the partial derivatives $\partial_i f_j(a)$ exist. More over the matrix of the linear transformation $\mathrm{D}\, f(a)$ also denoted by $\mathrm{D}\, f(a)$ is given by

$$\mathrm{D}\, f(a)_{i,j} = \partial_j f_i(a) \ .$$

Thus the derivative matrix of f at a is an $m \times n$ matrix whose i-th row consists of all the partials of f_i at a and the j-th column consists of the j-th partials of $f_1, \ldots, f_m$ at a.

Theorem 1. *If all the partials $\partial_i f_j$ exist and are continuous in an open neighbourhood of a, then f is differentiable at a. Such a function is called continuously differentiable at a.*

Note that:

- Existence of partial derivatives $\partial_i f_j(a)$ does not guarantee differentiability of f at a.

- If f is differentiable at a then all the partials of the coordinate functions exist at a, but this does not guarantee they exist is some open neighbourhood of a.

1.3. **Directional derivatives.** Let $f : U \to \mathbb{R}$ where $U \subset \mathbb{R}^n$ is open. For any vector $u \in \mathbb{R}^n$ the directional derivative of f along u at a is defined as

$$\partial_u f(a) = \lim_{t \to 0} \frac{f(a+tu) - f(a)}{t} .$$

The directional derivative exists if the above limit exists.

If f is differentiable at a then all the directional derivatives of f exist at a and

$$\partial_u f(a) = \mathrm{D}\, f(a) u$$

$\mathrm{D}\, f(a)$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$ and $u \in \mathbb{R}^n$ so $\mathrm{D}\, f(a)u$ is a number.

Note that:

- Even if all the directional derivatives of a function exist at a point the function may not be differentiable at that point.
- Even if a function is differentiable at a point, all its directional derivatives will not be equal at that point.

1.4. **Higher derivatives.** Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$. Suppose f is continuously differentiable on U that is the functions $\partial_i f_j : U \to \mathbb{R}$ exist and are continuous on all of U.

Then we can take partial derivatives of $\partial_i f_j$ at a point in U if they exists. The derivatives $\partial_i \partial_j f_k(x)$ are called the second order derivatives of f_k. Similarly if these functions are differentiable on U we can take the third order derivatives and so on.

We say the function is $\mathcal{C}^1$ if it is continuously differentiable. It is called $\mathcal{C}^2$ if all the second order derivatives of the coordinate functions exist and are continuous. It is $\mathcal{C}^k$ if all the k-th order derivatives of the coordinate functions exist and are continuous.

All the k-th order derivatives of all the coordinate functions of f collectively are called the higher derivatives of f. If all the higher derivatives of f exist and are continuous then the function is called $\mathcal{C}^\infty$.

Continuous functions on the other hand are denoted by $\mathcal{C}^0$.

We shall also denote by $\mathcal{C}^k$ the set of all $\mathcal{C}^k$ functions. It is easy to check that $\mathcal{C}^k$ functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a vector space.

We have the following inclusions $\mathcal{C}^0 \supset \mathcal{C}^1 \supset \mathcal{C}^2 \supset \ldots \supset \mathcal{C}^\infty$.

Theorem 2. *Let $f : U \to \mathbb{R}$ where $U \subset \mathbb{R}^n$ is open and $\partial_i \partial_j f$ and $\partial_j \partial_i f$ exist and are continuous in an open neighbourhood of $a \in U$, then*

$$\partial_i \partial_j f(a) = \partial_j \partial_i f(a)$$

Similar result holds for other higher derivatives regarding changing the order of differentiation.

In different set of notations the higher derivatives are written as follows.

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \partial_i \partial_j f(a) \qquad \text{and} \qquad \frac{\partial^3 f}{\partial x_i \partial x_j^2}(a) = \partial_i \partial_j \partial_j f(a)$$

1.5. **Chain rule.** Let $h : U \to \mathbb{R}^m$ and $g : V \to \mathbb{R}^p$, where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open sets. Then $f = g \circ h$ is defined on $g^{-1}(V) \cap U$. Let $a \in g^{-1}(V) \cap U$.

