

MATH 247 COURSE NOTES

CALCULUS 3 (ADVANCED LEVEL)

ALEXANDRU NICA • FALL 2019 • UNIVERSITY OF WATERLOO

Table of Contents

1	Ring of sets, additive set-function	3
2	Divisions in a ring of sets	5
3	Integrable functions	7
4	Operations with integrable functions	11
5	Our X, \mathcal{C}, μ of choice: $\mathbb{R}^n, \widehat{\mathcal{R}}_n, \text{vol}_n$	17
6	Integration on \mathbb{R}^n	19
7	Integration of continuous functions	21
8	Interior, closure, boundary for subsets of \mathbb{R}^n	25
9	Open and closed sets in \mathbb{R}^n	27
10	A new look at interior and closure	30
11	Compact subsets of \mathbb{R}^n	31
12	Continuous functions on compact subsets of \mathbb{R}^n	33
13	Jordan measurable subsets of \mathbb{R}^n	36
14	(UC-mod-E) upgraded to (C-mod-N)	40
15	Calculation of integrals and volumes by slicing	44
16	Directional derivatives and the MVT	47
17	C^1 -functions and their linear approximation	50
18	C^1 implies differentiable implies continuous	53
19	C^2 -functions and their quadratic approximation	55
20	Points of local extremum	58
21	Operations with C^1 -functions and the chain rule	61

22	Lagrange multipliers	64
23	Change of variable for multiple integrals	66
	Additional definitions from assignments	68

1 Ring of sets, additive set-function

Definition 1.1 (Ring of sets). Let \bar{X} be a set and let \mathcal{C} be a collection of subsets of \bar{X} . (Note that \mathcal{C} is a set of sets.) We say that \mathcal{C} is a **ring of subsets of \bar{X}** when it satisfies the following:

- (RS-1) $\emptyset \in \mathcal{C}$
- (RS-2) $A, B \in \mathcal{C} \implies A \cup B \in \mathcal{C}$
- (RS-3) $A, B \in \mathcal{C} \implies A \setminus B \in \mathcal{C}$

Remark 1.2. Let \bar{X} be a set and let \mathcal{C} be a ring of subsets of \bar{X} .

1. If $A, B \in \mathcal{C}$, then $A \cap B \in \mathcal{C}$.

Proof. Observe that $A = (A \cap B) \cup (A \setminus B)$, a disjoint union. Hence $A \cap B = A \setminus (A \setminus B)$. We have $A \setminus B \in \mathcal{C}$ by RS-3. But then

$$A \in \mathcal{C} \text{ and } A \setminus B \in \mathcal{C} \implies A \setminus (A \setminus B) \in \mathcal{C} \quad (\text{by RS-3})$$

Hence $A \cap B = A \setminus (A \setminus B) \in \mathcal{C}$, as required. \square

2. For every $k \geq 2$ and $A_1, \dots, A_k \in \mathcal{C}$, we have that $A_1 \cup \dots \cup A_k \in \mathcal{C}$ and $A_1 \cap \dots \cap A_k \in \mathcal{C}$.

Proof. Let $k = 2$. Then since $A_1, A_2 \in \mathcal{C}$, we have that $A_1 \cup A_2 \in \mathcal{C}$ by RS-2 and $A_1 \cap A_2 \in \mathcal{C}$ from what was proved above. Now suppose $A_1, \dots, A_{k+1} \in \mathcal{C}$, $A_1 \cup \dots \cup A_k \in \mathcal{C}$, and $A_1 \cap \dots \cap A_k \in \mathcal{C}$. Observe that since $A_{k+1} \in \mathcal{C}$, we once again have $A_1 \cup \dots \cup A_k \cup A_{k+1} \in \mathcal{C}$ by RS-2, and $A_1 \cap \dots \cap A_k \cap A_{k+1} \in \mathcal{C}$ from the result above. \square

Remark 1.3. Related terminology:

1. Let \bar{X} be a set and let \mathcal{C} be a ring of subsets of \bar{X} . If $\bar{X} \in \mathcal{C}$, then we say that \mathcal{C} is a **field of subsets of \bar{X}** , or that it is an **algebra of subsets of \bar{X}** . (However, the following examples will have $\bar{X} \notin \mathcal{C}$.)
2. One also has a notion of **semi-rings of subsets of \bar{X}** (this is on Homework 1).

Definition 1.4 (Additive set-function). Let \bar{X} be a set and let \mathcal{C} be a ring of subsets of \bar{X} . A function $\mu : \mathcal{C} \rightarrow [0, \infty)$ is said to be an **additive set-function** when it satisfies

$$(\text{Add}) \quad \begin{array}{l} A, B \in \mathcal{C} \\ A \cap B = \emptyset \end{array} \implies \mu(A \cup B) = \mu(A) + \mu(B)$$

Remark 1.5. Let \bar{X} be a set, \mathcal{C} be a ring of subsets of \bar{X} , and $\mu : \mathcal{C} \rightarrow [0, \infty)$ be an additive set-function. If $A_1, \dots, A_k \in \mathcal{C}$ and if $A_i \cap A_j = \emptyset$ whenever $i \neq j$, then it follows that

$$\mu(A_1 \cup \dots \cup A_k) = \sum_{i=1}^k \mu(A_i)$$

Proof. For $k = 1$, we can clearly see that

$$\mu(A_1) = \sum_{i=1}^1 \mu(A_i)$$

Suppose this is true for some k . Let $A_{k+1} \in \mathcal{C}$. We have that $A_1 \cup \dots \cup A_{k+1} \in \mathcal{C}$ and $A_i \cap A_{k+1} = \emptyset$ for any $i \neq k+1$, thus $(A_1 \cup \dots \cup A_k) \cap A_{k+1} = \emptyset$. By the definition of an additive set-function, we then have

$$\begin{aligned} \mu(A_1 \cup \dots \cup A_{k+1}) &= \mu(A_1 \cup \dots \cup A_k) + \mu(A_{k+1}) \\ &= \sum_{i=1}^k \mu(A_i) + \mu(A_{k+1}) \\ &= \sum_{i=1}^{k+1} \mu(A_i) \end{aligned}$$

as desired. □

Further consequences of (Add) (e.g. "inclusion-exclusion principle") will be on Homework 1.

Idea 1.6. The integration method of Riemann works (and is more transparent) in the framework of \bar{X} , \mathcal{C} , and μ .

Example 1.7. Let \bar{X} be an infinite set. Let $\mathcal{C} = \{A \subseteq \bar{X} \mid A \text{ is finite}\}$, and let $\mu : \mathcal{C} \rightarrow [0, \infty)$ be $\mu(A) = |A|$ (the number of elements of A), $A \in \mathcal{C}$.

Example 1.8. Let $\bar{X} = \mathbb{R}$ and let $\mathcal{C} = \{A \subseteq \mathbb{R} \mid \exists p \in \mathbb{N} \text{ and } a_1 < b_1 \leq a_2 < b_2 \leq \dots \leq a_p < b_p \in \mathbb{R} \text{ such that } A = [a_1, b_1) \cup [a_2, b_2) \cup \dots \cup [a_p, b_p)\} \cup \{\emptyset\}$. For A as above, define

$$\begin{aligned} \mu(A) &= \sum_{i=1}^p b_i - a_i && \text{("length" of } A) \\ \mu(\emptyset) &= 0 \end{aligned}$$

2 Divisions in a ring of sets

Definition 2.1. Let $\emptyset \neq A \in \mathcal{C}$. By **division** of A , we understand a set $\Delta = \{A_1, \dots, A_p\}$ where A_1, \dots, A_p are non-empty sets in \mathcal{C} such that $A_1 \cup \dots \cup A_p = A$ and $A_i \cap A_j = \emptyset$ for $i \neq j$ (they are pairwise disjoint).

Definition 2.2. Let $\emptyset \neq A \in \mathcal{C}$. Let $\Delta = \{A_1, \dots, A_p\}$ and $\Gamma = \{B_1, \dots, B_q\}$ be divisions of A . We say that Γ **refines** Δ , written as " $\Gamma \prec \Delta$ ", to mean that for every $1 \leq j \leq q$, there exists $1 \leq i \leq p$ such that $B_j \subseteq A_i$. (This can be thought of as " Γ has smaller pieces".)

Exercise 2.3. Let $\Delta = \{A_1, \dots, A_p\}$ and $\Gamma = \{B_1, \dots, B_q\}$ with $\Gamma \prec \Delta$ as in Definition 2.2. Prove that we can rewrite $\Gamma = \{B_{1,1}, \dots, B_{1,q_1}, B_{2,1}, \dots, B_{2,q_2}, \dots, B_{p,1}, \dots, B_{p,q_p}\}$ (Note: $\sum_{i=1}^p q_i = q$) such that $\{B_{1,1}, \dots, B_{1,q_1}\}$ is a division of A_1 , \dots , $\{B_{p,1}, \dots, B_{p,q_p}\}$ is a division of A_p .

Proof. Consider the set A_i in Δ , and let $B_{i,1}, \dots, B_{i,q_i}$ be the sets that are included in A_i . Then

$$B_{1,1}, \dots, B_{1,q_1}, \dots, B_{p,1}, \dots, B_{p,q_p}$$

is a complete list of sets for Γ , since Γ has the property given in Definition 2.2.

Now, we must verify that $B_{i,1}, \dots, B_{i,q_i}$ is indeed a division of A_i . These sets are clearly pairwise disjoint, so we only need to verify that $A_i \subseteq B_{i,1} \cup \dots \cup B_{i,q_i}$.

Suppose for the sake of contradiction that the inclusion $A_i \subseteq B_{i,1} \cup \dots \cup B_{i,q_i}$ does not hold. Then there exists a point $x \in A_i \setminus (B_{i,1} \cup \dots \cup B_{i,q_i})$.

Since Γ is a division of A , there exists $1 \leq j_0 \leq q$ such that $x \in B_{j_0}$. Moreover, since Γ and Δ satisfy Definition 2.2, there exists $1 \leq i_0 \leq p$ such that $B_{j_0} \subseteq A_{i_0}$. We consider two possible cases for i_0 .

CASE 1. $i_0 = i$.

In this case, $B_{j_0} \subseteq A_i$, hence B_{j_0} is one of the sets $B_{i,1}, \dots, B_{i,q_i}$. But then $x \in B_{j_0}$ means that $x \in B_{i,1} \cup \dots \cup B_{i,q_i}$, which is a contradiction with $x \in A_i \setminus (B_{i,1} \cup \dots \cup B_{i,q_i})$.

CASE 2. $i_0 \neq i$.

In this case, from $x \in B_{j_0} \subseteq A_{i_0}$, it follows that $A_i \cap A_{i_0} \neq \emptyset$ (since the point x is in the intersection). This contradicts the fact that the sets A_1, \dots, A_p are pairwise disjoint.

Both these cases lead to a contradiction, and thus the inclusion $A_i \subseteq B_{i,1} \cup \dots \cup B_{i,q_i}$ holds. \square

Proposition 2.4. Let $\emptyset \neq A \in \mathcal{C}$. Let Δ' and Δ'' be any two divisions of A . Then there exists a division Γ of A such that $\Gamma \prec \Delta'$ and $\Gamma \prec \Delta''$ (a common refinement).

Proof. Write $\Delta' = \{A'_1, \dots, A'_p\}$ and $\Delta'' = \{A''_1, \dots, A''_q\}$. Put

$$\Gamma := \{A'_i \cap A''_j \mid 1 \leq i \leq p, 1 \leq j \leq q, \text{ and } A'_i \cap A''_j \neq \emptyset\}$$

Note: From Remark 1.2.1, $A'_i \cap A''_j \in \mathcal{C}$.

CLAIM 1. Γ is a division of A .

VERIFICATION OF CLAIM 1. Γ is made of non-empty sets of \mathcal{C} . We must check two things.

- The union of the sets from Γ is equal to A .

To check this, we have

$$\begin{aligned}
 \bigcup_{\substack{1 \leq i \leq p, 1 \leq j \leq q \\ \text{such that } A'_i \cap A''_j \neq \emptyset}} (A'_i \cap A''_j) &= \bigcup_{\substack{1 \leq i \leq p, \\ 1 \leq j \leq q}} (A'_i \cap A''_j) \\
 &= \left(\bigcup_{i=1}^p A'_i \right) \cap \left(\bigcup_{j=1}^q A''_j \right) \\
 &= A \cap A = A
 \end{aligned}$$

The first step comes from adding in copies of \emptyset to the union, causing no change.

The second step comes from the distributivity of intersection with respect to the union.

- The sets in Γ are pairwise disjoint.

To check this, let $B = A'_i \cap A''_j$, $C = A'_k \cap A''_\ell$ be in Γ where $B \neq C$. We want to prove that $B \cap C = \emptyset$. Since $B \neq C$, we have that either $i \neq k$ or $j \neq \ell$. We have

$$\begin{aligned}
 B \cap C &= (A'_i \cap A''_j) \cap (A'_k \cap A''_\ell) \\
 &= (A'_i \cap A'_k) \cap (A''_j \cap A''_\ell)
 \end{aligned}$$

If $i \neq k$, then $A'_i \cap A'_k = \emptyset$. Similarly, if $j \neq \ell$, then $A''_j \cap A''_\ell = \emptyset$. Hence we have $B \cap C = \emptyset$ (since one of the two is \emptyset).

This proves Claim 1.

CLAIM 2. We have $\Gamma \prec \Delta'$ and $\Gamma \prec \Delta''$.

VERIFICATION OF CLAIM 2. For every $B = A'_i \cap A''_j \in \Gamma$, we have $B \subseteq A'_i \in \Delta'$. Hence $\Gamma \prec \Delta'$. Likewise (by using $A'_i \cap A''_j \subseteq A''_j$) we find that $\Gamma \prec \Delta''$.

This proves Claim 2, and we are done. □

3 Integrable functions

For this section, we fix a set X , a ring of subsets \mathcal{C} of X , and an additive set-function $\mu : \mathcal{C} \rightarrow [0, \infty)$.

Definition 3.1 (Upper and lower sums). Let A be a non-empty subset in \mathcal{C} and let $f : A \rightarrow \mathbb{R}$ be a bounded function. Let $\Delta = \{A_1, \dots, A_p\}$ be a division of A .

The number

$$U(f, \Delta) := \sum_{i=1}^p \mu(A_i) \cdot \sup_{A_i}(f)$$

is called the **upper sum** (sometimes called "upper Darboux sum") for f and Δ , and the number

$$L(f, \Delta) := \sum_{i=1}^p \mu(A_i) \cdot \inf_{A_i}(f)$$

is called the **lower sum** (sometimes called "lower Darboux sum") for f and Δ .

Remark 3.2. Let $A \in \mathcal{C}$, $f : A \rightarrow \mathbb{R}$, and the division Δ be as in Definition 3.1. It is obvious that we have $L(f, \Delta) \leq U(f, \Delta)$. Moreover, we can write an explicit formula for the difference of these sums:

$$U(f, \Delta) - L(f, \Delta) = \sum_{i=1}^p \mu(A_i) \cdot \operatorname{osc}_{A_i}(f).$$

Lemma 3.3. Let A be a non-empty set in \mathcal{C} and let $f : A \rightarrow \mathbb{R}$ be a bounded function. Let Δ, Γ be two divisions of A such that Γ refines Δ . Then we have

$$U(f, \Gamma) \leq U(f, \Delta) \text{ and } L(f, \Gamma) \geq L(f, \Delta)$$

As a consequence of the above, we also have that

$$U(f, \Gamma) - L(f, \Gamma) \leq U(f, \Delta) - L(f, \Delta).$$

Proof. We will prove the inequality for upper sums (the one for lower sums is analogous).

Write (as in Exercise 2.3) $\Delta = \{A_1, \dots, A_p\}$ and $\Gamma = \{B_{1,1}, \dots, B_{1,q_1}, \dots, B_{p,1}, \dots, B_{p,q_p}\}$ with $B_{i,1} \cup \dots \cup B_{i,q_i} = A_i$.

Then we have

$$U(f, \Gamma) = \sum_{i=1}^p \left(\sum_{j=1}^{q_i} \mu(B_{i,j}) \cdot \sup_{B_{i,j}}(f) \right)$$

For every $1 \leq i \leq p$ and $1 \leq j \leq q_i$, note that $B_{i,j} \subseteq A_i$. This implies $\sup_{B_{i,j}}(f) \leq \sup_{A_i}(f)$. Then,

$$\begin{aligned} U(f, \Gamma) &\leq \sum_{i=1}^p \left(\sum_{j=1}^{q_i} \mu(B_{i,j}) \cdot \sup_{A_i}(f) \right) \\ &= \sum_{i=1}^p \left[\left(\sum_{j=1}^{q_i} \mu(B_{i,j}) \right) \cdot \sup_{A_i}(f) \right] \\ &= \sum_{i=1}^p \mu(A_i) \cdot \sup_{A_i}(f) = U(f, \Delta) \end{aligned}$$

as required. □

Proposition 3.4. Let A be a non-empty set in \mathcal{C} , let $f : A \rightarrow \mathbb{R}$ be a bounded function, and let Δ', Δ'' be any two divisions of A . Then

$$U(f, \Delta') \geq L(f, \Delta'').$$

Proof. Let Γ be a common refinement of Δ' and Δ'' . Then by Remark 3.2 and Lemma 3.3, we have

$$L(f, \Delta'') \leq L(f, \Gamma) \leq U(f, \Gamma) \leq U(f, \Delta')$$

which was to be demonstrated. \square

Proposition 3.5. Let A be a non-empty set in \mathcal{C} and let $f : A \rightarrow \mathbb{R}$ be a bounded function.

1. The set of real numbers $T := \{U(f, \Delta) \mid \Delta \text{ is a division of } A\}$ is bounded from below. The number $\inf(T)$ is called the **upper integral of f on A** and is denoted as $\bar{\int}_A f$.
2. The set of real numbers $S := \{L(f, \Delta) \mid \Delta \text{ is a division of } A\}$ is bounded from above. The number $\sup(S)$ is called the **lower integral of f on A** and is denoted as $\underline{\int}_A f$.
3. We have $\underline{\int}_A f \leq \bar{\int}_A f$.

Proof.

STEP 1. Fix a division Δ'' of A . Then $L(f, \Delta'')$ is a lower bound for $T := \{U(f, \Delta') \mid \Delta' \text{ division of } A\}$ (by Proposition 3.4). Hence T is bounded from below by $\inf(T) \geq L(f, \Delta'')$.

CONCLUSION OF STEP 1. Thus, $\bar{\int}_A f := \inf(T) \in \mathbb{R}$ is well defined, and we have $\bar{\int}_A f \geq L(f, \Delta'')$ for any division Δ'' of A .

STEP 2. From Step 1, we get that $\bar{\int}_A f$ is an upper bound for the set $S := \{L(f, \Delta'') \mid \Delta'' \text{ division of } A\}$. Hence S is bounded from above, with $\sup(S) \leq \bar{\int}_A f$.

CONCLUSION OF STEP 2. $\underline{\int}_A f := \sup(S)$ is well defined, and $\underline{\int}_A f \leq \bar{\int}_A f$ as required. \square

Definition 3.6. Let A be a non-empty set in \mathcal{C} and let $f : A \rightarrow \mathbb{R}$ be a bounded function. We say that f is **integrable on A** to mean that one has $\underline{\int}_A f = \bar{\int}_A f$.

If f is integrable, then the common value of $\underline{\int}_A f$ and $\bar{\int}_A f$ is called the **integral of f on A** , denoted as $\int_A f$ (sometimes also denoted as $\int_A f(x) dx$, or as $\int_A f(x) d\mu(x)$ if there is a possibility of ambiguity on what our additive set-function μ is on \mathcal{C}).

Theorem 3.7. Let A be a non-empty set in \mathcal{C} and let $f : A \rightarrow \mathbb{R}$ be a bounded function. The following three statements are equivalent.

- (1) f is integrable as in Definition 3.6.
- (2) For every $\varepsilon > 0$, there exists a division Δ of A such that $U(f, \Delta) - L(f, \Delta) < \varepsilon$.
- (3) There exists a sequence $(\Delta_k)_{k=1}^\infty$ of divisions of A such that

$$\lim_{k \rightarrow \infty} U(f, \Delta_k) - L(f, \Delta_k) = 0.$$

Proof.

(1) \Rightarrow (2). Suppose f is integrable, that is, $\int_A f = \bar{\int}_A f =: I \in \mathbb{R}$. Given $\varepsilon > 0$, we want to find a division Δ of A such that $U(f, \Delta) - L(f, \Delta) < \varepsilon$.

CLAIM 1. There exists a division Δ' of A such that $I \leq U(f, \Delta') < I + \frac{\varepsilon}{2}$.

VERIFICATION OF CLAIM 1. We have $I = \bar{\int}_A f = \inf\{U(f, \Delta) \mid \Delta \text{ division of } A\}$. In particular, $I \leq U(f, \Delta)$ for all divisions Δ of A .

Assume by contradiction that $U(f, \Delta) \geq I + \frac{\varepsilon}{2}$ for all Δ . Then it follows that

$$I + \frac{\varepsilon}{2} \leq \inf\{U(f, \Delta) \mid \Delta \text{ division of } A\} = I,$$

a contradiction. So there must exist a division Δ' such that $I \leq U(f, \Delta') < I + \frac{\varepsilon}{2}$. This Δ' satisfies Claim 1.

CLAIM 2. There exists a division Δ'' of A such that $I \geq L(f, \Delta'') > I - \frac{\varepsilon}{2}$.

VERIFICATION OF CLAIM 2. This is analogous to the verification of Claim 1, where in this case we begin with $I = \int_A f = \sup\{L(f, \Delta) \mid \Delta \text{ division of } A\}$.

Let Δ be a common refinement of Δ' from Claim 1 and Δ'' from Claim 2. Then we have

$$I - \frac{\varepsilon}{2} < L(f, \Delta'') \leq L(f, \Delta) \leq U(f, \Delta) \leq U(f, \Delta') < I + \frac{\varepsilon}{2}$$

From these inequalities, we can see that

$$U(f, \Delta) - L(f, \Delta) < \left(I + \frac{\varepsilon}{2}\right) - \left(I - \frac{\varepsilon}{2}\right) = \varepsilon$$

which proves the first implication.

(2) \Rightarrow (3). We know for every $\varepsilon > 0$, there exists a division Δ of A such that $U(f, \Delta) - L(f, \Delta) < \varepsilon$.

Define $\varepsilon_k = 1/k$ and observe that $(\varepsilon_k)_{k=1}^\infty$ is a sequence of positive reals that converges to 0.

For each ε_k , there exists a division Δ_k of A such that $U(f, \Delta_k) - L(f, \Delta_k) < \varepsilon_k$. So there exists a sequence $(\Delta_k)_{k=1}^\infty$ of divisions of A , where

$$\lim_{k \rightarrow \infty} U(f, \Delta_k) - L(f, \Delta_k) = 0$$

which is what we wanted to show.

(3) \Rightarrow (1). Suppose there exists a sequence $(\Delta_k)_{k=1}^\infty$ of divisions of A such that

$$\lim_{k \rightarrow \infty} U(f, \Delta_k) - L(f, \Delta_k) = 0$$

We know that for all $k \in \mathbb{N}$,

$$U(f, \Delta_k) \geq \int_A f \geq \underline{\int}_A f \geq L(f, \Delta_k)$$

Hence

$$\left| \int_A f - \underline{\int}_A f \right| \leq U(f, \Delta_k) - L(f, \Delta_k) \xrightarrow{k \rightarrow \infty} 0$$

so it follows that $\bar{\int}_A f = \underline{\int}_A f$. Thus, f is integrable, as required. \square

Definition 3.8. Let A be a non-empty set in \mathcal{C} and let $f : A \rightarrow \mathbb{R}$ be a bounded function. Suppose that f is integrable on A . A sequence of divisions $(\Delta_k)_{k=1}^\infty$ of A such that

$$\lim_{k \rightarrow \infty} U(f, \Delta_k) - L(f, \Delta_k) = 0$$

as in (3) of Theorem 3.7 will be called a **witnessing sequence of divisions** for the integrability of f on A .

Proposition 3.9. Let A be a non-empty set in \mathcal{C} , let $f : A \rightarrow \mathbb{R}$ be a bounded integrable function, and let $(\Delta_k)_{k=1}^\infty$ be a witnessing sequence of divisions for the integrability of f on A . Then one has

$$\lim_{k \rightarrow \infty} U(f, \Delta_k) = \int_A f = \lim_{k \rightarrow \infty} L(f, \Delta_k).$$

Proof. We have $U(f, \Delta_k) \geq \bar{\int}_A f = \int_A f = \underline{\int}_A f \geq L(f, \Delta_k)$ for all $k \in \mathbb{N}$. Hence

$$\left| U(f, \Delta_k) - \int_A f \right| \leq U(f, \Delta_k) - L(f, \Delta_k) \xrightarrow{k \rightarrow \infty} 0.$$

It follows that $\lim_{k \rightarrow \infty} U(f, \Delta_k) = \int_A f$ by Squeeze Theorem.

Similarly, we get that $\lim_{k \rightarrow \infty} L(f, \Delta_k) = \int_A f$. □

4 Operations with integrable functions

Throughout this lecture, we fix a set X , a ring \mathcal{C} of subsets of X , and an additive set-function $\mu : \mathcal{C} \rightarrow [0, \infty)$. We also fix for the whole lecture a non-empty set $A \in \mathcal{C}$. We will discuss operations with integrable functions on this set A .

Notation 4.1. We set $F_b(A, \mathbb{R}) := \{f : A \rightarrow \mathbb{R} \mid f \text{ is bounded}\}$.

Remark 4.2. We observe several "pointwise" operations that can be done with functions in $F_b(A, \mathbb{R})$.

For $f, g \in F_b(A, \mathbb{R})$, we define the new functions

$$f + g, f \cdot g, f \vee g, f \wedge g \in F_b(A, \mathbb{R})$$

The formulas defining these new functions are given by

$$\begin{aligned} f + g : A &\rightarrow \mathbb{R} & (f + g)(x) &= f(x) + g(x) \text{ for all } x \in A; \\ f \cdot g : A &\rightarrow \mathbb{R} & (f \cdot g)(x) &= f(x) \cdot g(x) \text{ for all } x \in A; \\ f \vee g : A &\rightarrow \mathbb{R} & (f \vee g)(x) &= \max(f(x), g(x)) \text{ for all } x \in A; \\ f \wedge g : A &\rightarrow \mathbb{R} & (f \wedge g)(x) &= \min(f(x), g(x)) \text{ for all } x \in A. \end{aligned}$$

If we are only given one function $f \in F_b(A, \mathbb{R})$, then we can define the absolute value function $|f| \in F_b(A, \mathbb{R})$. The formula is given by

$$|f| : A \rightarrow \mathbb{R} \quad |f|(x) = |f(x)| \text{ for all } x \in A.$$

If we are given a function $f \in F_b(A, \mathbb{R})$ and a scalar $\alpha \in \mathbb{R}$, we can define the scalar amplification $\alpha f \in F_b(A, \mathbb{R})$, with the formula

$$\alpha f : A \rightarrow \mathbb{R} \quad (\alpha f)(x) = \alpha \cdot f(x) \text{ for all } x \in A.$$

Notation 4.3. We set $\text{Int}_b(A, \mathbb{R}) := \{f \in F_b(A, \mathbb{R}) \mid f \text{ is integrable}\}$.

Example. Let $\Delta = \{A_1, \dots, A_p\}$ be a division of A , and let $c_1, \dots, c_p \in \mathbb{R}$.

Define $f : A \rightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} c_1 & \text{if } x \in A_1 \\ \vdots & \vdots \\ c_p & \text{if } x \in A_p \end{cases}$$

Observe that for every A_i , we have $\sup_{A_i}(f) = \inf_{A_i}(f) = c_i$. Then, let $(\Delta_k)_{k=1}^\infty$ be a sequence of divisions of A with $\Delta_k = \Delta$ for all $k \in \mathbb{N}$. Then,

$$U(f, \Delta_k) = \sum_{i=1}^p \mu(A_i) \cdot \sup_{A_i}(f) = \sum_{i=1}^p c_i \mu(A_i) = \sum_{i=1}^p \mu(A_i) \cdot \inf_{A_i}(f) = L(f, \Delta_k)$$

That is, $U(f, \Delta_k) - L(f, \Delta_k) = 0$ for all $k \in \mathbb{N}$.

In particular, we have a sequence of divisions such that

$$\lim_{k \rightarrow \infty} U(f, \Delta_k) - L(f, \Delta_k) = 0$$

Thus $f \in \text{Int}_b(A, \mathbb{R})$, with

$$\int_A f(x) dx = c_1 \mu(A_1) + \cdots + c_p \mu(A_p).$$

Theorem 4.4. Let $f, g \in \text{Int}_b(A, \mathbb{R})$. Then

- (1) $\alpha f + \beta g \in \text{Int}_b(A, \mathbb{R})$;
- (2) $f \cdot g \in \text{Int}_b(A, \mathbb{R})$;
- (3) $f \vee g, f \wedge g \in \text{Int}_b(A, \mathbb{R})$.

In connection to part (1) of Theorem 4.4, we also have a result concerning the "linear behaviour" of the integral when one has linear combinations of functions.

Theorem 4.5. Let $f, g \in \text{Int}_b(A, \mathbb{R})$, let $\alpha, \beta \in \mathbb{R}$, and consider the new function $\alpha f + \beta g \in \text{Int}_b(A, \mathbb{R})$. Then

$$\int_A \alpha f + \beta g = \alpha \cdot \int_A f + \beta \cdot \int_A g.$$

Lemma 4.6.

- (1) For all $f, g \in F_b(A, \mathbb{R})$, we have

$$\int_A^{\bar{}} f + g \leq \int_A^{\bar{}} f + \int_A^{\bar{}} g$$

- (2) For all $f, g \in F_b(A, \mathbb{R})$, we have

$$\int_A^{\underline{}} f + g \leq \int_A^{\underline{}} f + \int_A^{\underline{}} g$$

Proof. We focus on (1) (the proof of (2) is analogous). We are given two functions $f, g \in F_b(A, \mathbb{R})$ and we need to examine the upper integrals of f, g and $f + g$. We first look at upper Darboux sums, and note the following inequality.

CLAIM 1. For any division $\Delta = \{A_1, \dots, A_p\}$ of A we have

$$U(f + g, \Delta) \leq U(f, \Delta) + U(g, \Delta)$$

VERIFICATION OF CLAIM 1. For all A_i in Δ , we know that $\sup_{A_i}(f + g) \leq \sup_{A_i}(f) + \sup_{A_i}(g)$. Thus, we obtain

$$U(f + g, \Delta) = \sum_{i=1}^p \mu(A_i) \cdot \sup_{A_i}(f + g) \leq \sum_{i=1}^p \mu(A_i) \cdot \left(\sup_{A_i}(f) + \sup_{A_i}(g) \right) = U(f, \Delta) + U(g, \Delta)$$

In order to make good use of the inequality from Claim 1, we will prove the following fact, which follows from the definition of the upper integral as an infimum of upper sums.

CLAIM 2. For every $k \in \mathbb{N}$, we can find divisions Δ'_k and Δ''_k of A such that

$$U(f, \Delta'_k) < \frac{1}{k} + \int_A f \quad \text{and} \quad U(g, \Delta''_k) < \frac{1}{k} + \int_A g$$

VERIFICATION OF CLAIM 2. **Fill this in.**

For every $k \in \mathbb{N}$, we pick a division Δ_k of A which is a common refinement of Δ'_k and Δ''_k from Claim 2, and we write the following inequalities:

$$\begin{aligned} \int_A f + g &\leq U(f + g, \Delta_k) && \text{(by definition of upper integral)} \\ &\leq U(f, \Delta_k) + U(g, \Delta_k) && \text{(by Claim 1)} \\ &\leq U(f, \Delta'_k) + U(g, \Delta''_k) && \text{(by Lemma 3.3)} \\ &< \left(\frac{1}{k} + \int_A f \right) + \left(\frac{1}{k} + \int_A g \right) && \text{(by Claim 2)} \\ &= \frac{2}{k} + \int_A f + \int_A g \end{aligned}$$

Since this holds for all $k \in \mathbb{N}$, we in particular have

$$\int_A f + g \leq \int_A f + \int_A g$$

which is what we wanted. □

Lemma 4.7. Let $f, g \in \text{Int}_b(A, \mathbb{R})$. Then $f + g \in \text{Int}_b(A, \mathbb{R})$ and $\int_A f + g = \int_A f + \int_A g$.

Idea of Proof. Observe the following inequalities:

$$\int_A f + g \leq \bar{\int}_A f + \bar{\int}_A g = \underline{\int}_A f + \underline{\int}_A g \leq \underline{\int}_A f + g \leq \bar{\int}_A f + g$$

Justify every step of the sequence of inequalities.

Since the sequence of inequalities begins and ends with the same quantity $\bar{\int}_A f + g$, we must have that all the inequalities are in fact equalities. Hence $\underline{\int}_A f + g = \bar{\int}_A f + g$ implies that $f + g \in \text{Int}_b(A, \mathbb{R})$, with $\int_A f + g = \int_A f + \int_A g$. □

Lemma 4.8. Let $f \in \text{Int}_b(A, \mathbb{R})$ and let $\alpha \in \mathbb{R}$. Then $\alpha f \in \text{Int}_b(A, \mathbb{R})$ and $\int_A \alpha f = \alpha \int_A f$.

Idea of Proof. Consider separately the cases when $\alpha > 0$, $\alpha = 0$, and $\alpha < 0$. Check in each case what happens to the inf and sup quantities used in the definition of the lower and upper integral. When in the $\alpha < 0$ case, it is useful to note the equalities

$$\begin{aligned} \sup_B(\alpha f) &= \alpha \inf_B(f) \\ \inf_B(\alpha f) &= \alpha \sup_B(f) \end{aligned}$$

holding for every set $\emptyset \neq B \in \mathcal{C}$ such that $B \subseteq A$. □

Proofs of Theorem 4.4 (1) and Theorem 4.5. Let $f, g \in \text{Int}_b(A, \mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$ be given. From Lemma 4.8 it follows that αf and βg are in $\text{Int}_b(A, \mathbb{R})$, with

$$\int_A \alpha f = \alpha \int_A f \text{ and } \int_A \beta g = \beta \int_A g$$

Then from Lemma 4.7 applied to the functions αf and βg it then follows that $\alpha f + \beta g \in \text{Int}_b(A, \mathbb{R})$, with

$$\int_A \alpha f + \beta g = \int_A \alpha f + \int_A \beta g = \alpha \int_A f + \beta \int_A g \quad \square$$

Lemma 4.9. Let $f \in \text{Int}_b(A, \mathbb{R})$. Then

- (1) $f^2 \in \text{Int}_b(A, \mathbb{R})$, and
- (2) $|f| \in \text{Int}_b(A, \mathbb{R})$.

Proof.

- (1) We know that $f \in \text{Int}_b(A, \mathbb{R})$.

Pick $\alpha, \beta \in \mathbb{R}$ such that $\alpha \leq f(x) \leq \beta$ for all $x \in A$. Let $u : [\alpha, \beta] \rightarrow \mathbb{R}$, $u(t) = t^2$.

Note. When we have a composition $g = u \circ f$, we get

$$g(x) = u(f(x)) = (f(x))^2 = f^2(x), \text{ for all } x \in A.$$

So $g = f^2$.

CLAIM. u is c -Lipschitz for $c = 2 \max(|\alpha|, |\beta|)$.

VERIFICATION OF CLAIM. For $s, t \in [\alpha, \beta]$ we have

$$\begin{aligned} |u(s) - u(t)| &= |s^2 - t^2| = |s - t| \cdot |s + t| \\ &\leq |s - t| \cdot (|s| + |t|) \\ &\leq |s - t| \cdot 2 \max(|\alpha|, |\beta|) \end{aligned}$$

Invoke Problem 3 on Homework 2, and we get $f^2 = g \in \text{Int}_b(A, \mathbb{R})$.

- (2) We know that $f \in \text{Int}_b(A, \mathbb{R})$.

Pick $\alpha, \beta \in \mathbb{R}$ such that $\alpha \leq f(x) \leq \beta$ for all $x \in A$. Let $u : [\alpha, \beta] \rightarrow \mathbb{R}$, $u(t) = |t|$.

Then $g = u \circ f$ becomes $|f|$.

CLAIM. u is c -Lipschitz for $c = 1$.

VERIFICATION OF CLAIM. We check that $|u(s) - u(t)| \leq |s - t|$ for all $s, t \in [\alpha, \beta]$ (**fill this in**). \square

Proof of Theorem 4.4 (2). We have $f, g \in \text{Int}_b(A, \mathbb{R})$.

Observe that

$$f \cdot g = \frac{1}{4} \left((f + g)^2 - (f - g)^2 \right)$$

Then

$$\begin{aligned}
 f, g \in \text{Int}_b(A, \mathbb{R}) &\implies f + g, f - g \in \text{Int}_b(A, \mathbb{R}) && \text{(by Theorem 4.4.1)} \\
 &\implies (f + g)^2, (f - g)^2 \in \text{Int}_b(A, \mathbb{R}) && \text{(by Lemma 4.9.1)} \\
 &\implies \frac{1}{4}(f + g)^2 + (-\frac{1}{4})(f - g)^2 \in \text{Int}_b(A, \mathbb{R}) && \text{(by Theorem 4.4.1)} \\
 &\implies f \cdot g \in \text{Int}_b(A, \mathbb{R}) && \square
 \end{aligned}$$

Proof of Theorem 4.4 (3). Observe that

$$\begin{aligned}
 (f \vee g)(x) + (f \wedge g)(x) &= \max(f(x), g(x)) + \min(f(x), g(x)) \\
 &= f(x) + g(x) \\
 &= (f + g)(x)
 \end{aligned}$$

and that

$$\begin{aligned}
 (f \vee g)(x) - (f \wedge g)(x) &= \max(f(x), g(x)) - \min(f(x), g(x)) \\
 &= |f(x) - g(x)| \\
 &= |f - g|(x)
 \end{aligned}$$

Combining these, we obtain the formulas

$$\begin{aligned}
 f \vee g &= \max(f, g) = \frac{f + g + |f - g|}{2} \\
 f \wedge g &= \min(f, g) = \frac{f + g - |f - g|}{2}
 \end{aligned}$$

Now suppose that $f, g \in \text{Int}_b(A, \mathbb{R})$. Then

$$\begin{aligned}
 f, g \in \text{Int}_b(A, \mathbb{R}) &\implies f + g, f - g \in \text{Int}_b(A, \mathbb{R}) && \text{(by Theorem 4.4.1)} \\
 &\implies f + g, |f - g| \in \text{Int}_b(A, \mathbb{R}) && \text{(by Lemma 4.9.2)} \\
 &\implies \frac{f + g}{2} \pm \frac{|f - g|}{2} \in \text{Int}_b(A, \mathbb{R}) && \text{(by Theorem 4.4.1)} \\
 &\implies f \vee g, f \wedge g \in \text{Int}_b(A, \mathbb{R}) && \square
 \end{aligned}$$

In connection to parts (2) and (3) of Theorem 4.4, we cannot expect to have exact formulas for the integrals of $f \cdot g$, $f \vee g$, and $f \wedge g$. Nevertheless, we can prove some inequalities.

Proposition 4.10.

- (1) Let $f \in \text{Int}_b(A, \mathbb{R})$ be such that $f \geq 0$. Then $\int_A f \geq 0$.
- (2) Let $f, g \in \text{Int}_b(A, \mathbb{R})$ be such that $f \leq g$. Then $\int_A f \leq \int_A g$.
- (3) For any $f \in \text{Int}_b(A, \mathbb{R})$ we have the inequality $|\int_A f| \leq \int_A |f|$.

Proof.

- (1) Let $f \in \text{Int}_b(A, \mathbb{R})$ such that $f \geq 0$. For any division $\Delta = \{A_1, \dots, A_p\}$ of A , the lower sum $L(f, \Delta) = \sum_{i=1}^p \mu(A_i) \cdot \inf_{A_i}(f) \geq 0$. Hence $\int_A f = \underline{\int}_A f \geq 0$.

- (2) Let $f, g \in \text{Int}_b(A, \mathbb{R})$ such that $f \leq g$. Let $h = g - f \in \text{Int}_b(A, \mathbb{R})$ (by Theorem 4.4 (1)). We have

$$\begin{aligned} \int_A h &= \int_A g - \int_A f && \text{(by Theorem 4.5)} \\ &= \int_A g - f \\ &\geq 0 \end{aligned}$$

which implies $\int_A f \leq \int_A g$.

- (3) Let $f \in \text{Int}_b(A, \mathbb{R})$. Hence $|f| \in \text{Int}_b(A, \mathbb{R})$ as well by Lemma 4.9 (2).

For all $x \in A$, we have $-|f(x)| \leq f(x) \leq |f(x)|$. Hence $-|f| \leq f \leq |f|$. By (2), this implies that

$$\int_A -|f| = -\int_A |f| \leq \int_A f \leq \int_A |f|$$

That is,

$$\left| \int_A f \right| \leq \int_A |f|$$

as required. □

Proposition 4.11 (Cauchy-Schwarz inequality for integrals). For any $f, g \in \text{Int}_b(A, \mathbb{R})$, we have the inequality $\left| \int_A f \cdot g \right| \leq \sqrt{\int_A f^2} \cdot \sqrt{\int_A g^2}$.

Proof. Note that $\text{Int}_b(A, \mathbb{R})$ is a vector space. For $f, g \in \text{Int}_b(A, \mathbb{R})$, let us denote $\langle f, g \rangle := \int_A fg \in \mathbb{R}$ (which makes sense by Theorem 4.4 (2)).

Then $\langle f, g \rangle$ has the properties:

- **bilinearity** (by Theorem 4.5, fill this in)
- **symmetry** ($fg = gf$, hence $\int_A fg = \int_A gf$)
- **non-negativity** ($\langle f, f \rangle = \int_A f^2 \geq 0$ by Proposition 4.10 (1))

Invoke the general Cauchy-Schwarz inequality to obtain

$$|\langle f, g \rangle| \leq \sqrt{\langle f, f \rangle} \cdot \sqrt{\langle g, g \rangle}$$

That is,

$$\left| \int_A f \cdot g \right| \leq \sqrt{\int_A f^2} \cdot \sqrt{\int_A g^2}$$

□

5 Our X, \mathcal{C}, μ of choice: $\mathbb{R}^n, \widehat{\mathcal{R}}_n, \text{vol}_n$

Notation 5.1. Throughout this lecture we fix a positive integer n (a dimension). We denote

$$\mathcal{R}_n = \{[a_1, b_1) \times \cdots \times [a_n, b_n) \mid a_1 < b_1, \dots, a_n < b_n \in \mathbb{R}\} \cup \{\emptyset\}$$

Moreover, we introduce the notation

$$\widehat{\mathcal{R}}_n = \left\{ A \subseteq \mathbb{R}^n \mid \begin{array}{l} \exists p \in \mathbb{N} \text{ and } S_1, \dots, S_p \in \mathcal{R}_n \\ \text{such that } A = S_1 \cup \cdots \cup S_p \end{array} \right\}$$

Proposition 5.2.

- (1) \mathcal{R}_n is a semi-ring of subsets of \mathbb{R}^n .
- (2) $\widehat{\mathcal{R}}_n$ is a ring of subsets of \mathbb{R}^n . (Moreover, $\widehat{\mathcal{R}}_n$ is the smallest possible ring of subsets of \mathbb{R}^n that contains \mathcal{R}_n .)
- (3) Every $A \in \widehat{\mathcal{R}}_n$ can be written as $A = T_1 \cup \dots \cup T_q$ with $T_1, \dots, T_q \in \mathcal{R}_n$ such that $T_i \cap T_j = \emptyset$ for $i \neq j$.

Proof. All of these statements were proven in Homework 1. □

Remark 5.3. When dealing with the ring of sets $\widehat{\mathcal{R}}_n$, we have at our disposition some ideas pertaining to the *geometry of half-open rectangles* which can be very useful.

In particular, let us note that the very "regular" shape of the rectangles in \mathcal{R}_n allows us to consider a special type of divisions which we will call **grid-divisions**. To make the notation not overly complicated, the following definition states what a grid-division is in the special case of the dimension $n = 2$.

Definition 5.4 (Grid-division of a half-open rectangle, case of dimension $n = 2$).

Let $S = [a, b) \times [c, d)$ be a half-open rectangle in \mathbb{R}^2 (for some $a < b$ and $c < d$ in \mathbb{R}). We will use the name **grid-division** for a division of S into smaller half-open rectangles which is obtained by the following recipe:

- Choose some intermediate points $a = s_0 < s_1 < \cdots < s_p = b$, and we write $[a, b) = I_1 \cup \cdots \cup I_p$, where $I_1 = [s_0, s_1), \dots, I_p = [s_{p-1}, s_p)$.
- Choose some intermediate points $c = t_0 < t_1 < \cdots < t_q = d$, and we write $[c, d) = J_1 \cup \cdots \cup J_q$, where $J_1 = [t_0, t_1), \dots, J_q = [t_{q-1}, t_q)$.
- We divide S into $p \cdot q$ rectangles of the form $I_k \times J_\ell$, with $k \in \{1, \dots, p\}$ and $\ell \in \{1, \dots, q\}$.

Exercise 5.5 (Refinement to grid-division). Let $S = [a, b) \times [c, d)$ be a half-open rectangle in \mathbb{R}^2 , and let $\Delta = \{S_1, \dots, S_m\}$ be any division of S into half-open rectangles. That is: $S_1, \dots, S_m \in \mathcal{R}_2$, with $S_i \cap S_j = \emptyset$ for $i \neq j$, and with $S_1 \cup \cdots \cup S_m = S$.

Prove that there exists a grid-division Γ of S such that Γ refines Δ .

Idea of Proof. Extend all horizontal and vertical lines. Every S_i gets to become a union of the pieces $I_k \times J_\ell$ of the grid-division. □

Proposition 5.6. Let $\mu_0 : \mathcal{R}_n \rightarrow [0, \infty)$ be defined as follows: $\mu_0(\emptyset) = 0$, and for every proper half-open rectangle $S = [a_1, b_1) \times \cdots \times [a_n, b_n)$ with $a_1 < b_1, \dots, a_n < b_n$ in \mathbb{R} we put

$$\mu_0(S) := (b_1 - a_1) \cdots (b_n - a_n)$$

Then μ_0 is a pre-additive set function as described in Homework 2 Problem 4.

Idea of Proof. We have to verify the condition (PRE-ADD).

This can be done in two steps:

- (1) Reduce the verification of (PRE-ADD) to the case when S_1, \dots, S_k form a grid-division of S .
- (2) Verify the equality $\mu_0(S) = \sum_{i=1}^k \mu_0(S_i)$ when S_1, \dots, S_k form a grid-division of S .

When $\{S_1, \dots, S_k\}$ is a grid-division, we have that $\{S_1, \dots, S_k\}$ are $\{I_h \times J_\ell \mid 1 \leq h \leq p \text{ and } 1 \leq \ell \leq q\}$. Then

$$\begin{aligned} \mu_0(S_1) + \cdots + \mu_0(S_k) &= \sum_{h=1}^p \sum_{\ell=1}^q \mu_0(I_h \times J_\ell) \\ &= \sum_{h=1}^p \sum_{\ell=1}^q u_h v_\ell \\ &= \left(\sum_{h=1}^p u_h \right) \left(\sum_{\ell=1}^q v_\ell \right) \\ &= (b - a)(d - c) \\ &= \mu_0(S) \end{aligned}$$

Then, the general case of $\{S_1, \dots, S_k\}$ can be reduced to the case of a grid-division. Use Exercise 5.5 to get a refinement of the given grid-division. \square

Definition 5.7. Let $\mu_0 : \mathcal{R}_n \rightarrow [0, \infty)$ be the pre-additive set-function considered in Proposition 5.6. From Homework 2 Problem 4, we know that μ_0 can be extended, in a unique way, to an additive set function $\mu : \widehat{\mathcal{R}}_n \rightarrow [0, \infty)$. This additive set function μ is called the **n -dimensional volume on $\widehat{\mathcal{R}}_n$** , and will be denoted as " vol_n ".

6 Integration on \mathbb{R}^n

Recall that in Lecture 5, we determined our framework we want to use when doing integration on \mathbb{R}^n , namely $X = \mathbb{R}^n$, $\mathcal{C} = \widehat{\mathcal{R}}_n$, and $\mu = \text{vol}_n$.

The following is an example where the domain $A \subseteq \mathbb{R}^n$ is not from the field of sets $\widehat{\mathcal{R}}_n$.

Example 6.1. Suppose the dimension is $n = 2$.

Let $A = \bar{B}(\vec{0}; 1) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ and let $f : A \rightarrow \mathbb{R}$ be the function defined by

$$f((x, y)) = \sqrt{1 - (x^2 + y^2)}, \text{ for } (x, y) \in A$$

Is this function integrable? If it is, what is the value of the integral?

In this lecture, we will see that the notion of integrability can be considered for *any bounded function* defined on *any bounded subset* of \mathbb{R}^n .

Definition 6.2. A set $A \subseteq \mathbb{R}^n$ is said to be **bounded** when there exists $r > 0$ such that $\|\vec{x}\| \leq r$ for all $\vec{x} \in A$.

For a bounded set $\emptyset \neq A \subseteq \mathbb{R}^n$, we define its **diameter** to be

$$\text{diam}(A) := \sup\{\|\vec{x} - \vec{y}\| \mid \vec{x}, \vec{y} \in A\}$$

An exercise: prove that $A \subseteq \mathbb{R}^n$ is bounded if and only if there exists a half-open rectangle $S \in \mathcal{R}_n$ such that $A \subseteq S$.

Definition and Proposition 6.3. Let $\emptyset \neq A \subseteq \mathbb{R}^n$ be a set and let $f : A \rightarrow \mathbb{R}$ be a bounded function. Consider the following procedure:

- Choose a half-open rectangle $S \in \mathcal{R}_n$ such that $A \subseteq S$.
- Let $\tilde{f} : S \rightarrow \mathbb{R}$ be the function defined by

$$\tilde{f}(\vec{x}) = \begin{cases} f(\vec{x}), & \text{if } \vec{x} \in A, \\ 0, & \text{if } \vec{x} \in S \setminus A. \end{cases}$$

If \tilde{f} is integrable on S , then we will say that the original function f is **integrable on A** , and we will define the integral of f on A to be

$$\int_A f(\vec{x}) d\vec{x} := \int_S \tilde{f}(\vec{x}) d\vec{x}$$

Proof. Given that $\emptyset \neq A \subseteq \mathbb{R}^n$ is bounded and $f : A \rightarrow \mathbb{R}$ is a bounded function, we enclose $A \subseteq S$ for $S \in \mathcal{R}_n$ and extend f to \tilde{f} .

We can also enclose $A \subseteq T \in \mathcal{R}_n$ and extend f to $\tilde{\tilde{f}}$ defined by

$$\tilde{\tilde{f}}(\vec{x}) := \begin{cases} f(\vec{x}) & \text{if } \vec{x} \in A \\ 0 & \text{if } \vec{x} \in T \setminus A \end{cases}$$

We will prove that indeed, $\tilde{f}, \tilde{\tilde{f}} \in \text{Int}_b(A, \mathbb{R})$, and if $\tilde{f}, \tilde{\tilde{f}} \in \text{Int}_b(A, \mathbb{R})$, then $\int_S \tilde{f} = \int_T \tilde{\tilde{f}}$.

Look at the intersection $S \cap T \in \mathcal{R}_n$, where $A \subseteq S \cap T$. Let $\widehat{f}: S \cap T \rightarrow \mathbb{R}$ be defined as

$$\widehat{f}(\vec{x}) := \begin{cases} f(\vec{x}) & \text{if } \vec{x} \in A \\ 0 & \text{if } \vec{x} \in (S \cap T) \setminus A \end{cases}$$

Then we have

$$\widetilde{f} \in \text{Int}_b(S, \mathbb{R}) \iff \widehat{f} \in \text{Int}_b(S \cap T, \mathbb{R}) \iff \widetilde{\widehat{f}} \in \text{Int}_b(T, \mathbb{R})$$

since $S, S \cap T \in \widehat{\mathcal{R}}_n$ and \widetilde{f} extends \widehat{f} by 0 on $S \setminus (S \cap T)$, and the result follows from A2P2. This is similar for $\widetilde{\widehat{f}}$.