Theorem 3 (Chain rule)**.** *Suppose h is continuously differentiable in an open neighbourhood of a, and g is differentiable at $h(a)$. Then $g \circ h$ is differentiable at a and*

$$\mathrm{D}(g \circ h)(a) = \mathrm{D}\, g(h(a))\, \mathrm{D}\, h(a) \ .$$

Let $f = g \circ h$ and $b = h(a)$, then in terms of the partial derivatives this translates to the following.

$$\partial_i f_j(a) = \sum_{k=1}^{m} \partial_k g_j(b) \partial_i h_k(a)$$

Written differently yet if $z_j = f_j(x)$ and $y_k = h_k(x)$ then

$$\frac{\partial z_j}{\partial x_i} = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k} \frac{\partial y_k}{\partial x_i} \ .$$

In this last notation you have to be careful about where the derivatives are taken.

1.6. **Inverse and Implicit function theorems.**

Theorem 4 (Inverse function theorem)**.** *Let $U \subset \mathbb{R}^n$ be some open set and $f : U \to \mathbb{R}^n$ be a $\mathcal{C}^1$ function such that $\mathrm{D}\, f(a)$ is invertible for some $a \in U$. Then there are open sets $V \subset U$, $a \in V$ and $W \subset \mathbb{R}^n$, $f(a) \in W$ such that*

- *f maps V bijectively onto W;*
- *the inverse function $f^{-1} : W \to \mathbb{R}^m$ is $\mathcal{C}^1$; and*
- *$\mathrm{D}\, f^{-1}(x) = \left[\mathrm{D}\, f\left(f^{-1}(x)\right)\right]^{-1}$ for all $x \in W$.*

In fact if the function f is $\mathcal{C}^k$ then the inverse function is also $\mathcal{C}^k$.

Theorem 5 (Implicit function theorem)**.** *Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ be $\mathcal{C}^1$ and $(a, b) \in \mathbb{R}^n \times \mathbb{R}^m$ such that the sub-matrix of $\mathrm{D}\, f(a, b)$ consisting of the last m columns is an invertible $m \times m$ matrix, then there are open sets $A \subset \mathbb{R}^n$ containing a and $B \subset \mathbb{R}^m$ containing b such that for every $x \in A$ there is a unique $g(x) \in B$ satisfying $f(x, g(x)) = f(a, b)$. The function $g : A \to B$ is $\mathcal{C}^1$.*

Again as before if f is $\mathcal{C}^k$ the function g is also C^k.

1.7. **Immersions and Submersions.** Let $U \subset \mathbb{R}^n$ be an open set.

If $n \leq m$ a function $f : U \to \mathbb{R}^m$ is called an **immersion** if it is $\mathcal{C}^1$ and the linear transformation $\mathrm{D}\, f(x)$ is injective for all $x \in U$, i.e. $\mathrm{D}\, f(x)$ has rank n (full rank).

We showed in class that for any $a \in U$ there is an open set $V \subset U$ containing a and an open set $W \subset \mathbb{R}^m$ containing $f(a)$ and a $\mathcal{C}^1$ function $g : W \to \mathbb{R}^n$ which satisfies

$$g \circ f(x) = x \text{ for all } x \in V \ .$$

It follows that f is injective on V. So any immersion is locally injective.

If $n \geq m$ a function $f : U \to \mathbb{R}^m$ is called a **submersion** if it is $\mathcal{C}^1$ and the linear transformation $\mathrm{D}\, f(x)$ is surjective for all $x \in U$, i.e. $\mathrm{D}\, f(x)$ has rank m (full rank).

We showed that for any $a \in U$, there are open sets $V \subset U$ containing a and $W \subset \mathbb{R}^m$ containing $f(a)$ and a $\mathcal{C}^1$ function $g : W \to \mathbb{R}^n$ which satisfies

$$f \circ g(y) = y \text{ for all } y \in W \ .$$

It follows that f is an open map.

Note that:

- Immersions may not be injective globally.
- Submersions may bot be surjective globally.

2. Integration

2.1. **Riemann Integral.** A closed rectangle $A \subset \mathbb{R}^n$ is a product of closed intervals $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ where $a_i < b_i$. We define the volume of such a rectangle as

$$v(A) = (b_1 - a_1) \cdots (b_n - a_n) \ .$$

An open rectangle $B \subset \mathbb{R}^n$ is a product of open intervals $A = (a_1, b_1) \times \cdots \times (a_n, b_n)$ where $a_i < b_i$. We define the volume of such a rectangle also as

$$v(B) = (b_1 - a_1) \cdots (b_n - a_n) \ .$$

If $f : A \to \mathbb{R}$ is a bounded function then define

$$m_A(f) = \inf\{f(x) \mid x \in A\} \qquad \text{and} \qquad M_A(f) = \sup\{f(x) \mid x \in A\} \ .$$

Recall that a partition of A is a collection of partitions $\{P_i\}_{i=1}^n$ of the intervals $[a_i, b_i]$, $i = 1, \ldots, n$. A partition divides A into smaller sub-rectangles. The partition P will also denote the set of sub-rectangles determined by P.