Also, if \widehat{f} , \widetilde{f} , and $\widetilde{\widehat{f}}$ are integrable, then by A2P2, an extension of 0 does not affect the value of the integral, hence

$$\int_S \widetilde{f} = \int_{S \cap T} \widehat{f} = \int_T \widetilde{\widehat{f}}$$

This number is taken as definition as $\int_A f$. □

Notation 6.4. Let $A \subseteq \mathbb{R}^n$ be a non-empty bounded set. We will use the notation $\text{Int}_b(A, \mathbb{R})$ for the set of all bounded functions $f: A \rightarrow \mathbb{R}$ which are integrable in the sense of the preceding definition.

Remark 6.5. Suppose $A \in \widehat{\mathcal{R}}_n$. Then $\text{Int}_b(A, \mathbb{R})$ is precisely the space of functions coming from the considerations in Lecture 4.

(To see this, use Homework 2 Problem 2.)

Proposition 6.6. Let $A \subseteq \mathbb{R}^n$ be a non-empty bounded set.

- (1) The set of functions $\text{Int}_b(A, \mathbb{R})$ is stable under all the operations with functions that were considered in Theorem 4.4.
- (2) The linearity property of the integral still holds. That is, for $f, g \in \text{Int}_b(A, \mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$, we have $\alpha f + \beta g \in \text{Int}_b(A, \mathbb{R})$, and

$$\int_A \alpha f + \beta g = \alpha \int_A f + \beta \int_A g$$

- (3) The inequalities from Propositions 4.10 and 4.11 still hold for functions in $\text{Int}_b(A, \mathbb{R})$.

Proof. Fix an $S \in \mathcal{R}_n$ such that $A \subseteq S$. For every $f \in \text{Int}_b(A, \mathbb{R})$, denote $\widetilde{f} \in \text{Int}_b(S, \mathbb{R})$ to be the extension of f by 0 on $S \setminus A$.

Observe that $\widetilde{f + g} = \widetilde{f} + \widetilde{g}$, $\widetilde{f \cdot g} = \widetilde{f} \cdot \widetilde{g}$, $\widetilde{f \vee g} = \widetilde{f} \vee \widetilde{g}$, and $\widetilde{f \wedge g} = \widetilde{f} \wedge \widetilde{g}$ (it is easy to show this).

Then we can use the results from Lecture 4 in connection to $\text{Int}_b(S, \mathbb{R})$.

For example,

$$\begin{aligned} f, g \in \text{Int}_b(A, \mathbb{R}) &\Rightarrow \widetilde{f}, \widetilde{g} \in \text{Int}_b(S, \mathbb{R}) \\ &\Rightarrow \widetilde{f} \cdot \widetilde{g} \in \text{Int}_b(S, \mathbb{R}) \\ &\Rightarrow \widetilde{f \cdot g} \in \text{Int}_b(S, \mathbb{R}) \\ &\Rightarrow f \cdot g \in \text{Int}_b(A, \mathbb{R}) \end{aligned}$$

The same can be done for the other operations. □

7 Integration of continuous functions

Definition 7.1. Let A be a subset of \mathbb{R}^n , let \vec{a} be a point in A , and let $f : A \rightarrow \mathbb{R}$ be a function.

- (1) We say that f is **continuous** at \vec{a} to mean that for every $\varepsilon > 0$, one can find a $\delta > 0$ such that whenever $\vec{a} \in A$ has $\|\vec{x} - \vec{a}\| < \delta$, it follows that $|f(\vec{x}) - f(\vec{a})| < \varepsilon$.
- (2) We say that f is **sequentially continuous** at \vec{a} to mean that whenever $(\vec{x}_k)_{k=1}^{\infty}$ is a sequence in A with $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{a}$, it follows that $\lim_{k \rightarrow \infty} f(\vec{x}_k) = f(\vec{a})$.

Proposition 7.2. Let A be a subset of \mathbb{R}^n , let \vec{a} be a point in A , and let $f : A \rightarrow \mathbb{R}$ be a function. Then f is continuous at \vec{a} if and only if f is sequentially continuous at \vec{a} .

Proof. This is Problem 4 on Assignment 3. □

Definition 7.3. Let A be a non-empty subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a function.

- (1) We say that f is **continuous** on A to mean that it is continuous at every point $\vec{a} \in A$.
- (2) We say that f is **uniformly continuous** on A to mean that for every $\varepsilon > 0$, one can find a $\delta > 0$ such that whenever $\vec{x}, \vec{y} \in A$ have $\|\vec{x} - \vec{y}\| < \delta$, it follows that $|f(\vec{x}) - f(\vec{y})| < \varepsilon$.

Definition 7.4. Let A be a non-empty subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a function.

- (1) Let $c \in [0, \infty)$. If we have that

$$|f(\vec{x}_1) - f(\vec{x}_2)| \leq c \cdot \|\vec{x}_1 - \vec{x}_2\|, \forall \vec{x}_1, \vec{x}_2 \in A$$

then we say that f is **c -Lipschitz on A** .

- (2) We will simply say that f is **Lipschitz on A** to mean that there exists $c \in [0, \infty)$ such that f is c -Lipschitz on A .

Remark 7.5. Relations between the notions related to continuity that were considered in Definitions 7.3 and 7.4.

Let $\emptyset \neq A \subseteq \mathbb{R}^n$ be a set and $f : A \rightarrow \mathbb{R}$ be a function. Then

$$(1) f \text{ is Lipschitz on } A \Rightarrow (2) f \text{ is uniformly continuous on } A \Rightarrow (3) f \text{ is continuous on } A$$

(1) \Rightarrow (2). We have that f is c -Lipschitz for some $c > 0$. Choose $\delta = \frac{\varepsilon}{1+c}$ so that $\|\vec{x} - \vec{a}\| < \delta$ for $\vec{x}, \vec{a} \in A$ implies

$$|f(\vec{x}) - f(\vec{a})| \leq c\|\vec{x} - \vec{a}\| < c\delta = \frac{c}{c+1}\varepsilon < \varepsilon$$

Hence f is uniformly continuous on A .

The converse is not true in general. For a counterexample, consider the function in Example 6.1.

(2) \Rightarrow (3). Use the corresponding δ for a given $\vec{a} \in A$.

The converse is not true. Consider the function given in A3Q5.

Definition 7.6. Let $\emptyset \neq A \subseteq \mathbb{R}^n$ and let $f : A \rightarrow \mathbb{R}$ be a function. We will say that f is **uniformly continuous modulo exceptional sets** on A to mean that it satisfies the following condition:

$$(\text{UC-mod-E}) \quad \begin{cases} \text{For every } \varepsilon > 0 \text{ one can find a subset } E_\varepsilon \subseteq A \text{ such that:} \\ \text{(i) } E_\varepsilon \in \widehat{\mathcal{R}}_n \text{ and } \text{vol}(E_\varepsilon) < \varepsilon; \text{ and} \\ \text{(ii) } f \text{ is uniformly continuous on } A \setminus E_\varepsilon. \end{cases}$$

Theorem 7.7. Let $A = [a_1, b_1) \times \cdots \times [a_n, b_n) \subseteq \mathbb{R}^n$ be a half-open rectangle, and let $f : A \rightarrow \mathbb{R}$ be a function. If f is bounded and is uniformly continuous modulo exceptional sets, then $f \in \text{Int}_b(A, \mathbb{R})$.

Corollary 7.8. Let A be a half-open rectangle in \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a function. If f is bounded and uniformly continuous on A , then $f \in \text{Int}_b(A, \mathbb{R})$.

Proof. If f is uniformly continuous on A , then it is in particular uniformly continuous modulo exceptional sets. Thus Theorem 7.7 applies, and hence f is integrable. \square

Example 7.9. Let the dimension be $n = 2$. Let $A := [0, 1) \times [0, 1) \subseteq \mathbb{R}^2$, and let $f : A \rightarrow \mathbb{R}$ be the function considered in A3Q5:

$$f((s, t)) = \sin\left(\frac{1-s}{1-t}\right), \quad \text{for } 0 \leq s, t \leq 1$$

We cannot invoke Corollary 7.8 to see if this function is integrable, since A3Q5(b) says that f is not uniformly continuous. However, we could apply Theorem 7.7 to get the result.

Example 7.10. Let the dimension be $n = 2$. Let $S := [-2, 2) \times [-2, 2) \subseteq \mathbb{R}^2$, and let $\tilde{f} : S \rightarrow \mathbb{R}$ be defined by

$$\tilde{f}((x, y)) = \begin{cases} \sqrt{1 - (x^2 + y^2)}, & \text{if } x^2 + y^2 \leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

Then \tilde{f} is bounded and uniformly continuous on S (**explain why**).

Invoking Corollary 7.8 ensures us that $\tilde{f} \in \text{Int}_b(S, \mathbb{R})$.

This example is related to Example 6.1. We had $A = \overline{B}(\vec{0}, 1) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$, and we had the function $f : A \rightarrow \mathbb{R}$ defined by $f((x, y)) := \sqrt{1 - (x^2 + y^2)}$. Now, we know that this function is integrable.

Remark 7.11. In Theorem 7.7, the domain A of the function f was picked to be a half-open rectangle, and the theorem would still be valid if A was any non-empty set picked from the ring of sets $\widehat{\mathcal{R}}_n$ (that is, if A was a finite union of half-open rectangles). But Theorem 7.7 couldn't be extended, in its present form, to the case when A is an arbitrary bounded subset of \mathbb{R}^n .

In order to see what happens, let us consider another example in the dimension $n = 2$. Let $A = \{(s, t) \in \mathbb{R}^2 \mid 0 < s, t < 1 \text{ and } s, t \in \mathbb{Q}\}$, where \mathbb{Q} is the set of rational numbers. Let $f : A \rightarrow \mathbb{R}$ be defined by putting $f(\vec{x}) = 1$ for all $\vec{x} = (s, t) \in A$. The function f is then bounded and uniformly continuous (even Lipschitz) on A . But is it true that $f \in \text{Int}_b(A, \mathbb{R})$?

In order to determine that f belongs to $\text{Int}_b(A, \mathbb{R})$, we follow the procedure described in Lecture 6:

- Enclose A in a half-open rectangle S : in particular, we can take $S = [0, 1) \times [0, 1)$.

- Extend f to a function $\tilde{f} : S \rightarrow \mathbb{R}$, by putting

$$\tilde{f}(\vec{x}) = \begin{cases} f(\vec{x}), & \text{if } \vec{x} \in A \\ 0, & \text{if } \vec{x} \in S \setminus A \end{cases}$$

- Examine the integrability of \tilde{f} on S .

We have that \tilde{f} is not integrable on $[0, 1) \times [0, 1)$. To see this, for any division Δ of $[0, 1) \times [0, 1)$, we find that $U(\tilde{f}, \Delta) = 1$ and $L(\tilde{f}, \Delta) = 0$. Hence, we cannot obtain

$$U(\tilde{f}, \Delta) - L(\tilde{f}, \Delta) < \varepsilon$$

for all $\varepsilon > 0$. It follows that f is not integrable on A .

Definition 7.12. Let A be a non-empty set in the ring of sets $\widehat{\mathcal{R}}_n$, and let $\Delta = \{A_1, \dots, A_p\}$ be a division of A . We define the **mesh** of the division Δ to be the number

$$\|\Delta\| := \max(\text{diam}(A_1), \dots, \text{diam}(A_p))$$

Exercise 7.13. Let A be a non-empty set in the ring of sets $\widehat{\mathcal{R}}_n$, and let $\sigma > 0$ be a given constant. Prove that there exists a division Δ of A such that $\|\Delta\| < \sigma$.

Idea of Proof.

- Pick a half-open rectangle S such that $A \subseteq S$.
- Find a grid-division $\Gamma = \{G_1, \dots, G_u\}$ of S with $\text{diam}(G_j) < \sigma$ for all $1 \leq j \leq u$.
- Put $\Delta = \{G_j \cap A \mid 1 \leq j \leq u, G_j \cap A \neq \emptyset\}$. Then Δ is a division of A with $\text{diam}(G_j \cap A) \leq \text{diam}(G_j) < \sigma$ for all $1 \leq j \leq u$, with $G_j \cap A \neq \emptyset$. \square

Proof of Theorem 7.7. A way to remember the proof: "Two ways of being small".

We have $A = [a_1, b_1) \times \dots \times [a_n, b_n) \subseteq \mathbb{R}^n$ and a bounded function $f : A \rightarrow \mathbb{R}$ with the (UC-mod-E) property.

We want to show that f is integrable. We will use the ε - Δ criterion for integrability (as in Theorem 3.7).

Fix, for the entire proof, an $\varepsilon > 0$. We need to find a division Δ of A such that

$$U(f, \Delta) - L(f, \Delta) < \varepsilon$$

For convenience, let us denote

$$V = (b_1 - a_1) \cdots (b_n - a_n) = \text{vol}_n(A)$$

$$c = \sup\{|f(\vec{x})| \mid \vec{x} \in A\} < \infty$$

STEP 1. Use the (UC-mod-E) property of f in connection to $\varepsilon' = \frac{\varepsilon}{4(c+1)}$.

Note: This ε' is chosen so that $2c\varepsilon' < \frac{\varepsilon}{2}$.

We find a set $E \subseteq A$, $E \in \widehat{\mathcal{R}}_n$, with $\text{vol}_n(E) < \varepsilon'$, and such that f is uniformly continuous on $A \setminus E$.

STEP 2. Use the (UNIF-CONT) property of f on $A \setminus E$, in connection to $\varepsilon'' = \frac{\varepsilon}{2V}$ (where $V = \text{vol}_n(A)$).

We find a $\delta > 0$ such that

$$(\vec{x}, \vec{y} \in A \setminus E \text{ and } \|\vec{x} - \vec{y}\| < \delta) \Rightarrow |f(\vec{x}) - f(\vec{y})| < \varepsilon''$$

STEP 3. Pick a division $\Delta_0 = \{A_1, \dots, A_p\}$ of $A \setminus E$ such that $\|\Delta_0\| < \delta$ (with $\delta > 0$ as in Step 2).

This is possible since $A \setminus E \in \widehat{\mathcal{R}}_n$ and Exercise 7.13 can be applied.

STEP 4. Take $\Delta = \{A_1, \dots, A_p, E\}$ to be a division of A .

CLAIM. The Δ found in Step 4 satisfies the ε - Δ criterion for integrability.

VERIFICATION OF CLAIM. By Remark 3.2, we have

$$U(f, \Delta) - L(f, \Delta) = \sum_{i=1}^p \text{vol}_n(A_i) \cdot \text{osc}_{A_i}(f) + \text{vol}_n(E) \cdot \text{osc}_E(f)$$

For every $1 \leq i \leq p$, note that $\text{osc}_{A_i}(f) \leq \varepsilon''$, since

$$\text{osc}_{A_i}(f) = \sup_{\vec{x}, \vec{y} \in A_i} |f(\vec{x}) - f(\vec{y})|$$

For every $\vec{x}, \vec{y} \in A_i$, we have

$$\|\vec{x} - \vec{y}\| \leq \text{diam}(A_i) \leq \|\Delta\| < \delta$$

Hence

$$(\vec{x}, \vec{y} \in A_i \subseteq A \setminus E \text{ and } \|\vec{x} - \vec{y}\| < \delta) \Rightarrow |f(\vec{x}) - f(\vec{y})| < \varepsilon''$$

and taking the supremum, we get $\text{osc}_{A_i}(f) \leq \varepsilon''$.

Also, we have $f(\vec{x}) \in [-c, c]$ for all $\vec{x} \in A$, hence $\text{osc}_E(f) \leq 2c$. Then,

$$\begin{aligned} U(f, \Delta) - L(f, \Delta) &\leq \sum_{i=1}^p \text{vol}_n(A_i) \cdot \varepsilon'' + \text{vol}_n(E) \cdot 2c \\ &= \left(\sum_{i=1}^p \text{vol}_n(A_i) \right) \varepsilon'' + \text{vol}_n(E) \cdot 2c \\ &\leq V \cdot \varepsilon'' + \varepsilon' \cdot 2c \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

This proves the claim, and we are done. □

8 Interior, closure, boundary for subsets of \mathbb{R}^n

Definition 8.1. Let A be a subset of \mathbb{R}^n .

- (1) A point $\vec{a} \in A$ is said to be an **interior point** of A when there exists $r > 0$ such that $B(\vec{a}; r) \subseteq A$.

The set of all interior points of A is called the **interior** of A , denoted as $\text{int}(A)$.

- (2) A point $\vec{b} \in A$ is said to be **adherent** to A when it has the property that

$$B(\vec{b}; r) \cap A \neq \emptyset, \forall r > 0$$

The set of all points that are adherent to A is called the **closure** of A , denoted as $\text{cl}(A)$.

Remark and Definition 8.2. For every subset $A \subseteq \mathbb{R}^n$ we have

$$\text{int}(A) \subseteq A \subseteq \text{cl}(A)$$

The set-difference $\text{cl}(A) \setminus \text{int}(A)$ is called the **boundary** of A , and is denoted by $\text{bd}(A)$.

Remark 8.3. Let M and N be two subsets of \mathbb{R}^n such that $M \subseteq N$. Then one has the inclusions

$$\text{int}(M) \subseteq \text{int}(N) \text{ and } \text{cl}(M) \subseteq \text{cl}(N)$$

The proofs of these follow immediately from the definitions of interior and closure, and are left as an exercise (A4Q4a).

With these inclusions, it is easy to derive some good properties which the interior and closure turn out to have in connection to set-operations. This is discussed in parts (b) and (c) of A4Q4.

Example 8.4. Let the dimension be $n = 2$. Let $A = [0, 2) \times [0, 1)$. Then

$$\begin{aligned} \text{int}(A) &= \{(s, t) \mid 0 < s < 2, 0 < t < 1\} \\ \text{cl}(A) &= \{(s, t) \mid 0 \leq s \leq 2, 0 \leq t \leq 1\} \end{aligned}$$

Thus, we obtain

$$\begin{aligned} \text{bd}(A) &= \text{cl}(A) \setminus \text{int}(A) \\ &= \{(0, t) \mid 0 \leq t \leq 1\} \cup \{(2, t) \mid 0 \leq t \leq 1\} \cup \{(s, 0) \mid 0 \leq s \leq 2\} \cup \{(s, 1) \mid 0 \leq s \leq 2\} \end{aligned}$$

Proposition 8.5 (duality between interior and closure). For every $A \subseteq \mathbb{R}^n$, we have the relations

$$\begin{aligned} \text{int}(\mathbb{R}^n \setminus A) &= \mathbb{R}^n \setminus \text{cl}(A) \\ \text{cl}(\mathbb{R}^n \setminus A) &= \mathbb{R}^n \setminus \text{int}(A) \end{aligned}$$

Proof. We will check the first equality, as the second one is similar.

We proceed by double-inclusion.

(\subseteq) Take $\vec{b} \in \text{int}(\mathbb{R}^n \setminus A)$. There exists $r > 0$ such that $B(\vec{b}; r) \subseteq \mathbb{R}^n \setminus A$. For this $r > 0$, we get that $B(\vec{b}; r) \cap A = \emptyset$. Hence $\vec{b} \notin \text{cl}(A)$, thus we must have that $\vec{b} \in \mathbb{R}^n \setminus \text{cl}(A)$.

(\supseteq) Take $\vec{b} \in \mathbb{R}^n \setminus \text{cl}(A)$. Then $\vec{b} \notin \text{cl}(A)$, hence there exists $r > 0$ such that $B(\vec{b}; r) \cap A = \emptyset$. For this $r > 0$, we must have $B(\vec{b}; r) \subseteq \mathbb{R}^n \setminus A$. Thus $\vec{b} \in \text{int}(\mathbb{R}^n \setminus A)$ by definition. \square

Proposition 8.6. Let A be a subset of \mathbb{R}^n and let $\vec{b} \in \mathbb{R}^n$. We have that

$$\left(\vec{b} \in \text{cl}(A) \right) \Leftrightarrow \left(\begin{array}{l} \text{There exists a sequence } (\vec{x}_k)_{k=1}^\infty \text{ in } A \\ \text{such that } \lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}. \end{array} \right)$$

Proof.

(\Rightarrow) Suppose $\vec{b} \in \text{cl}(A)$. Hence for every $k \in \mathbb{N}$, we have $B(\vec{b}; \frac{1}{k}) \cap A \neq \emptyset$. Pick a point $\vec{x}_k \in B(\vec{b}; \frac{1}{k}) \cap A$. In this way, we get a sequence $(\vec{x}_k)_{k=1}^\infty$ in A such that $\|\vec{x}_k - \vec{b}\| < \frac{1}{k}$ for all $k \in \mathbb{N}$. It follows that $\lim_{k \rightarrow \infty} \|\vec{x}_k - \vec{b}\| = 0$ (by squeeze), hence $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$.

(\Leftarrow) Suppose that there exists a sequence $(\vec{x}_k)_{k=1}^\infty$ in A such that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$. Fix an $r > 0$. Since $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$, there exists $k_0 \in \mathbb{N}$ such that $\|\vec{x}_k - \vec{b}\| < r$ for all $k \geq k_0$. So $\vec{x}_k \in B(\vec{b}; r)$ for all $k \geq k_0$. In particular, we have $\vec{x}_{k_0} \in B(\vec{b}; r) \cap A$, hence $B(\vec{b}; r) \cap A \neq \emptyset$, as required. \square

Corollary 8.7. For $A \subseteq \mathbb{R}^n$, we have the statements concerning the boundary of A :

$$(1) \text{ bd}(A) = \text{cl}(A) \cap \text{cl}(\mathbb{R}^n \setminus A)$$

$$(2) \text{ bd}(A) = \left\{ \vec{b} \in \mathbb{R}^n \left| \begin{array}{l} \text{There exists a sequence } (\vec{x}_k)_{k=1}^\infty \text{ in } A \text{ such that } \lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}, \text{ and} \\ \text{there exists a sequence } (\vec{y}_k)_{k=1}^\infty \text{ in } \mathbb{R}^n \setminus A \text{ such that } \lim_{k \rightarrow \infty} \vec{y}_k = \vec{b}. \end{array} \right. \right\}$$

$$(3) \text{ bd}(A) = \text{bd}(\mathbb{R}^n \setminus A)$$

Proof.

(1) We have

$$\begin{aligned} \text{bd}(A) &= \text{cl}(A) \setminus \text{int}(A) && \text{(by Definition 8.2)} \\ &= \text{cl}(A) \cap (\mathbb{R}^n \setminus \text{int}(A)) && \text{(operations with sets)} \\ &= \text{cl}(A) \cap \text{cl}(\mathbb{R}^n \setminus A) && \text{(duality of int and cl)} \end{aligned}$$

(2) Apply (1), and then use the description of $\text{cl}(A)$ using sequences.

(3) This is clear from either (1) or (2). \square

9 Open and closed sets in \mathbb{R}^n

Definition 9.1.

- (1) A set $A \subseteq \mathbb{R}^n$ is said to be **open** when it satisfies $A = \text{int}(A)$.

That is, A is open if and only if for every $\vec{a} \in A$, there exists $r > 0$ such that $B(\vec{a}; r) \subseteq A$.

- (2) A set $A \subseteq \mathbb{R}^n$ is said to be **closed** when it satisfies $A = \text{cl}(A)$.

That is, A is closed if and only if there is no point $\vec{b} \in \mathbb{R}^n \setminus A$ such that \vec{b} is adherent to A .

Note: For most subsets A of \mathbb{R}^n , both inclusions $\text{int}(A) \subseteq A \subseteq \text{cl}(A)$ are strict. That is, A is neither open nor closed. Also, if $\text{int}(A) = A = \text{cl}(A)$, then A is **clopen**. In \mathbb{R}^n , the only clopen sets are the empty set \emptyset and \mathbb{R}^n itself.

Proposition 9.2. Let A be a subset of \mathbb{R}^n . We have that

$$(A \text{ is closed}) \Leftrightarrow (\mathbb{R}^n \setminus A \text{ is open})$$

Proof.

(\Rightarrow) We have that A is closed, so $\text{cl}(A) = A$. Thus

$$\text{int}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus \text{cl}(A) = \mathbb{R}^n \setminus A$$

so $\mathbb{R}^n \setminus A$ is open.

(\Leftarrow) We have that $\mathbb{R}^n \setminus A$ is open, so

$$\text{int}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus A = \mathbb{R}^n \setminus \text{cl}(A)$$

Hence $\mathbb{R}^n \setminus A = \mathbb{R}^n \setminus \text{cl}(A)$. Taking complements on both sides of the equality, we obtain $A = \text{cl}(A)$. Thus A is closed. \square

Definition 9.3. A set $A \subseteq \mathbb{R}^n$ has the "no-escape property for sequences" when it satisfies the following:

$$(\text{NO-ESC}) \quad \begin{cases} \text{Whenever } (\vec{x}_k)_{k=1}^{\infty} \text{ is a sequence in } A \\ \text{such that } \lim_{k \rightarrow \infty} \vec{x}_k = \vec{b} \in \mathbb{R}^n, \\ \text{it follows that } \vec{b} \in A. \end{cases}$$

Proposition 9.4. Let $A \subseteq \mathbb{R}^n$. We have that

$$(A \text{ is closed}) \Leftrightarrow (A \text{ has the (NO-ESC) property})$$

Proof.

(\Rightarrow) We know that $A = \text{cl}(A)$. Pick a sequence $(\vec{x}_k)_{k=1}^{\infty}$ in A such that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b} \in \mathbb{R}^n$. We need to show that $\vec{b} \in A$.

By the description of $\text{cl}(A)$ using sequences, we have that the existence of a sequence $(\vec{x}_k)_{k=1}^{\infty}$ in A with $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$ implies that $\vec{b} \in \text{cl}(A)$. Since $A = \text{cl}(A)$, we get that $\vec{b} \in A$.

(\Leftarrow) Suppose that A satisfies the (NO-ESC) property. It suffices to show that $\text{cl}(A) \subseteq A$ (since $A \subseteq \text{cl}(A)$ always holds).

Pick $\vec{b} \in \text{cl}(A)$. By the description of $\text{cl}(A)$ using sequences, there exists a sequence $(\vec{x}_k)_{k=1}^\infty$ in A such that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$. Then

$$\left(\begin{array}{l} (\vec{x}_k)_{k=1}^\infty \text{ in } A \\ \lim_{k \rightarrow \infty} \vec{x}_k = \vec{b} \\ A \text{ satisfies (NO-ESC)} \end{array} \right) \Rightarrow (\vec{b} \in A)$$

This proves the inclusion, as required. \square

Remark 9.5. Let $A \subseteq \mathbb{R}^n$. For the notion of an open subset of \mathbb{R}^n , we have one description (Definition 9.1.1). But for the notion of a closed subset of \mathbb{R}^n , we have three equivalent descriptions:

Definition 9.1.2: $A = \text{cl}(A)$

Proposition 9.2: $\mathbb{R}^n \setminus A$ is open

Proposition 9.4: A satisfies (NO-ESC)

Any of these can be used as definition, but the other properties must be proved from that definition.

Example 9.6. Fix a point $\vec{c} \in \mathbb{R}^n$ (a center) and a number $r > 0$ (a radius). Then

- (1) The open ball $B(\vec{c}; r) = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x} - \vec{c}\| < r\}$ is an open subset of \mathbb{R}^n .
- (2) The closed ball $\overline{B}(\vec{c}; r) = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x} - \vec{c}\| \leq r\}$ is a closed subset of \mathbb{R}^n .

Proof.

- (1) Let $A = B(\vec{c}; r)$. Pick a point $\vec{a} \in A$. We need an $r' > 0$ such that $B(\vec{a}; r') \subseteq A$.

Take $r' = r - \|\vec{c} - \vec{a}\| > 0$.

CLAIM. The r' chosen above satisfies $B(\vec{a}; r') \subseteq B(\vec{c}; r)$.

VERIFICATION OF CLAIM. Pick a point $\vec{x} \in B(\vec{a}; r')$. Then $\|\vec{x} - \vec{a}\| < r'$, hence

$$\begin{aligned} \|\vec{x} - \vec{c}\| &\leq \|\vec{x} - \vec{a}\| + \|\vec{a} - \vec{c}\| \\ &< r' + \|\vec{a} - \vec{c}\| \\ &= r \end{aligned}$$

Hence $\vec{x} \in B(\vec{c}; r)$, so we are done.

- (2) Left as an exercise. (Hint: Show that it has the (NO-ESC) property.) \square

Remark 9.7. The collection of sets $\mathcal{T}_n := \{A \subseteq \mathbb{R}^n \mid A \text{ is open}\}$ is called the **topology** of \mathbb{R}^n .

\mathcal{T}_n has some stability properties under operations with sets. For instance:

- (1) $A, B \in \mathcal{T}_n \Rightarrow A \cup B \in \mathcal{T}_n$
- (2) $A, B \in \mathcal{T}_n \Rightarrow A \cap B \in \mathcal{T}_n$

We will prove (2), and leave (1) as an exercise. We have $\text{int}(A) = A$ and $\text{int}(B) = B$. Then $\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B) = A \cap B$, where the first equality comes from A4Q4. Hence $\text{int}(A \cap B) = A \cap B$, thus $A \cap B$ is open.

Could it be that \mathcal{T}_n is a ring of subsets of \mathbb{R}^n ? We will verify the properties:

(RS-1) $\emptyset \in \mathcal{T}_n$

We have $\text{int}(A) \subseteq A$, hence $\text{int}(\emptyset) \subseteq \emptyset$. Thus we have $\text{int}(\emptyset) = \emptyset$, so \emptyset is open.

(RS-2) $A, B \in \mathcal{T}_n \Rightarrow A \cup B \in \mathcal{T}_n$

This property holds.

(RS-3) $A, B \in \mathcal{T}_n \Rightarrow A \setminus B \in \mathcal{T}_n$

\mathcal{T}_n is not stable under set difference. Consider, for some fixed point $\vec{c} \in \mathbb{R}^n$,

$$B(\vec{c}; 2) \setminus B(\vec{c}; 1) = \{\vec{x} \in \mathbb{R}^n \mid 1 \leq \|\vec{x} - \vec{c}\| < 2\}$$

This is not open.

Since (RS-3) does not hold, \mathcal{T}_n is not a ring of subsets of \mathbb{R}^n .

10 A new look at interior and closure

Lemma 10.1. Let $A \subseteq \mathbb{R}^n$. Suppose D is an open subset of \mathbb{R}^n such that $D \subseteq A$. Then $D \subseteq \text{int}(A)$.

Proof. By A4Q4a, $D \subseteq A$ implies that $\text{int}(A) \subseteq \text{int}(D)$. Since D is open, we have $D = \text{int}(D)$, and thus we obtain $D \subseteq \text{int}(A)$. \square

Lemma 10.2. Let $A \subseteq \mathbb{R}^n$. Then $\text{int}(A)$ is an open set.

Proof. Denote $T := \text{int}(A)$. We must prove that for every $\vec{x} \in T$, there exists $r > 0$ such that $B(\vec{x}; r) \subseteq T$.

Let $\vec{x} \in T$. Then there exists $r > 0$ such that $B(\vec{x}; r) \subseteq A$. From Example 9.6.1, we know that $B(\vec{x}; r)$ is an open set. Then by Lemma 10.1, $B(\vec{x}; r) \subseteq A$ implies that $B(\vec{x}; r) \subseteq \text{int}(A) = T$, as required. \square

Proposition 10.3. Let $A \subseteq \mathbb{R}^n$. Then $\text{int}(A)$ is the largest open subset of \mathbb{R}^n which is contained in A . That is:

- (i) $\text{int}(A)$ is open, and $\text{int}(A) \subseteq A$.
- (ii) Whenever D is an open subset of \mathbb{R}^n such that $D \subseteq A$, it follows that $D \subseteq \text{int}(A)$.

Proof. This proposition is a combination of Lemma 10.1 and Lemma 10.2. Lemma 10.2 implies (i), and Lemma 10.1 implies (ii). \square

Lemma 10.4. Let $A \subseteq \mathbb{R}^n$. Then $\text{cl}(A)$ is a closed set.

Proof. We will consider $\mathbb{R}^n \setminus \text{cl}(A) = \text{int}(\mathbb{R}^n \setminus A)$. By Lemma 10.2, $\text{int}(\mathbb{R}^n \setminus A)$ is open, hence $\mathbb{R}^n \setminus \text{cl}(A)$ is open. Thus, $\text{cl}(A)$ is closed by Proposition 9.2. \square

Lemma 10.5. Let $A \subseteq \mathbb{R}^n$. Suppose that F is a closed subset of \mathbb{R}^n such that $F \supseteq A$. Then $F \supseteq \text{cl}(A)$.

Proof. We have that $F \supseteq A$ implies that $\text{cl}(F) \supseteq \text{cl}(A)$ by A4Q4a. Since F is closed, $F = \text{cl}(F)$, thus $F \supseteq \text{cl}(A)$. \square

Proposition 10.6. Let $A \subseteq \mathbb{R}^n$. Then $\text{cl}(A)$ is the smallest closed subset of \mathbb{R}^n which contains A . That is, we have:

- (i) $\text{cl}(A) \supseteq A$ and $\text{cl}(A)$ is closed.
- (ii) Whenever F is a closed subset of \mathbb{R}^n such that $F \supseteq A$, it follows that $F \supseteq \text{cl}(A)$.

Proof. This proposition is a combination of Lemma 10.4 and Lemma 10.5. Statement (i) follows from Lemma 10.4, and statement (ii) follows from Lemma 10.5. \square

11 Compact subsets of \mathbb{R}^n

Definition 11.1. A set $A \subseteq \mathbb{R}^n$ is said to be **compact** when it is closed and bounded.

Definition 11.2. A set $A \subseteq \mathbb{R}^n$ is said to be **sequentially compact** when it has the following property:

$$(\text{SEQ-CP}) \quad \begin{cases} \text{For every sequence } (\vec{x}_k)_{k=1}^\infty \text{ in } A, \\ \text{one can find a convergent subsequence } (\vec{x}_{k(p)})_{p=1}^\infty \\ \text{where } \lim_{p \rightarrow \infty} \vec{x}_{k(p)} \text{ still belongs to } A. \end{cases}$$

Theorem 11.3. For a set $A \subseteq \mathbb{R}^n$, we have that

$$(A \text{ is compact}) \Leftrightarrow (A \text{ is sequentially compact})$$

Some comments concerning the proof of Theorem 11.3:

(\Rightarrow) Invoke the Bolzano-Weierstrass Theorem for sequences in \mathbb{R}^n .

(\Leftarrow) We will use the following facts about convergent sequences in \mathbb{R}^n :

FACT 1. The limit of a convergent sequence $(\vec{x}_k)_{k=1}^\infty$ in \mathbb{R}^n is uniquely determined.

VERIFICATION OF FACT 1. Suppose that one finds that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{a}$ and $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$. Then for every $k \geq 1$, we have

$$0 \leq \|\vec{a} - \vec{b}\| \leq \|\vec{a} - \vec{x}_k\| + \|\vec{x}_k - \vec{b}\|$$

Since $\lim_{k \rightarrow \infty} \|\vec{a} - \vec{x}_k\| + \|\vec{x}_k - \vec{b}\| = 0 + 0 = 0$, then by squeeze, we obtain $\|\vec{a} - \vec{b}\| = 0$, and thus $\vec{a} = \vec{b}$.

FACT 2. Let $(\vec{x}_k)_{k=1}^\infty$ be a convergent sequence of vectors in \mathbb{R}^n , and let $\lim_{k \rightarrow \infty} \vec{x}_k =: \vec{a}$. Then for any $1 \leq k(1) < k(2) < \dots < k(p) < \dots$ in \mathbb{N} , the subsequence $(\vec{x}_{k(p)})_{p=1}^\infty$ of $(\vec{x}_k)_{k=1}^\infty$ converges to the same limit \vec{a} .

VERIFICATION OF FACT 2. This is a direct consequence of the definition of a convergent sequence (**fill this in**).

FACT 3. A convergent sequence $(\vec{x}_k)_{k=1}^\infty$ in \mathbb{R}^n is sure to be a bounded sequence.

VERIFICATION OF FACT 3. Denote $\lim_{k \rightarrow \infty} \vec{x}_k =: \vec{a}$, and pick a $k_0 \in \mathbb{N}$ such that $\|\vec{x}_k - \vec{a}\| < 1$ for all $k \geq k_0$. Observe that for every $k \geq k_0$, we have

$$\|\vec{x}_k\| = \|\vec{x}_k - \vec{a} + \vec{a}\| \leq \|\vec{x}_k - \vec{a}\| + \|\vec{a}\| < 1 + \|\vec{a}\|$$

Then, consider the constant $c > 0$ defined by

$$c := (\|\vec{x}_1\| + \dots + \|\vec{x}_{k_0}\|) + (1 + \|\vec{a}\|)$$

We have that for all $k \in \mathbb{N}$, $\|\vec{x}_k\| \leq c$ (**fill this in**). Thus, $(\vec{x}_k)_{k=1}^\infty$ is bounded.

Proof.

(\Rightarrow) Pick a sequence $(\vec{x}_k)_{k=1}^\infty$ in A . Since A is bounded, $(\vec{x}_k)_{k=1}^\infty$ is a bounded sequence. By the Bolzano-Weierstrass Theorem, we can find indices $k(1) < k(2) < \dots < k(p) < \dots$ in \mathbb{N} such that

the subsequence $(\vec{x}_{k(p)})_{p=1}^{\infty}$ of $(\vec{x}_k)_{k=1}^{\infty}$ is convergent to a limit $\vec{a} \in \mathbb{R}^n$. Finally, we get:

$$\left(\begin{array}{c} \lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{a} \\ A \text{ is closed (hence has (NO-ESC) property)} \end{array} \right) \Rightarrow (\vec{a} \in A)$$

(\Leftarrow) Suppose that A has the (SEQ-CP) property.

CLOSED. We will prove that every $\vec{b} \in \text{cl}(A)$ must belong to A .

Pick $\vec{b} \in \text{cl}(A)$. By the description of the closure using sequences, there exists a sequence $(\vec{x}_k)_{k=1}^{\infty}$ in A such that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$. By the (SEQ-CP) property, we can find indices $k(1) < k(2) < \dots < k(p) < \dots$ in \mathbb{N} such that $\lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{a} \in A$.

Then, Fact 2 says that since $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{b}$, we must also have $\lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{b}$. By Fact 1, the limit of a convergent sequence is uniquely determined, so we indeed have $\vec{a} = \vec{b}$.

Hence $\vec{b} \in A$, so $\text{cl}(A) \subseteq A$, which implies that A is closed.

BOUNDED. Assume by contradiction that A is not bounded. Hence, there does not exist an $r > 0$ such that $\|\vec{x}\| \leq r$, for all $\vec{x} \in A$.

For $r = k$, we find $\vec{x}_k \in A$ such that $\|\vec{x}_k\| > k$. Now, we have a sequence $(\vec{x}_k)_{k=1}^{\infty}$ in A , and by the (SEQ-CP) property, we can find indices $k(1) < k(2) < \dots < k(p) < \dots$ in \mathbb{N} such that $\lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{a} \in A$.

Fact 3 says that $(\vec{x}_{k(p)})_{p=1}^{\infty}$ is bounded since it is convergent. Hence there exists $r > 0$ such that $\|\vec{x}_{k(p)}\| \leq r$ for all $p \in \mathbb{N}$.

For every $p \in \mathbb{N}$ we note that

$$r \geq \|\vec{x}_{k(p)}\| > k(p) \geq p$$

This r is greater than all the natural numbers, so we have a contradiction. Thus, A must be bounded, as required. \square

12 Continuous functions on compact subsets of \mathbb{R}^n

Recall that from Lecture 7, f being uniformly continuous on A implies that f is continuous on A , and that the converse of this implication is not generally true.

The following proposition tells us that the converse of the implication does become true if we make the additional hypothesis that A is a compact set.

Proposition 12.1. Let A be a compact subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a function. If f is continuous on A , then f is uniformly continuous on A .

Proof. By contradiction, suppose that f is not uniformly continuous on A . There exists an $\varepsilon > 0$ for which no $\delta > 0$ works in the (UNIF-CONT) condition. Fix this value of ε . For all $k \in \mathbb{N}$, we have that $\delta = 1/k$ does not work – that is, we can find $\vec{x}_k, \vec{y}_k \in A$ with $\|\vec{x}_k - \vec{y}_k\| < 1/k$, but $|f(\vec{x}_k) - f(\vec{y}_k)| \geq \varepsilon$.

This creates two sequences $(\vec{x}_k)_{k=1}^\infty$ and $(\vec{y}_k)_{k=1}^\infty$ in A . We focus on the first sequence. Since A is compact, hence sequentially compact,

$$(\diamond) \quad \begin{cases} \text{We can find indices} \\ k(1) < k(2) < \cdots < k(p) < \cdots \text{ in } \mathbb{N} \\ \text{such that } \lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{a} \in A. \end{cases}$$

CLAIM 1. For the same indices $k(1) < k(2) < \cdots < k(p) < \cdots$ and the same \vec{a} as in (\diamond) , we also have that $\lim_{p \rightarrow \infty} \vec{y}_{k(p)} = \vec{a}$.

VERIFICATION OF CLAIM 1. We have:

$$\begin{aligned} 0 \leq \|\vec{y}_{k(p)} - \vec{a}\| &\leq \|\vec{y}_{k(p)} - \vec{x}_{k(p)}\| + \|\vec{x}_{k(p)} - \vec{a}\| \\ &< \frac{1}{k(p)} + \|\vec{x}_{k(p)} - \vec{a}\| \xrightarrow{p \rightarrow \infty} 0 + 0 = 0 \end{aligned}$$

It follows that by squeeze, $\lim_{p \rightarrow \infty} \|\vec{y}_{k(p)} - \vec{a}\| = 0$. Hence, $\lim_{p \rightarrow \infty} \vec{y}_{k(p)} = \vec{a}$.

CLAIM 2. Let $k(1) < k(2) < \cdots < k(p) < \cdots$ and \vec{a} be as in Claim 1. Then

$$f(\vec{x}_{k(p)}) \xrightarrow{p \rightarrow \infty} f(\vec{a}) \quad \text{and} \quad f(\vec{y}_{k(p)}) \xrightarrow{p \rightarrow \infty} f(\vec{a})$$

VERIFICATION OF CLAIM 2. This follows from the fact that $\lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{a}$, $\lim_{p \rightarrow \infty} \vec{y}_{k(p)} = \vec{a}$, and the hypothesis that f is continuous (hence sequentially continuous) at $\vec{a} \in A$.

For the above indices $k(1) < k(2) < \cdots < k(p) < \cdots$ found above, we subtract the two convergent sequences of real numbers found in Claim 2, which gives

$$\lim_{p \rightarrow \infty} f(\vec{x}_{k(p)}) - f(\vec{y}_{k(p)}) = f(\vec{a}) - f(\vec{a}) = 0$$

But this is a contradiction with the inequality $|f(\vec{x}_{k(p)}) - f(\vec{y}_{k(p)})| \geq \varepsilon$ for all $p \in \mathbb{N}$, which was included in the way we selected the sequences $(\vec{x}_k)_{k=1}^\infty$ and $(\vec{y}_k)_{k=1}^\infty$. Thus, f must be uniformly continuous on A . \square

Corollary 12.2. Let $m, n \in \mathbb{N}$ be two dimensions, let A be a compact subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}^m$ be a function. If f is continuous on A , then f is uniformly continuous on A .

Proof. Let $f_1, \dots, f_m : A \rightarrow \mathbb{R}$ be the m component functions for the function f . Then we have:

$$\begin{aligned} f \text{ is continuous on } A &\Rightarrow \text{each of } f_1, \dots, f_m \text{ is continuous on } A \text{ (by upgrades to continuity)} \\ &\Rightarrow \text{each of } f_1, \dots, f_m \text{ is uniformly continuous on } A \text{ (by Proposition 12.1)} \\ &\Rightarrow f \text{ is uniformly continuous on } A \text{ (by upgrades to continuity)} \quad \square \end{aligned}$$

The following proposition tells us that the image of a compact set under a continuous function is sure to be compact as well.

Proposition 12.3. Let $m, n \in \mathbb{N}$ be two dimensions, let A be a compact subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}^m$ be a continuous function. Consider the image-set

$$M := f(A) = \{\vec{y} \in \mathbb{R}^m \mid \exists \vec{x} \in A \text{ such that } f(\vec{x}) = \vec{y}\}$$

Then M is a compact subset of \mathbb{R}^m .

Proof. We will verify that M is sequentially compact, hence compact. Fix a sequence $(\vec{y}_k)_{k=1}^\infty$ in M . We need to find a convergent subsequence $(\vec{y}_{k(p)})_{p=1}^\infty$ with its limit still in M .

STEP 1. For every $k \in \mathbb{N}$, we have $\vec{y}_k \in M = f(A)$. Hence, there exists $\vec{x}_k \in A$ such that $\vec{y}_k = f(\vec{x}_k)$.

STEP 2. A is compact, hence sequentially compact, so for the sequence $(\vec{x}_k)_{k=1}^\infty$ in A found in Step 1, we can find indices $k(1) < k(2) < \dots < k(p) < \dots$ in \mathbb{N} such that $\lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{a} \in A$.

STEP 3. We have that f is continuous on A , hence sequentially continuous at $\vec{a} \in A$ from Step 2. From $\lim_{p \rightarrow \infty} \vec{x}_{k(p)} = \vec{a}$, it follows that $\lim_{p \rightarrow \infty} f(\vec{x}_{k(p)}) = f(\vec{a})$ (convergence in \mathbb{R}^m). But $f(\vec{x}_{k(p)}) = \vec{y}_{k(p)}$ for all $p \in \mathbb{N}$. Thus we have found a subsequence $(\vec{y}_{k(p)})_{p=1}^\infty$ of $(\vec{y}_k)_{k=1}^\infty$ such that $\lim_{p \rightarrow \infty} \vec{y}_{k(p)} \in M$. \square

Definition 12.4. Let $A \subseteq \mathbb{R}^n$ and let $f : A \rightarrow \mathbb{R}$ be a function.

- (1) A point $\vec{a} \in A$ is said to be a **point of global minimum** for f on A when it has the property that $f(\vec{a}) \leq f(\vec{x})$ for all $\vec{x} \in A$.
- (2) A point $\vec{b} \in A$ is said to be a **point of global maximum** for f on A when it has the property that $f(\vec{b}) \geq f(\vec{x})$ for all $\vec{x} \in A$.

Remark 12.5. Points of global minimum and maximum may or may not exist. When they exist, they may or may not be unique.

For example, let $n = 2$ be a dimension, and consider the set $A = \{(s, t) \in \mathbb{R}^2 \mid s > 0, t \geq 1/s\}$. Let $f : A \rightarrow \mathbb{R}$ be a function defined by $f((s, t)) = t$.

As an exercise, verify these facts:

- A is closed, but not compact (verify that the (NO-ESC) property holds).
- f is c -Lipschitz for $c = 1$, hence f is continuous.

But f has no global minimum and no global maximum on A .

It has no global maximum, since $\{f(\vec{x}) \mid \vec{x} \in A\} = \{t \mid (s, t) \in A\} = (0, \infty)$ is not bounded from above. It has no global minimum, since $\inf\{f(\vec{x}) \mid \vec{x} \in A\} = \inf(0, \infty) = 0$ does not lie on the image-set $f(A)$.

Theorem 12.6 (Extreme Value Theorem). Let $A \subseteq \mathbb{R}^n$ and let $f : A \rightarrow \mathbb{R}$ be a function. If A is compact and f is continuous, then there exists $\vec{a}, \vec{b} \in A$ such that $f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b})$ for all $\vec{x} \in A$.

Proof. Consider the image-set $K = f(A) = \{f(\vec{x}) \mid \vec{x} \in A\} \subseteq \mathbb{R}$. Then K is a non-empty compact subset of \mathbb{R} (from Proposition 12.3). Applying Lemma 12.7 to this K , we get $\alpha, \beta \in K$ such that $\alpha \leq t \leq \beta$ for all $t \in K$. Hence there exists $\vec{a}, \vec{b} \in A$ such that $f(\vec{a}) = \alpha$ and $f(\vec{b}) = \beta$. Then for every $\vec{x} \in A$,

$$f(\vec{x}) \in f(A) = K \Rightarrow \alpha \leq f(\vec{x}) \leq \beta \Rightarrow f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b}). \quad \square$$

Lemma 12.7. Let K be a non-empty compact subset of \mathbb{R} . Then K must have a minimum element and a maximum element. That is, there exist two numbers $\alpha, \beta \in K$ such that $\alpha \leq t \leq \beta$ for all $t \in K$.

Proof. Since K is compact, it is bounded (above and below). Consider $\alpha = \inf K$ and $\beta = \sup K$, where $\alpha, \beta \in \mathbb{R}$. Then using the definition of \inf , for all $k \in \mathbb{N}$, we can find $t_k \in K$ such that $\alpha \leq t_k < \alpha + 1/k$. By squeeze, we get that $\lim_{k \rightarrow \infty} t_k = \alpha$. But K is closed, hence it has the (NO-ESC) property. So $t_k \in K$ for all $k \in \mathbb{N}$, and $\lim_{k \rightarrow \infty} t_k = \alpha$ implies that $\alpha \in K$. The proof that $\beta \in K$ is analogous. \square

13 Jordan measurable subsets of \mathbb{R}^n

Definition 13.1. A **null-set** is a set $N \subseteq \mathbb{R}^n$ with the following property:

$$(\text{NULL-SET}) \quad \begin{cases} \text{For every } \varepsilon > 0, \text{ one can find } A \in \widehat{\mathcal{R}}_n \\ \text{such that } N \subseteq A \text{ and } \text{vol}_n(A) < \varepsilon. \end{cases}$$

Proposition 13.2.

- (1) Let $N \subseteq \mathbb{R}^n$ be a null-set. Then N is a bounded set.
- (2) Let $M, N \subseteq \mathbb{R}^n$ be such that $M \subseteq N$. If N is a null-set, then M is a null-set as well.
- (3) Let $N_1, \dots, N_k \subseteq \mathbb{R}^n$, and let $U := N_1 \cup \dots \cup N_k$. If N_1, \dots, N_k are null-sets, then U is a null-set as well.
- (4) Every finite subset of \mathbb{R}^n is a null-set.
- (5) Let $N \subseteq \mathbb{R}^n$ be a null-set. Then $\text{int}(N) = \emptyset$.
- (6) Let $N \subseteq \mathbb{R}^n$ be a null-set. Then $\text{cl}(N)$ is a null-set as well.
- (7) Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a transformation which has the following properties:
 - (i) Whenever $P \subseteq \mathbb{R}^n$ is a half-open rectangle, it follows that $T(P)$ is a half-open rectangle as well.
 - (ii) There exists a constant $c > 0$ such that for every half-open rectangle $P \subseteq \mathbb{R}^n$, one has $\text{vol}_n(T(P)) \leq c \cdot \text{vol}_n(P)$.

Then whenever $N \subseteq \mathbb{R}^n$ is a null-set, it follows that $T(N)$ is a null-set.

Proof.

- (1) We need to find an $r > 0$ such that $\vec{x} \in N$ implies that $\|\vec{x}\| \leq r$. Take $\varepsilon = 1$. We can find $A \in \widehat{\mathcal{R}}_n$ with $\text{vol}_n(A) < 1$ such that $N \subseteq A$. Write $A = S_1 \cup \dots \cup S_p$ with $S_1, \dots, S_p \in \mathcal{R}_n$. We can find $r_1, \dots, r_p > 0$ such that $\vec{x} \in S_i$ implies that $\|\vec{x}\| \leq r_i$ for each $1 \leq i \leq p$. Take $r = \max(r_1, \dots, r_p)$, so it follows that whenever $\vec{x} \in A = S_1 \cup \dots \cup S_p$, there exists $1 \leq i \leq p$ such that $\vec{x} \in S_i$. For this i , we get that $\|\vec{x}\| \leq r_i \leq r$, so we are done.
- (2) This is obvious (find $A \in \widehat{\mathcal{R}}_n$ that covers N , it will cover M as well).
- (3) Let $\varepsilon > 0$. For every $1 \leq i \leq k$, we can find $A_i \in \widehat{\mathcal{R}}_n$ such that $N_i \subseteq A_i$ and $\text{vol}_n(A_i) < \varepsilon/k$. Let $A = A_1 \cup \dots \cup A_k \in \widehat{\mathcal{R}}_n$. By sub-additivity of vol_n , we obtain

$$\text{vol}_n(A) \leq \sum_{i=1}^k \text{vol}_n(A_i) < \sum_{i=1}^k \frac{\varepsilon}{k} = \varepsilon$$

Then $U \subseteq A$ with $\text{vol}_n(A) < \varepsilon$, so U is a null-set.