If $f : A \to \mathbb{R}$ is bounded and P a partition of A define the lower and upper sums of f over P by

$$L(f, P) = \sum_{R \in P} m_R(f) v(R) \qquad \text{and} \qquad U(f, P) = \sum_{R \in P} M_R(f) v(R)$$

f is integrable on A if $\sup_P L(f, P) = \inf_P U(f, P)$. In that case we define the integral of f on A to be

$$\int_A f = \sup_P L(f, P) = \inf_P U(f, P) \ .$$

In any case let $\underline{\int_A} f = \sup_P L(f, P)$ be the lower integral and $\overline{\int_A} f = \inf_P U(f, P)$ be the upper integral. Clearly

$$\underline{\int_A} f \leq \overline{\int_A} f$$

Theorem 6. *Let A be a closed rectangle in $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ a bounded function, then f is integrable on A if and only if for any $\epsilon > 0$ there is a partition P of A such that $U(f, P) - L(f, P) < \epsilon$.*

If $S \subset \mathbb{R}^n$ is a bounded set then define $\chi_S : \mathbb{R}^n \to \mathbb{R}$ the characteristic (or indicator) function of S by

$$\chi_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

A bounded function $f : S \to \mathbb{R}$ is integrable on S if for some closed rectangle A containing S, $f\chi_S$ is integrable on A. In that case define

$$\int_S f = \int_A f\chi_S \ .$$

Similarly define the lower and upper integrals over S.

2.2. **Measure zero sets.** $S \subset \mathbb{R}^n$ has measure 0 if for every $\epsilon > 0$, there are countably many closed rectangles $\{R_i\}_{i=1}^{\infty}$ such that $S \subset \cup_{i=1}^{\infty} R_i$ and $\sum_{i=1}^{\infty} v(R_i) < \epsilon$.

Theorem 7. *$S \subset \mathbb{R}^n$ has measure 0 if and only if for every $\epsilon > 0$, there are countably many open rectangles $\{R_i\}_{i=1}^{\infty}$ such that $S \subset \cup_{i=1}^{\infty} R_i$ and $\sum_{i=1}^{\infty} v(R_i) < \epsilon$.*

Theorem 8. *(1) If $S \subset T \subset \mathbb{R}^n$ and T has measure 0, then S also has measure 0.*
(2) If $S \subset \mathbb{R}^n$ contains a closed rectangle then S does not have measure 0.
(3) Union of countably many measure 0 sets has measure 0.

It is a useful fact that any countable set has measure 0 where as a closed rectangle does not have measure 0.

Theorem 9 (Sard's theorem)**.** *Let $f : U \to \mathbb{R}^n$ be $\mathcal{C}^1$, where U is an open subset of $\mathbb{R}^n$. Define the set of critical points of f to be*

$$D = \{x \in U \mid \det \mathrm{D}\, f(x) = 0\} \ .$$

Then $f(D)$ has measure 0.

We have the following two corollaries.

Corollary 10. *If $f : U \to \mathbb{R}^m$ is $\mathcal{C}^1$, where U is an open subset of $\mathbb{R}^n$ and $n < m$, then $f(U)$ has measure 0.*

Corollary 11. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be $\mathcal{C}^1$ and $n \geq m$. For any $a \in \mathbb{R}^m$ define*

$$L_f(a) = \{x \in \mathbb{R}^n \mid f(x) = a\} = f^{-1}(\{a\}) \ .$$

Suppose $\mathrm{D}\, f(x) \neq 0$ for any $x \in L_f(a)$, then $L_f(a)$ has measure 0.

2.3. **Jordan domains.** A bounded set $S \subset \mathbb{R}^n$ is a Jordan domain if the constant function 1 is integrable on S. In that case the volume of S is

$$v(S) = \int_S 1$$

Theorem 12. *S is a Jordan domain if and only if $\mathrm{Bd}\, S$ the boundary of S has measure 0.*

If S_1 and S_2 are Jordan domains then $S_1 \cup S_2$, $S_1 \cap S_2$ and $S_1 \backslash S_2$ are also Jordan domains.