- (4) Let $F = \{\vec{a}_1, \dots, \vec{a}_k\} = \{\vec{a}_1\} \cup \dots \cup \{\vec{a}_k\}$ be a finite subset of \mathbb{R}^n . By (3), it suffices to check that a 1-element set $\{\vec{a}\} \subseteq \mathbb{R}^n$ is a null-set (fill this in).

(5) Proved in A4Q6.

(6) Also proved in A4Q6.

(7) Exercise. □

Note: If $N \subseteq \mathbb{R}^n$ is a null-set, then $\text{int}(\text{cl}(N)) = \emptyset$. Here, N is said to be **nowhere dense**.

Consider the set $B = \{(s, t) \in \mathbb{R}^2 \mid 0 < s, t < 1, s, t \in \mathbb{Q}\}$. We have that B is not a null-set, since

$$\text{int}(\text{cl}(B)) = \text{int}([0, 1] \times [0, 1]) = (0, 1) \times (0, 1) \neq \emptyset.$$

Remark 13.3. Let $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ be a Lipschitz path in \mathbb{R}^n . Let $c > 0$ be such that γ is c -Lipschitz: that is, we have

$$\|\gamma(s) - \gamma(t)\| \leq c|s - t|, \quad \forall s, t \in [0, 1]$$

We can prove the following statements:

- (A5Q1b) Let $I = [a, b]$ be a subinterval of $[0, 1]$, and let $\ell = b - a$ denote the length of I . Then there exists $\vec{p} \in \mathbb{R}^n$ such that $\{\gamma(t) \mid a \leq t < b\} \subseteq B(\vec{p}; c\ell)$.
- (A5Q1c) The set $\gamma(I)$ can be covered with a half-open rectangle $S = [a_1, b_1) \times \cdots \times [a_n, b_n) \in \mathcal{R}_n$ with $\text{vol}_n(S) \leq (2c\ell)^n$.
- (A5Q1d) The set $N := \gamma([0, 1]) = \{\gamma(t) \mid 0 \leq t \leq 1\} \subseteq \mathbb{R}^n$ is a null-set in \mathbb{R}^n .

These ideas can be further generalized, as seen in the following proposition.

Proposition 13.4. Let $m, n \in \mathbb{N}$ be two dimensions such that $m < n$. Let Y be a bounded subset of \mathbb{R}^m , and let $h : Y \rightarrow \mathbb{R}^n$ be a Lipschitz function. Consider the image-set

$$N := h(Y) = \{\vec{x} \in \mathbb{R}^n \mid \exists \vec{y} \in Y \text{ such that } h(\vec{y}) = \vec{x}\} \subseteq \mathbb{R}^n$$

Then N is a null-set in \mathbb{R}^n .

Proof. The proof can be written by following the same steps seen in Remark 13.3, which is just a special case of this proposition in the case where $m = 1$ and $Y = [0, 1]$. □

Example 13.5. For every $n \geq 2$, the sphere $S := \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\| = 1\}$ is a null-set in \mathbb{R}^n .

Consider the case where $n = 2$. We have $S = \gamma([0, 1])$ for $\gamma : [0, 1] \rightarrow \mathbb{R}^2$, where

$$\gamma(t) = (\cos(2\pi t), \sin(2\pi t)), \quad 0 \leq t \leq 1$$

We have that this is a Lipschitz path (verify this). Thus, we can apply Proposition 13.4 with $m = 1$.

For $n = 3$, we have $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$. Write $S = S_+ \cup S_-$, where

$$\begin{aligned} S_+ &= \{(x, y, z) \in S \mid z \geq 0\} \\ S_- &= \{(x, y, z) \in S \mid z \leq 0\} \end{aligned}$$

We will show that S_+ is a null-set. We observe that S_- is simply a transformation of S_+ , so it will follow that S_- is a null-set by Proposition 13.2.7. Invoking Proposition 13.2.3, we will have that S is a null-set.

Consider the set S_+ , and use Proposition 13.4 with dimension $m = 2$. Take

$$Y = \bar{B}(\vec{0}; 1) = \{(s, t) \in \mathbb{R}^2 \mid s^2 + t^2 \leq 1\}$$

Define $h : Y \rightarrow \mathbb{R}^3$ by

$$h((s, t)) = \frac{1}{1 + (s^2 + t^2)}(2s, 2t, 1 - (s^2 + t^2)), \quad (s, t) \in Y$$

Then h is Lipschitz, with $h(Y) = S_+$ (this is A6Q2).

In general, for any $n \geq 3$, we can set $m = n - 1$, and let

$$h(\vec{y}) = \frac{1}{1 + \|\vec{y}\|^2}(2\vec{y}, 1 - \|\vec{y}\|^2)$$

for every $\vec{y} \in Y \subseteq \mathbb{R}^m$.

Definition 13.6. Let A be a subset of \mathbb{R}^n . We say that A is **Jordan measurable** to mean that:

- (i) A is bounded, and
- (ii) $\text{bd}(A)$ is a null-set.

Example 13.7.

- (1) Every rectangle in \mathbb{R}^n (open, closed, or some mixture of the two) is Jordan measurable.

No matter what choices were used, we can verify that

$$\begin{aligned} \text{cl}(A) &= [a_1, b_1] \times \cdots \times [a_n, b_n] \\ \text{int}(A) &= (a_1, b_1) \times \cdots \times (a_n, b_n) \end{aligned}$$

It follows that $\text{bd}(A) = \text{cl}(A) \setminus \text{int}(A)$ is the union of the $2n$ sets of dimension $n - 1$. Applying Proposition 13.4, we see that each of the faces can be written as the image of a Lipschitz path. Then applying Proposition 13.2.3, we obtain that $\text{bd}(A)$ is a null-set.

- (2) Every ball in \mathbb{R}^n (open or closed) is Jordan measurable.

Let $\vec{p} \in \mathbb{R}^n$ and $r > 0$. Then suppose that $B(\vec{p}; r) \subseteq A \subseteq \bar{B}(\vec{p}; r)$. It follows that

$$\bar{B}(\vec{p}; r) = \text{cl}(B(\vec{p}; r)) \subseteq \text{cl}(A) \subseteq \text{cl}(\bar{B}(\vec{p}; r)) = \bar{B}(\vec{p}; r)$$

Hence, $\text{cl}(A) = \bar{B}(\vec{p}; r)$. Similarly, we can find that $\text{int}(A) = B(\vec{p}; r)$. Then

$$\text{bd}(A) = \text{cl}(A) \setminus \text{int}(A) = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x} - \vec{p}\| = r\}$$

is a null-set by Example 13.5 and Proposition 13.2.7.

Proposition 13.8. Let \mathcal{J}_n denote the collection of all Jordan measurable subsets of \mathbb{R}^n . Then:

- (1) \mathcal{J}_n is a ring of subsets of \mathbb{R}^n .
- (2) $\mathcal{J}_n \supseteq \widehat{\mathcal{R}}_n$.

The following are facts that will be useful in proving this proposition:

(a) $\text{bd}(A \cup B) \subseteq \text{bd}(A) \cup \text{bd}(B)$

Proof: Consider the description of boundary using sequences.

(b) $\text{bd}(A \cap B) \subseteq \text{bd}(A) \cup \text{bd}(B)$

Proof:

$$\begin{aligned} \text{bd}(A \cap B) &= \text{bd}(\mathbb{R}^n \setminus (A \cap B)) \\ &= \text{bd}((\mathbb{R}^n \setminus A) \cup (\mathbb{R}^n \setminus B)) \\ &\subseteq \text{bd}(\mathbb{R}^n \setminus A) \cup \text{bd}(\mathbb{R}^n \setminus B) \\ &= \text{bd}(A) \cup \text{bd}(B) \end{aligned}$$

(c) $\text{bd}(A \setminus B) \subseteq \text{bd}(A) \cup \text{bd}(B)$

Proof:

$$\begin{aligned} \text{bd}(A \setminus B) &= \text{bd}(A \cap (\mathbb{R}^n \setminus B)) \\ &\subseteq \text{bd}(A) \cup \text{bd}(\mathbb{R}^n \setminus B) \\ &= \text{bd}(A) \cup \text{bd}(B) \end{aligned}$$

We now proceed to prove the proposition.

Proof.

(1) **(RS-1)** $\emptyset \in \mathcal{J}_n$ is immediate.

(RS-2) Let $A, B \in \mathcal{J}_n$. Then $A \cup B$ is bounded (since A and B are bounded), and we have $\text{bd}(A \cup B) \subseteq \text{bd}(A) \cup \text{bd}(B)$. Since $\text{bd}(A)$ and $\text{bd}(B)$ are null-sets, it follows that $\text{bd}(A \cup B)$ is a null-set by Proposition 13.2.3. Hence, $A \cup B \in \mathcal{J}_n$.

(RS-3) Let $A, B \in \mathcal{J}_n$. We have that $A \setminus B$ is bounded since it is a subset of A which is also bounded. Then noting that $\text{bd}(A \setminus B) \subseteq \text{bd}(A) \cup \text{bd}(B)$, it follows that $\text{bd}(A \setminus B)$ is a null-set by the same argument above. Thus $A \setminus B \in \mathcal{J}_n$.

Hence, \mathcal{J}_n forms a ring of sets.

(2) Fix $A \in \widehat{\mathcal{R}}_n$. We can write $A = S_1 \cup \cdots \cup S_k$ with $S_1, \dots, S_k \subseteq \mathbb{R}^n$ being half-open rectangles. From Example 13.7.1, we have that $S_1, \dots, S_k \in \mathcal{J}_n$. Since \mathcal{J}_n is a ring of sets, it follows that $A = S_1 \cup \cdots \cup S_k \in \mathcal{J}_n$, as required. \square

14 (UC-mod-E) upgraded to (C-mod-N)

Definition 14.1. Let A be a subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a function. We say that f is **continuous modulo a null-set** when it satisfies the following property:

$$(C\text{-mod-N}) \quad \left\{ \begin{array}{l} \text{There exists } N \subseteq A \text{ such that} \\ N \text{ is a null-set and such that} \\ f \text{ is continuous at } \vec{a} \in A \setminus N. \end{array} \right.$$

Theorem 14.2. Let A be a Jordan measurable subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a bounded function. If f satisfies (C-mod-N), then f is integrable on A .

Remark 14.3. A comparison of Theorem 14.2 and Theorem 7.7:

- Theorem 7.7 requires the domain of the function f to be a half-open rectangle, while Theorem 14.2 allows the domain of f to be any Jordan measurable subset of \mathbb{R}^n .
- Theorem 7.7 requires f to have the (UC-mod-E) property, while Theorem 14.2 requires f to have the (C-mod-N) property. It can be argued that (C-mod-N) is easier to verify than (UC-mod-E) in concrete examples.

Lemma 14.4. Let $P = [a_1, b_1] \times \cdots \times [a_n, b_n]$ be a half-open rectangle in \mathbb{R}^n . Then for every $\varepsilon > 0$, one can find a half-open rectangle Q such that $\text{cl}(Q) \subseteq P$ and such that $\text{vol}_n(Q) > \text{vol}_n(P) - \varepsilon$.

Proof. We have $\text{vol}_n(P) = \ell_1 \cdots \ell_n$, where $\ell_i := b_i - a_i$ for each $1 \leq i \leq n$. We pick Q to be $Q = [a_1, b_1 - \sigma_1] \times \cdots \times [a_n, b_n - \sigma_n]$, where $\sigma_i = \frac{1}{k} \ell_i$ with $k \geq 2$ in \mathbb{N} . Hence, we are sure to have $a_i < b_i - \sigma_i$ for all $1 \leq i \leq n$. We know that the closure of Q is

$$\text{cl}(Q) = [a_1, b_1 - \sigma_1] \times \cdots \times [a_n, b_n - \sigma_n]$$

and since $b_i - \sigma_i < b_i$ for each $1 \leq i \leq n$, it follows that $\text{cl}(Q) \subseteq P$.

Now, we have

$$\begin{aligned} \text{vol}_n(Q) &= ((b_1 - \sigma_1) - a_1) \cdots ((b_n - \sigma_n) - a_n) \\ &= (\ell_1 - \sigma_1) \cdots (\ell_n - \sigma_n) \\ &= (\ell_1 - \frac{1}{k} \ell_1) \cdots (\ell_n - \frac{1}{k} \ell_n) \\ &= (\frac{k-1}{k})^n \ell_1 \cdots \ell_n \end{aligned}$$

Hence, we obtain $\text{vol}_n(P) - \text{vol}_n(Q) = (1 - (\frac{k-1}{k})^n) \ell_1 \cdots \ell_n$.

Let $\varepsilon > 0$. We can simply choose $k \in \mathbb{N}$ such that

$$1 - \left(\frac{k-1}{k} \right)^n < \frac{\varepsilon}{\ell_1 \cdots \ell_n}$$

which is possible since $\lim_{k \rightarrow \infty} 1 - (\frac{k-1}{k})^n = 0$. □

Lemma 14.5. Let A be a set in $\widehat{\mathcal{R}}_n$. Then for every $\varepsilon > 0$, one can find $B \in \widehat{\mathcal{R}}_n$ such that $\text{cl}(B) \subseteq A$ and such that $\text{vol}_n(B) > \text{vol}_n(A) - \varepsilon$.

Proof. Let $\varepsilon > 0$. Since $A \in \widehat{\mathcal{R}}_n$, we can write $A = P_1 \cup \cdots \cup P_k$ where $P_1, \dots, P_k \subseteq \mathbb{R}^n$ are half-open rectangles and such that $P_i \cap P_j = \emptyset$ for $i \neq j$.

By Lemma 14.4, for each $1 \leq i \leq k$, we can find a half-open rectangle Q_i such that $\text{cl}(Q_i) \subseteq P_i$ and such that $\text{vol}_n(Q_i) > \text{vol}_n(P_i) - \varepsilon/k$.

Put $B = Q_1 \cup \cdots \cup Q_k \in \widehat{\mathcal{R}}_n$. Then

$$\begin{aligned} \text{cl}(B) &= \text{cl}(Q_1 \cup \cdots \cup Q_k) \\ &= \text{cl}(Q_1) \cup \cdots \cup \text{cl}(Q_k) \\ &\subseteq P_1 \cup \cdots \cup P_k = A \end{aligned}$$

Hence $\text{cl}(B) \subseteq A$. We also have

$$\begin{aligned} \text{vol}_n(B) &= \text{vol}_n(Q_1) + \cdots + \text{vol}_n(Q_k) \\ &> (\text{vol}_n(P_1) - \varepsilon/k) + \cdots + (\text{vol}_n(P_k) - \varepsilon/k) \\ &= (\text{vol}_n(P_1) + \cdots + \text{vol}_n(P_k)) - \varepsilon \\ &= \text{vol}_n(A) - \varepsilon \end{aligned} \quad \square$$

Proposition 14.6. Let $A \in \widehat{\mathcal{R}}_n$, and let $f : A \rightarrow \mathbb{R}$ be continuous on A . Then f has the (UC-mod-E) property.

Proof. Let $\varepsilon > 0$. We want to find $E \subseteq A$ such that $E \in \widehat{\mathcal{R}}_n$ and $\text{vol}_n(E) < \varepsilon$, and f is uniformly continuous on $A \setminus E$. By Lemma 14.5, pick $B \in \widehat{\mathcal{R}}_n$ such that $\text{cl}(B) \subseteq A$ and $\text{vol}_n(B) > \text{vol}_n(A) - \varepsilon$. Take $E := A \setminus B \in \widehat{\mathcal{R}}_n$. Then

$$\text{vol}_n(E) = \text{vol}_n(A) - \text{vol}_n(B) < \varepsilon$$

Then, we need to verify that f is uniformly continuous on $A \setminus E = B$. Observe that $\text{cl}(B)$ is compact (bounded since $\text{cl}(B) \subseteq A$ and A is bounded, closed since it is a closure). Hence f is defined and continuous on $\text{cl}(B)$. By Proposition 12.1, it follows that f is uniformly continuous on $\text{cl}(B)$. We have $\text{cl}(B) \supseteq B$, so f is uniformly continuous on B , completing the proof. \square

Now, we have all the tools to proceed with the proof of the main theorem of this lecture.

Proof of Theorem 14.2. Since A is Jordan measurable, it is bounded, so we can find $a_1 < b_1, \dots, a_n < b_n$ in \mathbb{R} such that

$$A \subseteq [a_1, b_1) \times \cdots \times [a_n, b_n)$$

It will be helpful to extend this half-open rectangle slightly, say

$$S := [a_1 - 1, b_1 + 1) \times \cdots \times [a_n - 1, b_n + 1)$$

Then S is a half-open rectangle such that $\text{cl}(A) \subseteq S$. Since $\text{bd}(A) \subseteq \text{cl}(A)$, we obtain $\text{bd}(A) \subseteq S$.

Define the function $\tilde{f} : S \rightarrow \mathbb{R}$ by

$$\tilde{f}(\vec{x}) = \begin{cases} f(\vec{x}) & \text{if } \vec{x} \in A \\ 0 & \text{if } \vec{x} \in S \setminus A \end{cases}$$

We examine the integrability of \tilde{f} on S , and we will use Theorem 7.7. We need to show that \tilde{f} is bounded and has the (UC-mod-E) property.

It is clear that \tilde{f} is bounded, since it is an extension of f (which is bounded) by 0.

Let $\varepsilon > 0$. We need to find a set $E \in \widehat{\mathcal{R}}_n$ with $\text{vol}_n(E) < \varepsilon$ such that \tilde{f} is uniformly continuous on $S \setminus E$.

STEP 1. Find a preliminary exceptional set.

Since $A \in \mathcal{J}_n$, we have that $\text{bd}(A)$ is a null-set. Hence, we can find $U \in \widehat{\mathcal{R}}_n$ with $\text{vol}_n(U) < \varepsilon/3$, such that $\text{bd}(A) \subseteq U$.

Since f has the (C-mod-N) property of A , we can find $N \subseteq A$ such that N is a null-set and f is continuous on $A \setminus N$.

Then we can find $V \in \widehat{\mathcal{R}}_n$ with $\text{vol}_n(V) < \varepsilon/3$ and such that $N \subseteq V$.

Put $E' = U \cup V \in \widehat{\mathcal{R}}_n$, which has

$$\text{vol}_n(E') \leq \text{vol}_n(U) + \text{vol}_n(V) < 2\varepsilon/3$$

STEP 2. Let $M = S \setminus E' \in \widehat{\mathcal{R}}_n$.

CLAIM. The restriction of \tilde{f} to M is continuous (at every point in M).

VERIFICATION OF CLAIM. Pick \vec{x} and $(\vec{x}_k)_{k=1}^\infty$ in M such that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{x}$. We want to show that $\lim_{k \rightarrow \infty} f(\vec{x}_k) = f(\vec{x})$.

Note that $\vec{x} \in M = S \setminus E' = S \setminus (U \cup V)$. Since $\vec{x} \notin U$, we have $\vec{x} \notin \text{bd}(A)$. Since $\vec{x} \notin V$, we get $\vec{x} \notin N$.

From $\vec{x} \notin \text{bd}(A)$ it follows that either $\vec{x} \notin \text{cl}(A)$ or $\vec{x} \in \text{int}(A)$. Consider two cases:

CASE 1. $\vec{x} \notin \text{cl}(A)$.

There exists $r > 0$ such that $B(\vec{x}; r) \cap A = \emptyset$. For this $r > 0$, we can find $k_0 \in \mathbb{N}$ such that $\|\vec{x}_k - \vec{x}\| < r$ for all $k \geq k_0$. Then, for all $k \geq k_0$, we have:

$$\vec{x}_k \in B(\vec{x}; r) \Rightarrow \vec{x}_k \notin A \Rightarrow \vec{x}_k \in S \setminus A \Rightarrow \tilde{f}(\vec{x}_k) = 0$$

In this case, we have that for all $k \geq k_0$,

$$0 = \lim_{k \rightarrow \infty} \tilde{f}(\vec{x}_k) = \tilde{f}(\vec{x}) = 0$$

CASE 2. $\vec{x} \in \text{int}(A)$.

There exists $r > 0$ such that $B(\vec{x}; r) \subseteq A$. For this $r > 0$, we can find $k_0 \in \mathbb{N}$ such that $\|\vec{x}_k - \vec{x}\| < r$ for all $k \geq k_0$. Then for all $k \geq k_0$, we get

$$\vec{x}_k \in B(\vec{x}; r) \subseteq A \Rightarrow \tilde{f}(\vec{x}_k) = f(\vec{x}_k)$$

Hence the limit $\lim_{k \rightarrow \infty} \tilde{f}(\vec{x}_k) = \tilde{f}(\vec{x})$ amounts to $\lim_{k \rightarrow \infty} f(\vec{x}_k) = f(\vec{x})$. This holds since $\vec{x} \in A \setminus N$, thus f is continuous at \vec{x} .

STEP 3. Use Proposition 14.6 on the restriction of \tilde{f} to M . Then \tilde{f} has the (UC-mod-E) property on M .

We can use (UC-mod-E) to find a set $E'' \subseteq M$ and $E'' \in \widehat{\mathcal{R}}_n$ with $\text{vol}_n(E'') < \varepsilon/3$, such that \tilde{f} is uniformly continuous on $M \setminus E''$.

Let $E = E' \cup E'' \in \widehat{\mathcal{R}}_n$. We have

$$\text{vol}_n(E) \leq \text{vol}_n(E') + \text{vol}_n(E'') < 2\varepsilon/3 + \varepsilon/3 = \varepsilon$$

Also, we have

$$S \setminus E = (S \setminus E') \setminus E'' = M \setminus E''$$

Hence \tilde{f} is uniformly continuous on $S \setminus E$. This is what we wanted to show, so we can conclude that f is integrable on A . \square

15 Calculation of integrals and volumes by slicing

Notation 15.1. Let $n, p, q \in \mathbb{N}$ be dimensions such that $n = p + q$. Let $S = [a_1, b_1) \times \cdots \times [a_n, b_n)$ be a half-open rectangle in \mathbb{R}^n , which we write as $S = P \times Q$, where $P = [a_1, b_1) \times \cdots \times [a_p, b_p)$ is a half-open rectangle in \mathbb{R}^p , and $Q = [a_{p+1}, b_{p+1}) \times \cdots \times [a_n, b_n)$ is a half-open rectangle in \mathbb{R}^q .

A vector $\vec{x} \in S$ is therefore of the form $\vec{x} = (\vec{v}, \vec{w})$, where $\vec{v} \in P$ and $\vec{w} \in Q$ (\vec{v} picks the first p components of \vec{x} , and \vec{w} picks the last q components of \vec{x}).

Theorem 15.2. Let $S = P \times Q$ as above, and let $f : S \rightarrow \mathbb{R}$ be a bounded function. For every $\vec{v} \in P$, we define the *slice* $f_{\vec{v}} : Q \rightarrow \mathbb{R}$ by putting

$$f_{\vec{v}}(\vec{w}) := f((\vec{v}, \vec{w})), \quad \forall \vec{w} \in Q$$

We make the following two assumptions:

- (i) The function f is integrable on S .
- (ii) For every $\vec{v} \in P$, the slice-function $f_{\vec{v}}$ is integrable on Q .

Then defining $F : P \rightarrow \mathbb{R}$ by $F(\vec{v}) = \int_Q f_{\vec{v}}(\vec{w}) d\vec{w}$, it follows that F is integrable on P , with

$$(\diamond) \quad \int_P F(\vec{v}) d\vec{v} = \int_S f(\vec{x}) d\vec{x}$$

Remark 15.3. The equation (\diamond) from Theorem 15.2 is a mechanism for the computation of integrals. We can read it as

$$(\diamond\diamond) \quad \int_S f((\vec{v}, \vec{w})) d(\vec{v}, \vec{w}) = \int_P \left(\int_Q f_{\vec{v}}(\vec{w}) d\vec{w} \right) d\vec{v}$$

This is computed as an iterated integral: first integrate $d\vec{w}$, then integrate $d\vec{v}$.

Note: Versions of Theorem 15.2 hold when the n components of $\vec{x} \in S$ are split into a group of p and a remaining group of q in some other way. That is, it is not necessary to take the first p components and the last q components.

Example 15.4. Let the dimension be $n = 3$. Let

$$A = \{(x, y, z) \in \mathbb{R}^3 \mid x, y, z \geq 0 \text{ and } x^2 + y^2 + z^2 < 1\}$$

Define $f : A \rightarrow \mathbb{R}$ to be

$$f(x, y, z) = \sqrt{1 - x^2 - y^2}$$

We notice that A is a Jordan measurable set. This is because we can write $A = G \cap H$, where $G = \{(x, y, z) \mid x^2 + y^2 + z^2 < 1\}$ and $H = [0, 1)^3$, and since $G, H \in \mathcal{J}_n$, it follows that $A = G \cap H \in \mathcal{J}_n$ because \mathcal{J}_n is a ring of sets.

Also, f is continuous on A (this is immediate from sequential continuity).

Then by Theorem 14.2, f is integrable on A .

For the concrete computation of the integral $\int_A f((x, y, z)) d(x, y, z)$, look at A7Q1.

Remark 15.5. Let $n \in \mathbb{N}$ and let A be a set in $\widehat{\mathcal{R}}_n$. Let $f : A \rightarrow \mathbb{R}$ be defined by $f(\vec{x}) = 1$ for all $\vec{x} \in A$. Then f is integrable on A , and $\int_A f = \text{vol}_n(A)$.

(This is because f is piecewise constant, and these types of functions are always integrable with an easy calculation for the value of the integral, as seen in Lecture 4.)

Definition 15.6. Let $n \in \mathbb{N}$ and let A be a set in \mathcal{J}_n . Consider the function $f : A \rightarrow \mathbb{R}$ defined by $f(\vec{x}) = 1$ for all $\vec{x} \in A$. This function is clearly bounded and has the (C-mod-N) property, hence it is integrable on A (by Theorem 14.2). We define the **volume** of A to be the number

$$\text{vol}_n(A) := \int_A f = \int_A 1 \, d\vec{x}$$

The quantity $\text{vol}_n(A)$ is also known as the **Jordan content** of the set A .

We have now extended the notion of volume from $\widehat{\mathcal{R}}_n$ to sets in the larger collection \mathcal{J}_n . Now, for $A \in \mathcal{J}_n$, we can calculate

$$\text{vol}_n(A) = \int_A 1 \, d\vec{x} = \int_S \tilde{f}(\vec{x}) \, d\vec{x}$$

where $S \in \mathcal{R}_n$ with $A \subseteq S$, and $\tilde{f} : S \rightarrow \mathbb{R}$ is the function defined by

$$\tilde{f}(\vec{x}) = \begin{cases} 1 & \vec{x} \in A \\ 0 & \vec{x} \in S \setminus A \end{cases}$$

The function \tilde{f} is known as the **indicator function** of A in S , and denoted $\tilde{f} = \chi_{A;S}$.

Proposition 15.7. The set-function $\text{vol}_n : \mathcal{J}_n \rightarrow \mathbb{R}$ defined above is additive. That is, we have

$$\left(\begin{array}{l} A_1, \dots, A_p \in \mathcal{J}_n \\ A_i \cap A_j = \emptyset \text{ for } i \neq j \end{array} \right) \Rightarrow \text{vol}_n(A_1 \cup \dots \cup A_p) = \sum_{i=1}^p \text{vol}_n(A_i)$$

Proof. We can "glue" functions together from A3Q2. □

Corollary 15.8. Let $n, p, q \in \mathbb{N}$ be dimensions such that $n = p + q$. Let S be a half-open rectangle in \mathbb{R}^n , written as $S = P \times Q$, where P is a half-open rectangle in \mathbb{R}^p and Q is a half-open rectangle in \mathbb{R}^q .

Let $M \subseteq S$, and for every $\vec{v} \in P$, consider the "slice of M " defined by

$$M_{\vec{v}} = \{\vec{w} \in Q \mid (\vec{v}, \vec{w}) \in M\} \subseteq Q$$

Suppose we have the following two hypotheses:

- (1) The set M is Jordan measurable in \mathbb{R}^n .
- (2) For every $\vec{v} \in P$, the set $M_{\vec{v}}$ is Jordan measurable in \mathbb{R}^q .

Then defining $F : P \rightarrow \mathbb{R}$ by $F(\vec{v}) := \text{vol}_q(M_{\vec{v}})$ for all $\vec{v} \in P$, it follows that F is integrable on P , and we have that

$$\int_P F(\vec{v}) \, d\vec{v} = \text{vol}_n(M)$$

In other words, we can write

$$\text{vol}_n(M) = \int_P \text{vol}_q(M_{\vec{v}}) d\vec{v}$$

Idea of Proof. Let f be the indicator function for M in S . That is, $f = \chi_{M;S}$, defined by

$$f(\vec{x}) = \begin{cases} 1 & \vec{x} \in M \\ 0 & \vec{x} \in S \setminus M \end{cases}$$

Verify that:

- Hypothesis (1) of Corollary 15.8 implies hypothesis (i) of Theorem 15.2.
- Hypothesis (2) of Corollary 15.8 implies hypothesis (ii) of Theorem 15.2.

Then, apply Theorem 15.2 to get the result. \square

Example 15.9. For every $n \in \mathbb{N}$, let $B_n = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\| < 1\}$, the unit ball sphere. We know that $B_n \in \mathcal{J}_n$ (by Example 13.7.2). Let us denote $V_n := \text{vol}_n(B_n)$.

We will find a recursive formula between V_n and V_{n+1} by slicing.

We have $B_{n+1} = \{\vec{x} \in \mathbb{R}^{n+1} \mid \|\vec{x}\| < 1\} \subseteq [-1, 1]^{n+1} =: S$. We can write $S = P \times Q$, where $P = [-1, 1]$ and $Q = [-1, 1]^n$ are half-open rectangles in \mathbb{R} and \mathbb{R}^n , respectively.

For every $v \in P$, consider the slice of B_{n+1} defined by

$$\begin{aligned} H_v &:= \{\vec{w} \in [-1, 1]^n \mid (v, \vec{w}) \in B_{n+1}\} \\ &= \{\vec{w} = (w_1, \dots, w_n) \in [-1, 1]^n \mid v^2 + w_1^2 + \dots + w_n^2 < 1\} \\ &= \{\vec{w} = (w_1, \dots, w_n) \in [-1, 1]^n \mid \|\vec{w}\| < \sqrt{1 - v^2}\} \\ &= B(\vec{0}; \sqrt{1 - v^2}) \end{aligned}$$

Hence, invoking A3Q3 (stretching of coordinates), we get

$$H_v = \{\sqrt{1 - v^2} \cdot \vec{y} \mid \vec{y} \in \mathbb{R}^n, \|\vec{y}\| < 1\}$$

Verify that: For every $v \in [-1, 1]$, the slice of B_{n+1} , H_v is Jordan measurable with $\text{vol}_n(H_v) = (\sqrt{1 - v^2})^n \cdot V_n$.

Now, apply Corollary 15.8 to get

$$V_{n+1} = \int_{-1}^1 \text{vol}_n(H_v) dv = \int_{-1}^1 (\sqrt{1 - v^2})^n \cdot V_n dv = V_n \cdot \int_{-1}^1 (1 - v^2)^{n/2} dv$$

Hence,

$$\frac{V_{n+1}}{V_n} = \int_{-1}^1 (1 - v^2)^{n/2} dv \quad \forall n \in \mathbb{N}$$

For $k \in \mathbb{N}$, the general formula is

$$V_{2k} = \frac{\pi^k}{k!} \quad V_{2k+1} = \frac{2^{k+1} \cdot \pi^k}{(1)(3) \cdots (2k+1)}$$

An interesting consequence is that $\lim_{n \rightarrow \infty} V_n = 0$ (extremely fast). Another fact is that

$$\int_0^{\pi/2} \cos^{2p}(t) dt = \frac{1}{2} \cdot \frac{\pi}{4^p} \binom{2p}{p} \quad \forall p \in \mathbb{N}$$

16 Directional derivatives and the MVT

Definition 16.1. Let A be a subset of \mathbb{R}^n , let \vec{a} be an interior point of A , and let $f : A \rightarrow \mathbb{R}$ be a function. Let us also consider a "direction" vector $\vec{v} \in \mathbb{R}^n$. The **1-dimensional reduction of the function f , in direction \vec{v} , around the point \vec{a}** is a function $h : (-c, c) \rightarrow \mathbb{R}$ of the form

$$h(t) := f(\vec{a} + t\vec{v}), \quad -c < t < c$$

where $c > 0$ is small enough to ensure that $\vec{a} + t\vec{v} \in A$ for every $t \in (-c, c)$ (namely, picking $r > 0$ such that $B(\vec{a}; r) \subseteq A$, it follows that $c = r/(1 + \|\vec{v}\|)$ is small enough).

Definition 16.2. Let $A \subseteq \mathbb{R}^n$, $\vec{a} \in \text{int}(A)$, and $\vec{v} \in \mathbb{R}^n$ be as in the preceding definition, and let $h : (-c, c) \rightarrow \mathbb{R}$ be a 1-dimensional reduction of f in direction \vec{v} , around the point \vec{a} . If h is differentiable at 0, then the number $h'(0)$ is called the **directional derivative of f at \vec{a} , in direction \vec{v}** , and is denoted by $\partial_{\vec{v}} f(\vec{a})$.

Proposition 16.3. Let A be a subset of \mathbb{R}^n , let \vec{a} be an interior point of A , and let $f : A \rightarrow \mathbb{R}$ be a function.

- (1) The directional derivative $\partial_{\vec{0}} f(\vec{a})$ exists, and is equal to 0.
- (2) Let \vec{v} be a non-zero vector in \mathbb{R}^n , let α be a real number, and put $\vec{w} := \alpha\vec{v}$. If the directional derivative $\partial_{\vec{v}} f(\vec{a})$ exists, then $\partial_{\vec{w}} f(\vec{a})$ exists as well, and we have the relation

$$(*) \quad \partial_{\vec{w}} f(\vec{a}) = \alpha \cdot \partial_{\vec{v}} f(\vec{a})$$

Proof. This is A7Q4. □

Remark 16.4. By using Proposition 16.3.2, we can always reduce the calculation of directional derivatives to the special case where the direction vector \vec{v} is such that $\|\vec{v}\| = 1$. This is because every non-zero vector $\vec{w} \in \mathbb{R}^n$ can be written in the form $\vec{w} = \alpha\vec{v}$ with $\alpha > 0$ and $\|\vec{v}\| = 1$.

Definition 16.5. For $\vec{x}, \vec{y} \in \mathbb{R}^n$, we denote

$$\text{Co}(\vec{x}, \vec{y}) := \{(1-t)\vec{x} + t\vec{y} \mid 0 \leq t \leq 1\} \subseteq \mathbb{R}^n$$

The set $\text{Co}(\vec{x}, \vec{y})$ is called the **line-segment** in \mathbb{R}^n with endpoints \vec{x} and \vec{y} .

Theorem 16.6 (Mean Value Theorem in direction \vec{v}). Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{R}$ be a function, and let \vec{v} be a direction vector in \mathbb{R}^n with $\|\vec{v}\| = 1$. We assume that the directional derivative $\partial_{\vec{v}} f(\vec{a})$ exists at every point $\vec{a} \in A$.

Let \vec{x}, \vec{y} be distinct points of A such that the following two hypotheses are satisfied:

- $\text{Co}(\vec{x}, \vec{y}) \subseteq A$
- The direction from \vec{x} to \vec{y} is given by \vec{v} . That is, there exists $\alpha > 0$ such that $\vec{y} - \vec{x} = \alpha\vec{v}$ (we must necessarily have $\alpha = \|\vec{y} - \vec{x}\|$).

Then there exists $\vec{z} \in \text{Co}(\vec{x}, \vec{y})$ such that

$$(\text{MVT-}\vec{v}) \quad \frac{f(\vec{y}) - f(\vec{x})}{\|\vec{y} - \vec{x}\|} = \partial_{\vec{v}} f(\vec{z})$$

Proof. Define $\gamma : [0, 1] \rightarrow \mathbb{R}$ by

$$\gamma(t) = f((1-t)\vec{x} + t\vec{y}), \quad 0 \leq t \leq 1$$

This function γ makes sense due to the hypothesis $\text{Co}(\vec{x}, \vec{y}) \subseteq A$.

Note that we have $\gamma(0) = f(\vec{x})$, $\gamma(1) = f(\vec{y})$, and for general $t \in [0, 1]$, $\gamma(t) = f(\vec{x} + t(\vec{y} - \vec{x}))$. Hence, we can write

$$(\diamond) \quad \gamma(t) = f(\vec{x} + t \cdot \alpha \vec{v}), \quad 0 \leq t \leq 1$$

CLAIM 1. For every $t \in [0, 1]$, γ is differentiable at t , with $\gamma'(t) = \partial_{\alpha \vec{v}} f((1-t)\vec{x} + t\vec{y})$.

VERIFICATION OF CLAIM 1. Fix $t \in [0, 1]$. Using (\diamond) , we examine the following limit:

$$\begin{aligned} \lim_{s \rightarrow 0} \frac{\gamma(t+s) - \gamma(t)}{s} &= \lim_{s \rightarrow 0} \frac{f(\vec{x} + (t+s)\alpha \vec{v}) - f(\vec{x} + t\alpha \vec{v})}{s} \\ &= \lim_{s \rightarrow 0} \frac{f((\vec{x} + t\alpha \vec{v}) + s\alpha \vec{v}) - f(\vec{x} + t\alpha \vec{v})}{s} \\ &= \lim_{s \rightarrow 0} \frac{f(\vec{p} + s\alpha \vec{v}) - f(\vec{p})}{s} \\ &= \partial_{\alpha \vec{v}} f(\vec{p}) \quad (\text{finish the verification}) \end{aligned}$$

CLAIM 2. There exists $0 < c < 1$ such that $\frac{\gamma(1)-\gamma(0)}{1-0} = \gamma'(c)$.

VERIFICATION OF CLAIM 2. Apply MVT to γ .

CLAIM 3. There exists $\vec{z} \in \text{Co}(\vec{x}, \vec{y})$ such that $f(\vec{y}) - f(\vec{x}) = \alpha \cdot \partial_{\vec{v}} f(\vec{z})$.

VERIFICATION OF CLAIM 3. Take c from Claim 2, and let $\vec{z} = (1-c)\vec{x} + c\vec{y}$. Then

$$\begin{aligned} f(\vec{y}) - f(\vec{x}) &= \gamma(1) - \gamma(0) \\ &= \gamma'(c) \\ &= \partial_{\alpha \vec{v}} f((1-c)\vec{x} + c\vec{y}) \\ &= \alpha \cdot \partial_{\vec{v}} f(\vec{z}) \end{aligned}$$

From Claim 3, for the \vec{z} we found, we get

$$\partial_{\vec{v}} f(\vec{z}) = \frac{f(\vec{y}) - f(\vec{x})}{\alpha} = \frac{f(\vec{y}) - f(\vec{x})}{\|\vec{y} - \vec{x}\|} \quad \square$$

Remark 16.7. The formula (MVT- \vec{v}) is not symmetric. However, this is not a problem. When the roles of \vec{x} and \vec{y} are switched, we now have the direction vector $-\vec{v}$, and we get

$$\frac{f(\vec{x}) - f(\vec{y})}{\|\vec{x} - \vec{y}\|} = -\frac{f(\vec{y}) - f(\vec{x})}{\|\vec{y} - \vec{x}\|} = -\partial_{\vec{v}} f(\vec{z}) = \partial_{-\vec{v}} f(\vec{z})$$

Definition 16.8. Let $A \subseteq \mathbb{R}^n$, let $\vec{a} \in \text{int}(A)$, and let $f : A \rightarrow \mathbb{R}$ be a function. Consider the special unit vectors

$$\vec{e}_1 = (1, 0, \dots, 0), \dots, \vec{e}_n = (0, \dots, 0, 1) \in \mathbb{R}^n$$

- (1) If for some $1 \leq i \leq n$, the directional derivative $\partial_{\vec{e}_i} f(\vec{a})$ exists, then this number is called the **i -th partial derivative** of f at \vec{a} , and is denoted $\partial_i f(\vec{a})$.

(2) Suppose that the partial derivative $\partial_i f(\vec{a})$ exists for all $1 \leq i \leq n$. Then the vector

$$(\partial_1 f(\vec{a}), \dots, \partial_n f(\vec{a})) \in \mathbb{R}^n$$

is called the **gradient vector** of f at \vec{a} , and is denoted by $\nabla f(\vec{a})$.

Remark 16.9. Some basic observations about partial derivatives:

- (1) Calculating partial derivatives is more straightforward than the calculation of general directional derivatives. In order to calculate the partial derivative $\partial_i f(\vec{a})$ at the point $\vec{a} = (a_1, \dots, a_n) \in A$, we simply differentiate the 1 variable function

$$u(t) = f((a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n))$$

which is defined on an interval $(a_i - r, a_i + r)$ (for r small enough) centered at the component a_i of the vector \vec{a} . The needed partial derivative then becomes

$$(*) \quad u'(a_i) = \partial_i f(\vec{a})$$

The proof that this formula works is on A8Q2.

- (2) Unlike in 1-dimensional calculus, the existence of partial derivatives (or even directional derivatives in all directions $\vec{v} \in \mathbb{R}^n$) does not generally imply the continuity of the function at that respective point. An example of this is discussed in A7Q5.

Corollary 16.10 (Mean Value Theorem in direction \vec{e}_i). Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{R}$ be a function, and let $i \in \{1, \dots, n\}$. We assume the partial derivative $\partial_i f(\vec{a})$ exists at every point $\vec{a} \in A$.

Let $\vec{x} = (x_1, \dots, x_n)$ and $\vec{y} = (y_1, \dots, y_n)$ be two points of A such that the following two hypotheses are satisfied:

- $\text{Co}(\vec{x}, \vec{y}) \subseteq A$
- $x_j = y_j$ for all $j \neq i$ in $\{1, \dots, n\}$, but $x_i \neq y_i$.

Then there exists $\vec{z} \in \text{Co}(\vec{x}, \vec{y})$ such that

$$(\text{MVT-}\vec{e}_i) \quad \frac{f(\vec{y}) - f(\vec{x})}{y_i - x_i} = \partial_i f(\vec{z})$$

Proof. If $x_i < y_i$, then we have $\vec{y} - \vec{x} = \alpha \vec{e}_i$ for $\alpha = y_i - x_i = \|\vec{y} - \vec{x}\| > 0$, and Theorem 16.6 gives us the existence of a vector $\vec{z} \in \text{Co}(\vec{x}, \vec{y})$ such that

$$\partial_i f(\vec{z}) = \frac{f(\vec{y}) - f(\vec{x})}{\|\vec{y} - \vec{x}\|} = \frac{f(\vec{y}) - f(\vec{x})}{y_i - x_i}$$

as required.

If $x_i > y_i$, then we rewrite $(f(\vec{y}) - f(\vec{x})) / (y_i - x_i)$ as $(f(\vec{x}) - f(\vec{y})) / (x_i - y_i)$, and we use the same argument above with the roles of \vec{x} and \vec{y} reversed. \square

17 C^1 -functions and their linear approximation

Definition 17.1. Let A be an open subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ be a function. Suppose that the partial derivatives $\partial_i f(\vec{a})$ exist for all points $\vec{a} \in A$ and for all $1 \leq i \leq n$. Then for every $1 \leq i \leq n$, we can consider the function $\partial_i f : A \rightarrow \mathbb{R}$, defined by

$$\vec{a} \mapsto \partial_i f(\vec{a}) \text{ for all } \vec{a} \in A$$

If all n functions $\partial_1 f, \dots, \partial_n f : A \rightarrow \mathbb{R}$ are continuous on A , then we say that f is a C^1 -function.

Theorem 17.2. Let A be an open subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ be a C^1 -function. Fix a point $\vec{a} \in A$. Then we have

$$(L\text{-Approx}) \quad \lim_{\vec{x} \rightarrow \vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle}{\|\vec{x} - \vec{a}\|} = 0$$

Remark 17.3. The statement of Theorem 17.2 refers to a limit (L-Approx) taken for $\vec{x} \rightarrow \vec{a}$. What does it mean when we talk about such a limit?

- (1) We can understand the limit (L-Approx) in terms of sequences. For every sequence $(\vec{x}_k)_{k=1}^\infty$ in A with $\vec{x}_k \neq \vec{a}$ for all $k \in \mathbb{N}$ and with $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{a}$, we have

$$(L\text{-Approx-II}) \quad \lim_{k \rightarrow \infty} \frac{f(\vec{x}_k) - f(\vec{a}) - \langle \vec{x}_k - \vec{a}, \nabla f(\vec{a}) \rangle}{\|\vec{x}_k - \vec{a}\|} = 0$$

- (2) Another way to understand the limit (L-Approx) is by using an ε - δ condition.

$$(L\text{-Approx-III}) \quad \left\{ \begin{array}{l} \text{For every } \varepsilon > 0, \text{ there exists a } \delta > 0 \text{ such that } B(\vec{a}; \delta) \subseteq A \\ \text{and such that whenever } \vec{x} \in B(\vec{a}; \delta), \text{ it follows that} \\ |f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle| \leq \varepsilon \|\vec{x} - \vec{a}\|. \end{array} \right.$$

Remark 17.4. For \vec{x} close to \vec{a} (by looking at (L-Approx-III) with small ε , say $\varepsilon = 1/100$), we have

$$f(\vec{x}) \approx f(\vec{a}) + \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle$$

This is a linear approximation, since we have

$$f(\vec{x}) \approx \alpha_1 x_1 + \dots + \alpha_n x_n + \beta$$

where $\vec{x} = (x_1, \dots, x_n)$, $\alpha_i = \partial_i f(\vec{a})$, and $\beta = f(\vec{a}) - \langle \vec{a}, \nabla f(\vec{a}) \rangle$.

Lemma 17.5. Let A be an open subset of \mathbb{R}^n , let \vec{a} be a point in A , and let $f : A \rightarrow \mathbb{R}$ be a C^1 -function. Pick a $\delta > 0$ such that $B(\vec{a}; \delta) \subseteq A$. Then for every $\vec{x} \in B(\vec{a}; \delta)$, we can find points $\vec{z}_1, \dots, \vec{z}_n \in B(\vec{a}; \delta)$ such that

$$f(\vec{x}) - f(\vec{a}) = \langle \vec{x} - \vec{a}, \vec{w} \rangle$$

with $\vec{w} = (\partial_1 f(\vec{z}_1), \dots, \partial_n f(\vec{z}_n)) \in \mathbb{R}^n$.

Proof. Fix $\vec{x} \in B(\vec{a}; \delta)$. Let us write concretely $\vec{a} = (a^{(1)}, \dots, a^{(n)})$ and $\vec{x} = (x^{(1)}, \dots, x^{(n)})$. Consider the vectors $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_n$ defined as follows:

$$\begin{aligned}\vec{x}_0 &= \vec{a} = (a^{(1)}, \dots, a^{(n)}) \\ \vec{x}_1 &= (x^{(1)}, a^{(2)}, \dots, a^{(n)}) \\ \vec{x}_i &= (x^{(1)}, \dots, x^{(i)}, a^{(i+1)}, \dots, a^{(n)}) \\ \vec{x}_n &= \vec{x} = (x^{(1)}, \dots, x^{(n)})\end{aligned}$$

Note that for every $1 \leq i \leq n$, we have

$$\|\vec{x}_i - \vec{a}\| \leq \|\vec{x} - \vec{a}\| < \delta$$

Hence $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_n \in B(\vec{a}; \delta) \subseteq A$.

CLAIM 1. For every $1 \leq i \leq n$, there exists $\vec{z}_i \in \text{Co}(\vec{x}_{i-1}, \vec{x}_i) (\subseteq B(\vec{a}; \delta) \subseteq A)$ such that $f(\vec{x}_i) - f(\vec{x}_{i-1}) = (x^{(i)} - a^{(i)}) \cdot \partial_i f(\vec{z}_i)$.

VERIFICATION OF CLAIM 1. We have

$$\begin{aligned}\vec{x}_i &= (x^{(1)}, \dots, x^{(i-1)}, x^{(i)}, a^{(i+1)}, \dots, a^{(n)}) \\ \vec{x}_{i-1} &= (x^{(1)}, \dots, x^{(i-1)}, a^{(i)}, a^{(i+1)}, \dots, a^{(n)})\end{aligned}$$

We observe that \vec{x}_i and \vec{x}_{i-1} only differ by one component. Verify the hypotheses of MVT and apply it in direction \vec{e}_i to get the result.

CLAIM 2. The vectors $\vec{z}_1, \dots, \vec{z}_n$ found in Claim 1 satisfy the conclusion of the Lemma.

VERIFICATION OF CLAIM 2. We can write

$$\begin{aligned}f(\vec{x}) - f(\vec{a}) &= f(\vec{x}_n) - f(\vec{x}_0) \\ &= \sum_{i=1}^n f(\vec{x}_i) - f(\vec{x}_{i-1}) \\ &= \sum_{i=1}^n (x^{(i)} - a^{(i)}) \cdot \partial_i f(\vec{z}_i) \\ &= \langle \vec{x} - \vec{a}, \vec{w} \rangle\end{aligned}$$

where $\vec{w} = (\partial_1 f(\vec{z}_1), \dots, \partial_n f(\vec{z}_n))$. □

Proof of Theorem 17.2. We will prove (L-Approx) using the ε - δ description. Fix some $\varepsilon > 0$. We need to find a $\delta > 0$ such that $B(\vec{a}; \delta) \subseteq A$, and whenever $\vec{x} \in B(\vec{a}; \delta)$, it follows that

$$|f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle| \leq \varepsilon \|\vec{x} - \vec{a}\|$$

We first choose $r > 0$ so that $B(\vec{a}; r) \subseteq A$. For every $1 \leq i \leq n$, we know that $\partial_i f : A \rightarrow \mathbb{R}$ is continuous at \vec{a} . Then there exists $\delta_i > 0$ such that whenever $\vec{z} \in A$ with $\|\vec{z} - \vec{a}\| < \delta_i$, it follows that

$$(\diamond_i) \quad |\partial_i f(\vec{z}) - \partial_i f(\vec{a})| < \varepsilon/n$$

Take $\delta := \min(r, \delta_1, \dots, \delta_n) > 0$.

CLAIM 1. The $\delta > 0$ found above is such that $B(\vec{a}; \delta) \subseteq A$ and such that whenever $\vec{z} \in B(\vec{a}; \delta)$, it follows that

$$(\diamond_+) \quad |\partial_i f(\vec{z}) - \partial_i f(\vec{a})| < \varepsilon, \quad \forall 1 \leq i \leq n$$

VERIFICATION OF CLAIM 1. We have $\delta \leq r$, so we have that $B(\vec{a}; \delta) \subseteq B(\vec{a}; r) \subseteq A$. Moreover, for any $\vec{z} \in B(\vec{a}; \delta)$, we have $\|\vec{z} - \vec{a}\| < \delta \leq \delta_i$ for every $1 \leq i \leq n$, hence (\diamond_i) applies to \vec{z} for all $1 \leq i \leq n$. Hence, (\diamond_+) holds.

CLAIM 2. The $\delta > 0$ found above is such that whenever $\vec{x} \in B(\vec{a}; \delta)$, it follows that

$$|f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle| \leq \varepsilon \|\vec{x} - \vec{a}\|$$

VERIFICATION OF CLAIM 2. Pick a point $\vec{x} \in B(\vec{a}; \delta)$. By Lemma 17.5, there exist points $\vec{z}_1, \dots, \vec{z}_n \in B(\vec{a}; \delta)$ such that

$$f(\vec{x}) - f(\vec{a}) = \langle \vec{x} - \vec{a}, \vec{w} \rangle$$

where $\vec{w} = (\partial_1 f(\vec{z}_1), \dots, \partial_n f(\vec{z}_n))$. Then:

$$\begin{aligned} |f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle| &= |\langle \vec{x} - \vec{a}, \vec{w} \rangle - \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle| \\ &= |\langle \vec{x} - \vec{a}, \vec{w} - \nabla f(\vec{a}) \rangle| \\ &\leq \|\vec{x} - \vec{a}\| \|\vec{w} - \nabla f(\vec{a})\| && \text{(Cauchy-Schwarz)} \\ &\leq \|\vec{x} - \vec{a}\| \cdot \sum_{i=1}^n |\partial_i f(\vec{z}_i) - \partial_i f(\vec{a})| \\ &\leq \|\vec{x} - \vec{a}\| \cdot \sum_{i=1}^n \varepsilon/n \\ &= \varepsilon \|\vec{x} - \vec{a}\| \end{aligned}$$

This proves the theorem. □

Remark 17.6. Note that in the definition of a C^1 -function, say f , we did not assume that:

- f itself is continuous, and
- f has directional derivatives in other directions besides the n special directions $\vec{e}_1, \dots, \vec{e}_n$ giving the partial derivatives.