Jordan domains are important because of the following result.

Proposition 13. *A bounded set S is a Jordan domain if and only if any continuous bounded function $f : S \to \mathbb{R}$ is integrable on S.*

Note that:

- Not every measure 0 set is a Jordan domain.
- If a measure 0 set is a Jordan domain then its volume is 0.
- Not every bounded set which is closed (hence compact) or open is a Jordan domain.

2.4. **Evaluating integrals.** The following theorem is an essential tool in calculating integrals of functions.

Theorem 14. *(1) If $f, g : S \to \mathbb{R}$ are integrable functions on $S \subset \mathbb{R}^n$ bounded, and $a, b \in \mathbb{R}$ then $af + bg$ is also integrable on S and $\int_S af + bg = a \int_S f + b \int_S g$.*
(2) If $f, g : S \to \mathbb{R}$ are integrable functions on $S \subset \mathbb{R}^n$ bounded, and $f(x) \geq g(x)$ for all $x \in A$, then $\int_S f \geq \int_S g$. It follows that $|f|$ is also integrable on S and $\int_S |f| \geq |\int_S f|$.
(3) If $S \subset \mathbb{R}^n$ is bounded, $T \subset S$ and $f : S \to \mathbb{R}$ is a non-negative function integrable on S and T then $\int_T f \leq \int_S f$.

(4) Let S_1, S_2 be bounded subsets of $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ is integrable on both S_1 and S_2 then f is integrable on $S_1 \cup S_2$ and $S_1 \cap S_2$ and

$$\int_{S_1 \cup S_2} f = \int_{S_1} f + \int_{S_2} f - \int_{S_1 \cap S_2} f \ .$$

In single variable integration theory the following theorem is the main tool for calculating integrals.

Theorem 15 (Fundamental theorem of Calculus)**.** *If $f : [a, b] \to \mathbb{R}$ is continuous then:*

(1) $F : [a, b] \to \mathbb{R}$ defined by $F(x) = \int_a^x f$ is continuously differentiable on (a, b) and $F'(x) = f(x)$.
(2) If $G : [a, b] \to \mathbb{R}$ is continuous and $G'(x) = f(x)$ for $x \in (a, b)$ then $\int_a^b f = G(b) - G(a)$.

In higher dimensions Fubini's theorem gives us a recipe to calculate integrals as iterated integrals.

Theorem 16 (Fubini's theorem)**.** *Let $A \times B \subset \mathbb{R}^{m+n}$ be a closed rectangle where A and B are closed rectangles in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. Let $f : A \times B \to \mathbb{R}$ be integrable on $A \times B$. For each $x \in A$ let $g_x : B \to \mathbb{R}$ be the function $g_x(y) = f(x, y)$. Define the functions $\mathcal{L}, \mathcal{U} : A \to \mathbb{R}$ by*

$$\mathcal{L}(x) = \underline{\int_B} g_x \qquad \text{and} \qquad \mathcal{U}(x) = \overline{\int_B} g_x \ .$$

Then $\mathcal{L}$ and $\mathcal{U}$ are integrable on A and

$$\int_{A \times B} f = \int_A \mathcal{L} = \int_A \mathcal{U} \ .$$

Note that in case g_x is integrable on B for each $x \in A$, the upper and lower integrals are equal to the integral of g_x. In that case we write this result as

$$\int_{A \times B} f = \int_A \int_B f dy dx$$

where $x \in A$ and $y \in B$.

If $f : A \to \mathbb{R}$ is continuous and $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is a closed rectangle in $\mathbb{R}^n$, then

$$\int_A f = \int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f dx_1 \dots dx_n \ .$$

Of course we can integrate in any different order and the integral will be the same by the previous theorem.

2.5. **Improper integral on open sets.** As we saw not all open sets are Jordan domains, which means continuous functions may not be integrable on even bounded open sets. In order to extend the theory of Riemann integration to arbitrary open sets we shall use partitions of unity.