We can prove that these hold as corollaries of Theorem 17.2.

Corollary 17.7. Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{R}$ be a C^1 -function, and let \vec{a} be a point in A . Then the directional derivative of f at \vec{a} in direction \vec{v} exists for every direction $\vec{v} \in \mathbb{R}^n$, and moreover, we have the formula

$$(*) \quad \partial_{\vec{v}} f(\vec{a}) = \langle \vec{v}, \nabla f(\vec{a}) \rangle \quad \forall \vec{v} \in \mathbb{R}^n$$

Proof. Assume that $\vec{v} \neq \vec{0}$ (the case where $\vec{v} = \vec{0}$ is obvious). Note that in (L-Approx), we can make " $\vec{x} \rightarrow \vec{a}$ " be of the form " $\vec{x} = \vec{a} + t\vec{v}$, where $t \rightarrow 0$ ". Hence, the absolute value of the fraction in (L-Approx) becomes:

$$\begin{aligned} \frac{|f(\vec{a} + t\vec{v}) - f(\vec{a}) - \langle \vec{a} + t\vec{v} - \vec{a}, \nabla f(\vec{a}) \rangle|}{\|\vec{a} + t\vec{v} - \vec{a}\|} &= \frac{|f(\vec{a} + t\vec{v}) - f(\vec{a}) - t\langle \vec{v}, \nabla f(\vec{a}) \rangle|}{|t| \cdot \|\vec{v}\|} \\ &= \frac{1}{\|\vec{v}\|} \cdot \left| \frac{f(\vec{a} + t\vec{v}) - f(\vec{a}) - t\langle \vec{v}, \nabla f(\vec{a}) \rangle}{t} \right| \\ &= \frac{1}{\|\vec{v}\|} \cdot \left| \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t} - \langle \vec{v}, \nabla f(\vec{a}) \rangle \right| \end{aligned}$$

The limit (L-Approx) tells us that this quantity goes to 0 as $t \rightarrow 0$. Hence:

$$\partial_{\vec{v}} f(\vec{a}) = \lim_{t \rightarrow 0} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t} = \langle \vec{v}, \nabla f(\vec{a}) \rangle \quad \square$$

18 C^1 implies differentiable implies continuous

Definition 18.1. Let A be a subset of \mathbb{R}^n , let \vec{a} be an interior point of A , and let $f : A \rightarrow \mathbb{R}$ be a function. We will say that f is **differentiable** at \vec{a} to mean that there exists a vector $\vec{g} \in \mathbb{R}^n$ such that

$$(\text{Diff-at-}\vec{a}) \quad \lim_{\vec{x} \rightarrow \vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \vec{g} \rangle}{\|\vec{x} - \vec{a}\|} = 0$$

Remark 18.2.

- (1) Similarly to the limit (L-Approx), we can understand the limit (Diff-at- \vec{a}) by using sequences or a suitable ε - δ condition.
- (2) The vector \vec{g} is uniquely determined. This is implied by the following proposition.

Proposition 18.3. Let A be a subset of \mathbb{R}^n , let \vec{a} be an interior point of A , and let $f : A \rightarrow \mathbb{R}$ be a function. Suppose we find a vector $\vec{g} \in \mathbb{R}^n$ such that the limit (Diff-at- \vec{a}) holds. Then $\partial_{\vec{v}} f(\vec{a})$ exists for all directions $\vec{v} \in \mathbb{R}^n$, and we have

$$(**) \quad \partial_{\vec{v}} f(\vec{a}) = \langle \vec{v}, \vec{g} \rangle \quad \forall \vec{v} \in \mathbb{R}^n$$

Proof. Similar to the proof of Corollary 17.7. □

Corollary 18.4. Let A be a subset of \mathbb{R}^n , let \vec{a} be an interior point of A , and let $f : A \rightarrow \mathbb{R}$ be a function. Suppose that f is differentiable at \vec{a} , and that \vec{g}_1 and \vec{g}_2 both satisfy the limit (Diff-at- \vec{a}). Then $\vec{g}_1 = \vec{g}_2$.

Proof. From (**) in Proposition 18.3, we have

$$\langle \vec{v}, \vec{g}_1 \rangle = \partial_{\vec{v}} f(\vec{a}) = \langle \vec{v}, \vec{g}_2 \rangle \quad \forall \vec{v} \in \mathbb{R}^n$$

Then putting $\vec{v} = \vec{e}_i$, we find that $\langle \vec{e}_i, \vec{g}_1 \rangle = \langle \vec{e}_i, \vec{g}_2 \rangle$ for all $1 \leq i \leq n$. This implies that $\vec{g}_1 = \vec{g}_2$, as required. □

Remark 18.5.

- (1) If f is differentiable at \vec{a} , then by putting $\vec{v} = \vec{e}_i$ in (**), we get

$$\partial_i f(\vec{a}) = \langle \vec{e}_i, \vec{g} \rangle = g_i \quad \forall 1 \leq i \leq n$$

Hence, the gradient vector $\nabla f(\vec{a})$ is sure to exist, and $\nabla f(\vec{a}) = \vec{g} = (g_1, \dots, g_n)$.

- (2) Our notion of differentiability works for the 1-dimensional case studied from MATH 147.

Proposition 18.6. Let $f : A \rightarrow \mathbb{R}$ be a C^1 -function on A . Then f is differentiable at every $\vec{a} \in A$.

Proof. This is immediate: use (L-Approx) with $\vec{g} = \nabla f(\vec{a})$. □

Proposition 18.7. Let A be an open subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ be a function. If f is differentiable at \vec{a} , then f is continuous at \vec{a} .

Proof. Let $\vec{a} \in A$ be such that f is differentiable at \vec{a} . Pick \vec{g} such that (Diff-at- \vec{a}) holds. We write (Diff-at- \vec{a}) using an ε - δ condition, and we choose $\varepsilon = 1$. That is:

$$(\square) \quad \left\{ \begin{array}{l} \text{There exists a } \delta > 0 \text{ such that } B(\vec{a}; \delta) \subseteq A \text{ and} \\ \text{such that whenever } \vec{x} \in B(\vec{a}; \delta), \text{ it follows that} \\ |f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \vec{g} \rangle| \leq 1 \cdot \|\vec{x} - \vec{a}\|. \end{array} \right.$$

CLAIM 1. If $\vec{x} \in B(\vec{a}; \delta)$, then $|f(\vec{x}) - f(\vec{a})| \leq \|\vec{x} - \vec{a}\| \cdot (1 + \|\vec{g}\|)$.

VERIFICATION OF CLAIM 1. We have

$$\begin{aligned} |f(\vec{x}) - f(\vec{a})| &= |(f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \vec{g} \rangle) + \langle \vec{x} - \vec{a}, \vec{g} \rangle| \\ &\leq |(f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \vec{g} \rangle)| + |\langle \vec{x} - \vec{a}, \vec{g} \rangle| \\ &\leq \|\vec{x} - \vec{a}\| + \|\vec{x} - \vec{a}\| \|\vec{g}\| \quad (\text{by } (\square) \text{ and Cauchy-Schwarz}) \\ &= \|\vec{x} - \vec{a}\| \cdot (1 + \|\vec{g}\|) \end{aligned}$$

CLAIM 2. f is sequentially continuous at \vec{a} .

VERIFICATION OF CLAIM 2. Let $(\vec{x}_k)_{k=1}^\infty$ be a sequence in A such that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{a}$. Consider the $\delta > 0$ chosen above. There exists $k_0 \in \mathbb{N}$ such that $\|\vec{x}_k - \vec{a}\| < \delta$ for all $k \geq k_0$. Then $k \geq k_0$ implies that $\vec{x}_k \in B(\vec{a}; \delta)$, and hence

$$0 \leq |f(\vec{x}_k) - f(\vec{a})| \leq \|\vec{x}_k - \vec{a}\| \cdot (1 + \|\vec{g}\|) \xrightarrow{k \rightarrow \infty} 0$$

By Squeeze Theorem, we obtain $|f(\vec{x}_k) - f(\vec{a})| \xrightarrow{k \rightarrow \infty} 0$, which proves the claim.

Since f is sequentially continuous at \vec{a} , we know that f is continuous at \vec{a} , as required. \square

19 C^2 -functions and their quadratic approximation

Definition 19.1. Let A be an open subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ be a function. Suppose that the partial derivatives $\partial_i f(\vec{a})$ exist for all $\vec{a} \in A$ and all $1 \leq i \leq n$, and consider the n functions $\partial_1 f, \dots, \partial_n f : A \rightarrow \mathbb{R}$. If all n functions are C^1 -functions, then we say that f is a **C^2 -function**.

Remark 19.2. Let A be an open subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a C^2 -function.

- (1) **C^2 -function implies C^1 -function.** Every partial derivative $\partial_i f : A \rightarrow \mathbb{R}$ is continuous on A (since C^1 -functions are continuous). Thus, f is a C^1 -function.
- (2) **Partial derivatives of second order.** For every $1 \leq i, j \leq n$, we can consider the function $\partial_j(\partial_i f) : A \rightarrow \mathbb{R}$. This function is continuous on A (since $\partial_i f$ is a C^1 -function). We will write this function without brackets, as $\partial_j \partial_i f$.

Proposition 19.3 (Interchanging the order of partial derivatives). Let A be an open subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ be a C^2 -function. Then for every $1 \leq i, j \leq n$, we have $\partial_i \partial_j f(\vec{x}) = \partial_j \partial_i f(\vec{x})$ for all $\vec{x} \in A$.

Idea of Proof. Consider the dimension $n = 2$, and we want to verify that $\partial_1 \partial_2 f = \partial_2 \partial_1 f$. Let us denote $u := \partial_1 \partial_2 f$ and $v := \partial_2 \partial_1 f$. We have that $u, v : A \rightarrow \mathbb{R}$ are two continuous functions, and we want to show that $u(\vec{x}) = v(\vec{x})$ for every $\vec{x} \in A$. By A8Q5, it suffices to verify that

$$\int_S u(\vec{x}) \, d\vec{x} = \int_S v(\vec{x}) \, d\vec{x}$$

for every half-open rectangle $S = [a_1, b_1] \times [a_2, b_2]$ such that $[a_1, b_1] \times [a_2, b_2] \subseteq A$.

Fix $S = [a_1, b_1] \times [a_2, b_2]$. We will prove that

$$\int_S u(\vec{x}) \, d\vec{x} = f(b_1, b_2) - f(b_1, a_2) - f(a_1, b_2) + f(a_1, a_2) = \int_S v(\vec{x}) \, d\vec{x}$$

We will show the calculation in relation to u . By slicing, we have

$$(*) \quad \int_S u(\vec{x}) \, d\vec{x} = \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} u(s, t) \, ds \right) dt$$

Recall that $u = \partial_1(\partial_2 f)$. Denote $\partial_2 f =: h$ (a help function) so that $u = \partial_1 h$. Pick $t \in [a_2, b_2]$ and consider the horizontal slice at level t for the functions u and h . We have $u_t, h_t : [a_1, b_1] \rightarrow \mathbb{R}$ defined by

$$u_t(s) = u(s, t) \qquad h_t(s) = h(s, t)$$

Observe that since $u = \partial_1 h$, it follows that $u_t = h'_t$ (by definition of $\partial_1 h$). Then, for our fixed t , we get

$$\begin{aligned} \int_{a_1}^{b_1} u(s, t) \, ds &= \int_{a_1}^{b_1} u_t(s) \, ds \\ &= \int_{a_1}^{b_1} h'_t(s) \, ds \\ &= h_t(b_1) - h_t(a_1) \quad (\text{by FTC}) \\ &= h(b_1, t) - h(a_1, t) \\ &= \partial_2 f(b_1, t) - \partial_2 f(a_1, t) \end{aligned}$$

Now, unfix t , and replace the value in $(*)$ to get

$$\begin{aligned}\int_S u(\vec{x}) d\vec{x} &= \int_{a_2}^{b_2} (\partial_2 f(b_1, t) - \partial_2 f(a_1, t)) dt \\ &= \int_{a_2}^{b_2} \partial_2 f(b_1, t) dt - \int_{a_2}^{b_2} \partial_2 f(a_1, t) dt\end{aligned}$$

For the function $v_{b_1} : [a_2, b_2]$ defined by $v_{b_1}(t) = f(b_1, t)$, we get $\partial_2 f(b_1, t) = v'_{b_1}(t)$. Hence

$$\int_{a_2}^{b_2} \partial_2 f(b_1, t) dt = \int_{a_2}^{b_2} v'_{b_1}(t) dt = v_{b_1}(b_2) - v_{b_1}(a_2) = f(b_1, b_2) - f(b_1, a_2)$$

Similarly, we have

$$\int_{a_1}^{b_1} \partial_2 f(a_1, t) dt = f(a_1, b_2) - f(a_1, a_2)$$

Thus, we can conclude that

$$\begin{aligned}\int_S u(\vec{x}) d\vec{x} &= (f(b_1, b_2) - f(b_1, a_2)) - (f(a_1, b_2) - f(a_1, a_2)) \\ &= f(b_1, b_2) - f(b_1, a_2) - f(a_1, b_2) + f(a_1, a_2)\end{aligned}$$

For the function $v(\vec{x})$, we can run a similar calculation with $\int_S v(\vec{x}) d\vec{x} = \int_{a_1}^{b_1} (\int_{a_2}^{b_2} v(s, t) dt) ds$. \square

Definition 19.4. Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{R}$ be a C^2 -function, and let \vec{a} be a point in A . Then one can consider the $n \times n$ matrix which, for every $1 \leq i, j \leq n$, has the (i, j) -entry equal to $\partial_i \partial_j f(\vec{a})$. This matrix is called the **Hessian matrix** of f at \vec{a} . We will denote it as $[Hf](\vec{a})$.

Remark 19.5. As a consequence of Proposition 19.3, the Hessian matrix $[Hf](\vec{a})$ is sure to be a symmetric matrix.

Theorem 19.6 (quadratic approximation for C^2 -functions). Let A be an open subset of \mathbb{R}^n . Let f be a C^2 -function, and let \vec{a} be a point in A . Denote $\vec{g} := \nabla f(\vec{a})$ and $H := [Hf](\vec{a}) \in \mathcal{M}_{n \times n}(\mathbb{R})$. Then we have

$$(Q\text{-Approx}) \quad \lim_{\vec{x} \rightarrow \vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \vec{g} \rangle - \langle H \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle}{\|\vec{x} - \vec{a}\|^2} = 0$$

Remark 19.7.

- (1) In the limit (Q-Approx) above, $\langle H \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle$ is a quadratic functions of the components of \vec{x} . Let us write explicitly $\vec{x} = (x_1, \dots, x_n)$, $\vec{a} = (a_1, \dots, a_n)$, and $H = [\partial_i \partial_j f(\vec{a})]_{i,j=1}^n$. Then we have

$$H \cdot (\vec{x} - \vec{a}) = \begin{bmatrix} \partial_1^2 f(\vec{a}) & \cdots & \partial_1 \partial_n f(\vec{a}) \\ \vdots & \ddots & \vdots \\ \partial_n \partial_1 f(\vec{a}) & \cdots & \partial_n^2 f(\vec{a}) \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n \partial_1 \partial_j f(\vec{a}) \cdot (x_j - a_j) \\ \vdots \\ \sum_{j=1}^n \partial_n \partial_j f(\vec{a}) \cdot (x_j - a_j) \end{bmatrix}$$

Hence, we get

$$\begin{aligned}\langle H \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle &= \sum_{i=1}^n \left(\sum_{j=1}^n \partial_i \partial_j f(\vec{a}) \cdot (x_j - a_j) \right) (x_i - a_i) \\ &= \sum_{i,j=1}^n \partial_i \partial_j f(\vec{a}) (x_i - a_i)(x_j - a_j)\end{aligned}$$

- (2) As with the limits (L-Approx) and (Diff-at- \vec{a}), we can think of the limit (Q-Approx) using sequences, or using a suitable ε - δ condition.
- (3) Using the ε - δ description and choosing a small $\varepsilon > 0$, we can find a suitable $\delta > 0$, and we have that

$$f(\vec{x}) \approx f(\vec{a}) + \langle \vec{x} - \vec{a}, \vec{g} \rangle + \langle H \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle$$

which is an improved approximation for $f(\vec{x})$ compared to (L-Approx) whenever $\vec{x} \in B(\vec{a}; \delta)$.

20 Points of local extremum

Definition 20.1 (points of local extremum). Let $A \subseteq \mathbb{R}^n$ and let $f : A \rightarrow \mathbb{R}$ be a function.

- A point $\vec{p} \in A$ is said to be a **point of local minimum** for f if there exists $r > 0$ such that $f(\vec{x}) \geq f(\vec{p})$ for all $\vec{x} \in B(\vec{p}; r) \cap A$.
- A point $\vec{q} \in A$ is said to be a **point of local maximum** for f if there exists $r > 0$ such that $f(\vec{x}) \leq f(\vec{q})$ for all $\vec{x} \in B(\vec{q}; r) \cap A$.
- We use the term **point of local extremum** for f to refer to a point of A which is either a point of local maximum or a point of local minimum.

Definition 20.2 (critical and regular points of a C^1 -function). Let $A \subseteq \mathbb{R}^n$ be an open set, let $f : A \rightarrow \mathbb{R}$ be a C^1 -function, and let \vec{a} be a point in A .

- If $\nabla f(\vec{a}) = \vec{0}$, then we say that \vec{a} is a **critical point** for f .
- If $\nabla f(\vec{a}) \neq \vec{0}$, then we say that \vec{a} is a **regular point** for f .

Proposition 20.3. Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \rightarrow \mathbb{R}$ be a C^1 -function. Let $\vec{a} \in A$ be a point of local extremum for f . Then \vec{a} is a critical point for f .

Proof. This is Assignment 9 Problem 2(a). □

Remark 20.4. Let $A \subseteq \mathbb{R}^n$ be open and let $f : A \rightarrow \mathbb{R}$ be a C^1 -function. Suppose we could determine the list of critical points for f on A . Which of these are points of local minimum or points or local maximum? The Hessian matrix could be of use, in the "second derivative test".

Remark 20.5. Let H be a symmetric matrix in $\mathcal{M}_{n \times n}(\mathbb{R})$, and let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigenvalues of H listed in decreasing order with multiplicities. (Note: All the eigenvalues of H are real, since it is a symmetric matrix with real entries.)

If $\lambda_n > 0$ (hence $\lambda_i > 0$ for all $1 \leq i \leq n$), then we say that H is **positive definite**.

If $\lambda_1 < 0$ (hence $\lambda_i < 0$ for all $1 \leq i \leq n$), then we say that H is **negative definite**.

We will use the following fact concerning quadratic functions associated with symmetric matrices that are positive or negative definite.

FACT. Use the notations as above.

- (1) If H is positive definite, then we have

$$(\square_+) \quad \langle H \cdot \vec{v}, \vec{v} \rangle \geq \lambda_n \|\vec{v}\|^2 (\geq 0), \quad \forall \vec{v} \in \mathbb{R}^n$$

- (2) If H is negative definite, then we have

$$(\square_-) \quad \langle H \cdot \vec{v}, \vec{v} \rangle \leq \lambda_1 \|\vec{v}\|^2 (\leq 0), \quad \forall \vec{v} \in \mathbb{R}^n$$

IDEA OF PROOF OF FACT. Consider the inequality (\square_+) . From linear algebra, we can find an orthonormal basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ of \mathbb{R}^n consisting of eigenvectors of H , where $H\vec{v}_i = \lambda_i \vec{v}_i$ for all

$1 \leq i \leq n$. Every vector $\vec{v} \in \mathbb{R}^n$ can be written as $\vec{v} = \sum_{i=1}^n \alpha_i \vec{v}_i$ for some coefficients $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Direct calculation then shows that

$$\begin{aligned} \langle H \cdot \vec{v}, \vec{v} \rangle &= \lambda_1 \alpha_1^2 + \dots + \lambda_n \alpha_n^2 \\ &\geq \lambda_n \alpha_1^2 + \dots + \lambda_n \alpha_n^2 \\ &= \lambda_n (\alpha_1^2 + \dots + \alpha_n^2) \\ &= \lambda_n \|\vec{v}\|^2 \end{aligned}$$

The proof of (\square_-) is similar.

Remark 20.6. Consider the "second derivative test" in one dimension.

Let $I \subseteq \mathbb{R}$ be an open interval and let $f : I \rightarrow \mathbb{R}$ be a twice differentiable function. Let $a \in I$, and suppose that $f'(a) = 0$ but $f''(a) \neq 0$. Then if $f''(a) > 0$, we have that a is a point of local minimum for f . If $f''(a) < 0$, then a is a point of local maximum for f .

We will generalize this to n dimensions by replacing the hypothesis: change $f'(a) = 0$ to $\nabla f(\vec{a}) = 0$ (that is, \vec{a} is a critical point for f), and change $f''(a) > 0$ (or $f''(a) < 0$) to the Hessian matrix $[Hf](\vec{a})$ being positive definite (or negative definite).

Proposition 20.7. Let A be an open subset of \mathbb{R}^n , let $f : A \rightarrow \mathbb{R}$ be a C^2 -function, and let \vec{a} be a critical point of A (for which we want to determine whether it is a point of local maximum, a point of local minimum, or neither). We consider the Hessian matrix $H := [Hf](\vec{a})$, and we assume that all the eigenvalues of H are non-zero. There are three possibilities:

- (a) If all the eigenvalues of H are positive (that is, H is positive definite), then \vec{a} is a point of local minimum for f .
- (b) If all the eigenvalues of H are negative (that is, H is negative definite), then \vec{a} is a point of local maximum for f .
- (c) If there exist eigenvalues λ', λ'' of H such that $\lambda' < 0 < \lambda''$, then \vec{a} is not a point of local extremum for f (it is called a **saddle point** for f).

Proof. We will prove (a), and the proof of (b) is analogous.

We will use the limit (Q-Approx) for f around the point \vec{a} , with the ε - δ condition. That is, given $\varepsilon > 0$, we can find $\delta > 0$ such that $B(\vec{a}; \delta) \subseteq A$ and for any $\vec{x} \in B(\vec{a}; \delta)$ we have

$$|f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}, \nabla f(\vec{a}) \rangle - \langle [Hf](\vec{a}) \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle| \leq \varepsilon \|\vec{x} - \vec{a}\|^2$$

But \vec{a} is a critical point for f , so $\nabla f(\vec{a}) = \vec{0}$. Hence, we can reduce the above inequality to

$$|f(\vec{x}) - f(\vec{a}) - \langle H \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle| \leq \varepsilon \|\vec{x} - \vec{a}\|^2$$

Take $\varepsilon = \lambda_n/2$. Let $\delta > 0$ be such that the above is satisfied. Then for $\vec{x} \in B(\vec{a}; \delta)$, we can write

$$-\varepsilon \|\vec{x} - \vec{a}\|^2 \leq f(\vec{x}) - f(\vec{a}) - \langle H \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle \leq \varepsilon \|\vec{x} - \vec{a}\|^2$$

We use the inequality on the left (the inequality on the right can be used for the proof of (b)). We obtain

$$\begin{aligned} f(\vec{x}) &\geq f(\vec{a}) + \langle H \cdot (\vec{x} - \vec{a}), \vec{x} - \vec{a} \rangle - \varepsilon \|\vec{x} - \vec{a}\|^2 \\ &\geq f(\vec{a}) + \lambda_n \|\vec{x} - \vec{a}\|^2 - \varepsilon \|\vec{x} - \vec{a}\|^2 & (\square_+) \\ &= f(\vec{a}) + \frac{\lambda_n}{2} \|\vec{x} - \vec{a}\|^2 \\ &\geq f(\vec{a}) \end{aligned}$$

Hence, for every $\vec{x} \in B(\vec{a}; \delta)$, we have $f(\vec{x}) \geq f(\vec{a})$, so \vec{a} is a point of local minimum for f .

The proof of (c) will be on Assignment 10. □

21 Operations with C^1 -functions and the chain rule

Notation 21.1. Let A be an open subset of \mathbb{R}^n . We denote

$$C^1(A, \mathbb{R}) := \{f : A \rightarrow \mathbb{R} \mid f \text{ is a } C^1\text{-function}\}$$

Proposition 21.2. Let A be an open subset of \mathbb{R}^n .

- (1) The set of functions $C^1(A, \mathbb{R})$ is stable under linear combinations. That is, for $f, g \in C^1(A, \mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$, we have that $\alpha f + \beta g \in C^1(A, \mathbb{R})$. Moreover, for every $1 \leq i \leq n$, we have the explicit formula

$$\partial_i (\alpha f + \beta g) = \alpha \partial_i f + \beta \partial_i g$$

- (2) The set of functions $C^1(A, \mathbb{R})$ is stable under multiplication. That is, if $f, g \in C^1(A, \mathbb{R})$, then $f \cdot g \in C^1(A, \mathbb{R})$, and for every $1 \leq i \leq n$ we have the explicit formula

$$\partial_i (f \cdot g) = (\partial_i f) \cdot g + f \cdot (\partial_i g)$$

Proof. This follows immediately from A8Q1. □

Remark 21.3. Recall the chain rule in one dimension. Let $A \subseteq \mathbb{R}$ be an open interval, and consider a C^1 -function $f : A \rightarrow \mathbb{R}$ (in this case, $f \in C^1(A, \mathbb{R})$ means that f is differentiable with continuous derivative). Suppose we are also given an open interval $B \subseteq \mathbb{R}$ and a function $g \in C^1(B, \mathbb{R})$ such that $g(B) \subseteq A$. Then we can consider the function $u := f \circ g : B \rightarrow \mathbb{R}$. The chain rule says that $u \in C^1(B, \mathbb{R})$, and for all $b \in B$ we have

$$u'(b) = f'(g(b)) \cdot g'(b)$$

We will generalize the chain rule to functions of several variables. The key idea is that the chain rule formula will remain the same, but we will multiply matrices instead of multiplying numbers.

What matrices will we be multiplying?

Also, there is an issue: for $f : A \rightarrow \mathbb{R}$ and $g : B \rightarrow \mathbb{R}$ with $A, B \subseteq \mathbb{R}^n$, we cannot even define the function $u = f \circ g$, since unless $n = 1$, it is impossible to have $g(B) \subseteq A$.

The resolution of the above issue is to consider functions that have values in some space \mathbb{R}^m (rather than being forced to have values in \mathbb{R}).