Theorem 17 (Partitions of unity)**.** *Given any open set $U \subset \mathbb{R}^n$ and $\{U_\alpha\}_{\alpha \in A}$ a collection of open sets such that $\cup_{\alpha \in A} U_\alpha = U$ we have a countable collection of $\mathcal{C}^\infty$ functions $\{\phi_1, \phi_2, \dots\}$ such that*

(1) $\phi_i : \mathbb{R}^n \to \mathbb{R}$ and $\phi_i(x) \geq 0$;
(2) for each ϕ_i there is some U_α such that $\phi_i = 0$ outside a compact subset of U_α;
(3) each $x \in U$ has an open neighbourhood on which all but finitely many ϕ_i are 0;
(4) $\sum_{i=1}^\infty \phi_i(x) = 1$ for each $x \in U$.

Such a collection $\{\phi_1, \phi_2, \dots\}$ is called a partition of unity subordinate to the open cover $\{U_\alpha\}_{\alpha \in A}$. Condition 3 is called local finiteness.

Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ be locally bounded. This means each $x \in U$ has an open neighbourhood on which f is bounded. Let further the set of discontinuities of f have measure 0.

Let $\{\phi_1, \phi_2, \ldots\}$ be a partition of unity for U subordinate to some open cover of U. Note that $|f|\phi_i$ is integrable on any closed rectangle. Choose a closed rectangle A_i outside which $\phi_i = 0$ (possible since ϕ_i vanishes outside a compact set). We say that f is integrable (in the extended sense) on U if the series $\sum_{i=1}^{\infty} \int_{A_i} |f|\phi_i$ converges.

Note that in this case
$$\sum_{i=1}^{\infty} \Big| \int_{A_i} f\phi_i \Big| \leq \sum_{i=1}^{\infty} \int_{A_i} |f|\phi_i$$
which shows that the series $\sum_{i=1}^{\infty} \int_{A_i} f\phi_i$ is absolutely convergent and we define
$$\int_U f = \sum_{i=1}^{\infty} \int_{A_i} f\phi_i \ .$$

It can be shown that this definition is independent of the choice of a partition of unity. If f and U are both bounded the integral exists, and if U is a Jordan domain and f is integrable on U in the usual sense and the two integrals are the same.

Whenever we deal with open sets we shall always consider this extended or improper integral.

There is another way of defining improper integrals for continuous functions which is more useful for computations. If $f : U \to \mathbb{R}$ is continuous and $f(x) \geq 0$ for all $x \in U$. We define
$$\int_U f = \sup\{\int_C f \mid C \subset U \text{ compact Jordan domain}\}$$
provided the supremum exists. For any continuous function $f : U \to \mathbb{R}$ define $f_+ = \max\{f, 0\}$ and $f_- = \max\{-f, 0\}$ we say f is integrable on U if both f_+, f_- are integrable on U and define
$$\int_U f = \int_U f_+ - \int_U f_- \ .$$

Both these definitions for improper integrals for continuous functions match and give the same integration value. Moreover we have the following result.

Theorem 18. *Suppose $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}$ is continuous. If $C_1, C_2, \ldots$ be compact subsets of U which are Jordan domains, $C_k \subset \operatorname{Int} C_{k+1}$ and $U = \cup_{k=1}^{\infty} C_k$ then f is integrable on U if and only if $\int_{C_k} |f|$ is a bounded sequence and in that case*
$$\int_U f = \lim_{k\to\infty} \int_{C_k} f \ .$$

2.6. **Change of variables.** Let $A \subset \mathbb{R}^n$ be open then a function $f : A \to \mathbb{R}^n$ is called a $\mathcal{C}^1$ diffeomorphism if f is $\mathcal{C}^1$ and injective and $\mathrm{D}\, f(x)$ is invertible for every $x \in A$. Note that this forces $f(A)$ to be an open set.

Theorem 19 (Change of variables)**.** *If $g : A \to \mathbb{R}^n$ is a $\mathcal{C}^1$ diffeomorphism and $f : g(A) \to \mathbb{R}$ is integrable then $(f \circ g)|\det \mathrm{D}\, g|$ is integrable on A and*
$$\int_{g(A)} f = \int_A (f \circ g)|\det \mathrm{D}\, g| \ .$$

In this theorem the condition that g is a diffeomorphism can be relaxed. In fact if it is only the case that g is $\mathcal{C}^1$ and injective and $g(A)$ is open then the statement of the theorem still holds (since the image of the set where $\det \mathrm{D}\, g$ vanishes has measure 0).

Note that: The above theorem fails when g is not injective.