Definition 21.4. Let $m, n \in \mathbb{N}$ be two dimensions, let A be an open subset of \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}^m$ be a function. For every $\vec{x} \in A$, write

$$f(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x})) \in \mathbb{R}^m$$

and consider the component functions $f_1, \dots, f_m : A \rightarrow \mathbb{R}$ defined in this way. If $f_1, \dots, f_m \in C^1(A, \mathbb{R})$, then we say that $f \in C^1(A, \mathbb{R}^m)$.

Definition 21.5. Let A be an open subset of \mathbb{R}^n , and let $f \in C^1(A, \mathbb{R}^m)$ with components $f_1, \dots, f_m \in C^1(A, \mathbb{R})$. For every $\vec{a} \in A$, the matrix

$$[Jf](\vec{a}) = \begin{bmatrix} \partial_1 f_1(\vec{a}) & \cdots & \partial_n f_1(\vec{a}) \\ \vdots & \ddots & \vdots \\ \partial_1 f_m(\vec{a}) & \cdots & \partial_n f_m(\vec{a}) \end{bmatrix} \in \mathcal{M}_{m \times n}(\mathbb{R})$$

is called the **Jacobian matrix** of f at \vec{a} .

Remark 21.6. Two special cases of the preceding definition:

- (1) The case where $m = 1$ (but general $n \in \mathbb{N}$). In this case, the function f has only one component, $f_1 = f : A \rightarrow \mathbb{R}$, and the Jacobian matrix is

$$[Jf](\vec{a}) := \begin{bmatrix} \partial_1 f(\vec{a}) & \cdots & \partial_n f(\vec{a}) \end{bmatrix} \in \mathcal{M}_{1 \times n}(\mathbb{R})$$

That is, $[Jf](\vec{a})$ is just the gradient vector $\nabla f(\vec{a})$ viewed as a row-matrix of size $1 \times n$.

- (2) The case where $n = 1$ (but general $m \in \mathbb{N}$). We have that $A \subseteq \mathbb{R}$ is open. Then $f : A \rightarrow \mathbb{R}^m$ is a C^1 -path, defined by

$$f(t) = (f_1(t), \dots, f_m(t)) \in \mathbb{R}^m$$

The functions $f_1, \dots, f_m : A \rightarrow \mathbb{R}$ are differentiable with continuous derivative. The Jacobian matrix in this case is

$$[Jf](t) = \begin{bmatrix} f'_1(t) \\ \vdots \\ f'_m(t) \end{bmatrix} \in \mathcal{M}_{m \times 1}(\mathbb{R})$$

Hence, $[Jf](t)$ is the velocity vector of f at time t , given by $f'(t) = (f'_1(t), \dots, f'_m(t)) \in \mathbb{R}^m$, viewed as a column matrix.

Theorem 21.7 (Chain Rule). Let $m, n, p \in \mathbb{N}$ be three dimensions. Suppose we are given a function $f \in C^1(A, \mathbb{R}^m)$, where $A \subseteq \mathbb{R}^n$ is open, and a function $g \in C^1(B, \mathbb{R}^p)$, where $B \subseteq \mathbb{R}^p$ is open. Suppose moreover that $g(B) \subseteq A$. Now, we can consider the function $u := f \circ g : B \rightarrow \mathbb{R}^m$, defined by $u(\vec{b}) := f(g(\vec{b}))$ for every $\vec{b} \in B$. Then $u \in C^1(B, \mathbb{R}^m)$, and we have the formula

$$(CR) \quad [Ju](\vec{b}) = [Jf](g(\vec{b})) \cdot [Jg](\vec{b}) \quad \forall \vec{b} \in B$$

Idea of Proof. Compose linear combinations.

The linear approximation of f says that for every $\vec{a} \in A$, we have

$$(\diamond) \quad \lim_{\vec{x} \rightarrow \vec{a}} \frac{1}{\|\vec{x} - \vec{a}\|} (f(\vec{x}) - f(\vec{a}) - [Jf](\vec{a}) \cdot (\vec{x} - \vec{a})) = \vec{0} \in \mathbb{R}^m$$

We can check this with sequences. Take $(\vec{x}_k)_{k=1}^\infty$ in A such that $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{a}$, but $\vec{x}_k \neq \vec{a}$ for all $k \in \mathbb{N}$. We know that for every $1 \leq i \leq n$,

$$\lim_{k \rightarrow \infty} \frac{1}{\|\vec{x}_k - \vec{a}\|} (f_i(\vec{x}_k) - f_i(\vec{a}) - \langle \nabla f_i(\vec{a}), \vec{x}_k - \vec{a} \rangle) = 0$$

This is (L-Approx) from Lecture 17 for the component function f_i . Each of the m limits goes to 0, so we have

$$\lim_{k \rightarrow \infty} \frac{1}{\|\vec{x}_k - \vec{a}\|} (f(\vec{x}_k) - f(\vec{a}) - [Jf](\vec{a}) \cdot (\vec{x}_k - \vec{a})) = \vec{0} \in \mathbb{R}^m$$

which is exactly the limit (\diamond) described with sequences.

Likewise, the linear approximation for g says that for every $\vec{b} \in B$, we have

$$(\diamond\diamond) \quad \lim_{\vec{y} \rightarrow \vec{b}} \frac{1}{\|\vec{y} - \vec{b}\|} (g(\vec{y}) - g(\vec{b}) - [Jg](\vec{b}) \cdot (\vec{y} - \vec{b})) = \vec{0} \in \mathbb{R}^p$$

Combining (\diamond) and $(\diamond\diamond)$, we find a linear approximation for the composite function $u = f \circ g$. The linear approximation for u can then be used to prove that u is a C^1 -function, with the Jacobian matrix $[Ju](\vec{b})$ of the form stated in the equation (CR) of the theorem. \square

Example 21.8. Let $A = (0, \infty)^3 \subseteq \mathbb{R}^3$, and let $u : A \rightarrow \mathbb{R}$ be defined by

$$u(x_1, x_2, x_3) = (x_1^{x_2})^{x_3}, \quad \text{for } x_1, x_2, x_3 \in (0, \infty)$$

What is the formula for $\partial_2 u(x_1, x_2, x_3)$?

Let $g : (0, \infty)^3 \rightarrow (0, \infty)^2$ be defined by

$$g(x_1, x_2, x_3) = (x_1^{x_2}, x_3)$$

and let $f : (0, \infty)^2 \rightarrow \mathbb{R}$ be defined by

$$f(s, t) = s^t$$

Then we have $u = f \circ g : (0, \infty)^3 \rightarrow \mathbb{R}$. Writing (CR), we have

$$[Ju](x_1, x_2, x_3) = [Jf](x_1^{x_2}, x_3) \cdot [Jg](x_1, x_2, x_3)$$

That is,

$$\begin{bmatrix} \partial_1 u(x_1, x_2, x_3) & \partial_2 u(x_1, x_2, x_3) & \partial_3 u(x_1, x_2, x_3) \end{bmatrix} = \begin{bmatrix} \partial_1 f(x_1^{x_2}, x_3) & \partial_2 f(x_1^{x_2}, x_3) \end{bmatrix} \begin{bmatrix} x_2 \cdot x_1^{x_2-1} & x_1^{x_2} \cdot \ln x_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, the formula is

$$\begin{aligned} \partial_2 u(x_1, x_2, x_3) &= \partial_1 f(x_1^{x_2}, x_3) \cdot x_1^{x_2} \cdot \ln x_1 \\ &= [x_3 \cdot (x_1^{x_2})^{x_3-1}] \cdot x_1^{x_2} \cdot \ln x_1 \\ &= x_3 \ln x_1 \cdot (x_1^{x_2})^{x_3} \end{aligned}$$

Corollary 21.9. (special case of Chain Rule where $m = p = 1$ and $n \in \mathbb{N}$ is arbitrary)

Let A be an open subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a C^1 -function. On the other hand, let $I \subseteq \mathbb{R}$ be an open interval and let $\gamma : I \rightarrow \mathbb{R}^n$ be a C^1 -path such that $\gamma(t) \in A$ for all $t \in I$. We consider the composite function $u : I \rightarrow \mathbb{R}$ defined by $u(t) := f(\gamma(t))$ for all $t \in I$. Then u is a C^1 -function, and we have the formula

$$u'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle, \quad \forall t \in I$$

Proof. Write the formula (CR) explicitly, and read it as an inner product. □

22 Lagrange multipliers

Example 22.1. Let $S := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$, and let $f : S \rightarrow \mathbb{R}$ be defined by $f(x, y, z) := xy^2z^2$ for $(x, y, z) \in S$. Since f is a C^1 -function, and in particular continuous on the compact set S , the Extreme Value Theorem tells us that f must have points of global maximum on S .

Our goal: What is the global maximum of f on S , and where is it achieved?

The method used in Lecture 20 will not work here. There are some critical points lying on S , but none of them can be a point of global maximum. This is because we will have that at least one of x , y , or z is 0, and hence $f(x, y, z) = 0$. We can easily see that this choice of (x, y, z) is not a point of global maximum.

Some terminology: The function f to be maximized is called an *objective function*. We need to maximize f on a level set of the form

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid \varphi(x, y, z) = 1\}$$

where $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$ is defined by $\varphi(x, y, z) = x^2 + y^2 + z^2$. This function φ is called a *constraint function* (since we are maximizing f subject to a constraint provided by φ).

Notation 22.2.

- *Constraint function:* We have an open set $\Omega \subseteq \mathbb{R}^n$, a C^1 -function $\varphi : \Omega \rightarrow \mathbb{R}$, and a level set

$$L := \{\vec{x} \in \Omega \mid \varphi(\vec{x}) = c\}$$

where c is one of the values assumed by φ . We will make the additional assumption that $\nabla\varphi(\vec{x}) \neq \vec{0}$ for all $\vec{x} \in L$ (all points in L are regular points for φ).

- *Objective function:* a C^1 -function $f : A \rightarrow \mathbb{R}$, where $A \subseteq \mathbb{R}^n$ is open, such that $A \supseteq L$.
- *Our goal:* Find the points of local extremum for the restriction of f to L .

Theorem 22.3 (Lagrange multipliers). Use the notation as above. Let $\vec{a} \in L$ be a point of local extremum for f on L . (That is, there exists $r > 0$ such that $f(\vec{x}) \geq f(\vec{a})$ for all $\vec{x} \in B(\vec{a}; r) \cap L$, or there exists $r > 0$ such that $f(\vec{x}) \leq f(\vec{a})$ for all $\vec{x} \in B(\vec{a}; r) \cap L$.) Then there exists $\lambda \in \mathbb{R}$ such that $\nabla f(\vec{a}) = \lambda \cdot \nabla\varphi(\vec{a})$.

Example 22.4 (Finalizing Example 22.1). We use the framework in Notation 22.2, where the constraint function is $\varphi(x, y, z) = x^2 + y^2 + z^2$, the level set is the sphere $S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$, and the objective function is $f(x, y, z) = xy^2z^2$. A point $\vec{a} = (x, y, z) \in S$ which maximizes the value of f must satisfy an equation of the form

$$(y^2z^2, 2xyz^2, 2xy^2z) = \lambda \cdot (2x, 2y, 2z)$$

for some $\lambda \in \mathbb{R}$. Thus, we get a system of four equations with unknowns x, y, z, λ :

$$\begin{cases} y^2z^2 &= 2\lambda x \\ 2xyz^2 &= 2\lambda y \\ 2xy^2z &= 2\lambda z \\ x^2 + y^2 + z^2 &= 1 \end{cases}$$

Solving the system gives us four points of global maximum on the sphere, which are of the form $(\sqrt{1/5}, \pm\sqrt{2/5}, \pm\sqrt{2/5})$.

Proposition 22.5. Consider the framework from Notation 22.2. Fix a point $\vec{a} \in L$, and denote $T := \{\vec{v} \in \mathbb{R}^n \mid \vec{v} \text{ is a tangent vector to } L \text{ at } \vec{a}\}$. Then we have

$$(\spadesuit) \quad T = \{\vec{v} \in \mathbb{R}^n \mid \vec{v} \perp \nabla\varphi(\vec{a})\}$$

Idea of Proof. The inclusion " \subseteq " comes from the fact that a tangent vector and a normal vector are always orthogonal to each other. This was proved in A9Q3. On the other hand, the inclusion " \supseteq " says something about the structure of L , namely that L must have a lot of tangent vectors at \vec{a} . This is a non-trivial consequence of the fact that the point \vec{a} is known to be regular for φ . \square

Corollary 22.6. Consider the same framework from Proposition 22.5, where we fix a point $\vec{a} \in L$ and we denote T as the set of vectors that are tangent to L at \vec{a} . Let $\vec{w} \in \mathbb{R}^n$ be a vector which is orthogonal to every $\vec{v} \in T$. Then there exists $\lambda \in \mathbb{R}$ such that $\vec{w} = \lambda \cdot \nabla\varphi(\vec{a})$.

Proof. Consider the 1-dimensional vector space

$$V := \{\lambda \cdot \nabla\varphi(\vec{a}) \mid \lambda \in \mathbb{R}\} \subseteq \mathbb{R}^n$$

The equality (\spadesuit) from Proposition 22.5 tells us that $T = V^\perp$ (that is, T is the *orthogonal complement* of the subspace V). On the other hand, the hypothesis tells us that $\vec{w} \in T^\perp$. Thus, we get

$$\vec{w} \in T^\perp = (V^\perp)^\perp = V$$

Hence, $\vec{w} \in V$, which is precisely what we wanted. \square

Proof of Theorem 22.3. Suppose that \vec{a} is a point of local maximum for f on L (the case where \vec{a} is a point of local minimum is similar). By Corollary 22.6, it suffices to show that $\nabla f(\vec{a}) \perp \vec{v}$ for every $\vec{v} \in T$, where T is the set of tangent vectors to L at \vec{a} .

Fix $\vec{v} \in T$. By definition, we can find an open interval $I \subseteq \mathbb{R}$ and a C^1 -path $\gamma : I \rightarrow \mathbb{R}^n$ such that $\gamma(t) \in L$ for all $t \in I$, and such that $0 \in I$, with $\gamma(0) = \vec{a}$ and $\gamma'(0) = \vec{v}$.

Define $u : I \rightarrow \mathbb{R}$ by $u(t) = f(\gamma(t))$ for all $t \in I$. By Corollary 21.9, we have that

$$u'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle \quad \forall t \in I$$

In particular, we get $u'(0) = \langle \nabla f(\vec{a}), \vec{v} \rangle$. Now, we only need to check that $u'(0) = 0$.

CLAIM. 0 is a point of local maximum for u on I .

VERIFICATION OF CLAIM. Since \vec{a} is a point of local maximum for f , there exists $r > 0$ such that whenever $\vec{x} \in B(\vec{a}; r) \cap L$, it follows that $f(\vec{x}) \leq f(\vec{a})$. Since γ is continuous at 0, we can find $\delta > 0$ such that $(-\delta, \delta) \subseteq I$ and such that $t \in (-\delta, \delta)$ implies $\|\gamma(t) - \gamma(0)\| < r$. Hence:

$$t \in (-\delta, \delta) \Rightarrow \gamma(t) \in B(\vec{a}; r) \cap L \Rightarrow f(\gamma(t)) \leq f(\vec{a}) \Rightarrow u(t) \leq u(0)$$

This proves the claim. Thus $u'(0) = 0$, so we are done. \square

23 Change of variable for multiple integrals

Remark 23.1. In this lecture, we will discuss a method of calculating an integral $\int_A f(\vec{x}) d\vec{x}$ in the following situation:

- A is a bounded open subset of \mathbb{R}^n which is Jordan measurable (that is, the boundary $\text{bd}(A)$ is a null-set).
- The function $f : A \rightarrow \mathbb{R}$ is bounded and integrable.
- There is a set $B \subseteq \mathbb{R}^n$ (which is also bounded, open, and Jordan measurable) and a bijective transformation $G : B \rightarrow A$ such that the substitution

$$\vec{x} = G(\vec{y}) \quad (\text{with } \vec{x} \in A, \vec{y} \in B)$$

appears to simplify the function $f(\vec{x})$ which we want to integrate.

In other words, instead of integrating $f(\vec{x})$ with \vec{x} in the set A , we would prefer to integrate $f(G(\vec{y}))$ with \vec{y} in the set B .

Definition 23.2 (Jacobian determinant of a C^1 -transformation). Let $B \subseteq \mathbb{R}^n$ be an open set and let G be a function in $C^1(B, \mathbb{R}^n)$. Recall from Lecture 21 that for every $\vec{b} \in B$, we have a Jacobian matrix $[JG](\vec{b})$, which is an $n \times n$ matrix with real entries. The real number $\det([JG](\vec{b}))$ is called the **Jacobian determinant** of G at the point \vec{b} . In short, we also call this the Jacobian of G at \vec{b} .

We are interested in the absolute value of the Jacobian determinant, and we will use the notation $|J|_G : B \rightarrow \mathbb{R}$ for the function defined by

$$(*) \quad |J|_G(\vec{y}) := |\det([JG](\vec{y}))|, \quad \vec{y} \in B$$

Remark 23.3. Let $B \subseteq \mathbb{R}^n$ and $f \in C^1(B, \mathbb{R}^n)$ be as above. Then the function $|J|_G : B \rightarrow \mathbb{R}$ defined by the formula $(*)$ is continuous. This is because we have polynomial formulas which express $\det([JG](\vec{y}))$ in terms of the entries of $[JG](\vec{y})$, and the entries are continuous functions of \vec{y} .

Definition 23.4 (C^1 -diffeomorphism). Let $G : B \rightarrow A$ be a bijective transformation, where A and B are open subsets of \mathbb{R}^n . Let us also consider the inverse transformation $G^{-1} : A \rightarrow B$. If both G and G^{-1} are C^1 -functions, then we say that G is a **C^1 -diffeomorphism** between A and B .

Theorem 23.5 (Substitution in a multiple integral). Consider the framework from Remark 23.1, where we now assume that $G : B \rightarrow A$ is a C^1 -diffeomorphism. We assume moreover that the function $|J|_G : B \rightarrow \mathbb{R}$ defined by the formula $(*)$ in Definition 23.2 is bounded.

Let $f : A \rightarrow \mathbb{R}$ be a bounded integrable function. Then $g : B \rightarrow \mathbb{R}$ defined by

$$g(\vec{y}) := f(G(\vec{y})) \cdot |J|_G(\vec{y}), \quad \vec{y} \in B$$

is bounded and integrable, and we have the equality of integrals $\int_A f = \int_B g$. That is, we have the formula

$$(\text{Change-Var}) \quad \int_A f(\vec{x}) d\vec{x} = \int_B f(G(\vec{y})) \cdot |J|_G(\vec{y}) d\vec{y}$$

Idea of Proof. The proof is long and was not covered in lecture. However, the main idea is that the function $|J|_G$ keeps track of how G distorts volumes between subsets of B and subsets of A . \square

Example 23.6 (Polar coordinates in \mathbb{R}^2). Let the dimension be $n = 2$, and fix some radii $r_2 > r_1 \geq 0$. We consider the transformation $G : B \rightarrow A$, where A , B , and the transformation G are defined as follows:

- $B := \{(r, \theta) \in \mathbb{R}^2 \mid r_1 < r < r_2, 0 < \theta < 2\pi\}$, a rectangle.
- $A := \{(s, t) \in \mathbb{R}^2 \mid r_1 < \sqrt{s^2 + t^2} < r_2\} \setminus \{(s, 0) \mid r_1 < s < r_2\}$, an annulus with a ray removed.
- $G : B \rightarrow A$ is defined by $G((r, \theta)) = (r \cos \theta, r \sin \theta)$, for $(r, \theta) \in B$.

Then G is a C^1 -diffeomorphism between B and A (this can be checked using trigonometry). An immediate calculation shows that the Jacobian of G at a point $(r, \theta) \in B$ is

$$\det \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} = r$$

Hence, the function $|J|_G : B \rightarrow \mathbb{R}$ from Definition 23.2 is simply defined by $|J|_G((r, \theta)) = r$.

Corollary 23.7 (Integration in polar coordinates). Let $r_2 > r_1 \geq 0$ be two radii, and let $f : C \rightarrow \mathbb{R}$ be a bounded integrable function, where $C = \{(s, t) \in \mathbb{R}^2 \mid r_1 \leq \sqrt{s^2 + t^2} \leq r_2\}$. Then we have

$$(\text{Pol-Coord}) \quad \int_C f(\vec{x}) \, d\vec{x} = \int_B f(r \cos \theta, r \sin \theta) \cdot r \, dr \, d\theta$$

where B is the rectangle $(r_1, r_2) \times (0, 2\pi)$.

Corollary 23.8. The Gaussian integral $\int_{-\infty}^{\infty} e^{-x^2} \, dx$ has value $\sqrt{\pi}$.

Proof. From A9Q5, we have that

$$(*) \quad \left(\int_{-\infty}^{\infty} e^{-x^2} \, dx \right)^2 = \lim_{k \rightarrow \infty} \int_{D_k} e^{-\|\vec{x}\|^2} \, d\vec{x}$$

where $D_k = \{\vec{x} \in \mathbb{R}^2 \mid \|\vec{x}\| < k\}$. For every $k \in \mathbb{N}$, we evaluate the integral on the right-hand side of the above equation $(*)$ by using the formula for integration in polar coordinates. The radii are now $r_1 = 0$ and $r_2 = k$. We also note that from $\vec{x} = (r \cos \theta, r \sin \theta)$, it follows that $\|\vec{x}\|^2 = r^2$. Hence:

$$\begin{aligned} \int_{D_k} e^{-\|\vec{x}\|^2} \, d\vec{x} &= \int_0^k \int_0^{2\pi} e^{-r^2} \cdot r \, d(r, \theta) \\ &= \int_0^k \left(\int_0^{2\pi} r e^{-r^2} \, d\theta \right) dr \\ &= \int_0^k 2\pi r e^{-r^2} \, dr \\ &= \pi \int_0^k (-e^{-r^2})' \, dr \\ &= \pi(1 - e^{-k^2}) \end{aligned}$$

Taking the limit and plugging it back into $(*)$, we get

$$\left(\int_{-\infty}^{\infty} e^{-x^2} \, dx \right)^2 = \lim_{k \rightarrow \infty} \int_{D_k} e^{-\|\vec{x}\|^2} \, d\vec{x} = \pi$$

and the corollary follows. □

Additional definitions from assignments

Definition (Semi-ring of sets). Let X be a set and let \mathcal{S} be a collection of subsets of X . We say that \mathcal{S} is a **semi-ring** of subsets of X when it fulfills the following conditions:

$$(S\text{-RS-1}) \quad \emptyset \in \mathcal{S}.$$

$$(S\text{-RS-2}) \quad \text{Whenever } S, T \in \mathcal{S}, \text{ it follows that } S \cap T \in \mathcal{S}.$$

$$(S\text{-RS-3}) \quad \text{For every } S, T \in \mathcal{S}, \text{ there exist } p \in \mathbb{N} \text{ and } U_1, \dots, U_p \in \mathcal{S} \text{ such that } U_i \cap U_j = \emptyset \text{ for } i \neq j \text{ and such that } S \setminus T = \cup_{i=1}^p U_i.$$

Definition (Oscillation). Let A be a non-empty set and let $f : A \rightarrow \mathbb{R}$ be a bounded function. For every non-empty subset $B \subseteq A$, we define the **oscillation** of f on B to be the number

$$\text{osc}_B(f) := \sup_B(f) - \inf_B(f)$$

Definition (Convex).

- (1) Let \vec{x} and \vec{y} be two vectors in \mathbb{R}^n . A vector of the form

$$\vec{v} = (1-t)\vec{x} + t\vec{y}, \quad 0 \leq t \leq 1$$

is called a **convex combination** of \vec{x} and \vec{y} . The set of convex combinations of \vec{x} and \vec{y} is called the **line-segment** with endpoints at \vec{x} and \vec{y} .

- (2) A subset $A \subseteq \mathbb{R}^n$ is said to be **convex** when it has the following property:

$$(Conv) \quad \begin{cases} \text{For every } \vec{x}, \vec{y} \in A \text{ and every } 0 \leq t \leq 1, \\ \text{it follows that } (1-t)\vec{x} + t\vec{y} \text{ still belongs to } A. \end{cases}$$

Definition (Lipschitz function). Let $n \in \mathbb{N}$ be a dimension, let A be a non-empty set in \mathbb{R}^n , and let $f : A \rightarrow \mathbb{R}$ be a function.

- (1) Let c be a number in $[0, \infty)$. If we have that

$$(c\text{-LIP}) \quad |f(\vec{x}_1) - f(\vec{x}_2)| \leq c\|\vec{x}_1 - \vec{x}_2\| \quad \forall \vec{x}_1, \vec{x}_2 \in A$$

then we say that f is **c-Lipschitz** on A .

- (2) We will simply say that f is **Lipschitz** on A to mean that there exists $c \in [0, \infty)$ such that f is c -Lipschitz on A .

Definition (Pre-additive). Let X be a set and \mathcal{S} be a semi-ring of subsets of X . A function $\mu_0 : \mathcal{S} \rightarrow [0, \infty)$ is said to be **pre-additive** when it has the following property:

$$(PRE\text{-}ADD) \quad \left(\begin{array}{l} S, S_1, \dots, S_k \in \mathcal{S} \\ \text{are such that } S = S_1 \cup \dots \cup S_k \\ \text{and such that } S_i \cap S_j = \emptyset \text{ for } 1 \leq i < j \leq k \end{array} \right) \Rightarrow \mu_0(S) = \sum_{i=1}^k \mu_0(S_i)$$

Definition (Path). Let $I \subseteq \mathbb{R}$ be an interval.

- (1) A function $\gamma : I \rightarrow \mathbb{R}^n$ (for some $n \in \mathbb{N}$) is said to be a **path**.
- (2) We have a notion of what it means for a function $\gamma : I \rightarrow \mathbb{R}^n$ to be continuous on I . In this situation, we say that γ is a **continuous path**.

Definition (Path-connected). A set $A \subseteq \mathbb{R}^n$ is said to be **path-connected** if for every $\vec{x}, \vec{y} \in A$, it is possible to find a continuous path $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ such that $\gamma(0) = \vec{x}$, $\gamma(1) = \vec{y}$, and $\gamma(t) \in A$ for all $0 \leq t \leq 1$.

Definition (Polygonally connected).

- (1) For $\vec{x}, \vec{y} \in \mathbb{R}^n$, we denote $\text{Co}(\vec{x}, \vec{y}) := \{(1-t)\vec{x} + t\vec{y} \mid 0 \leq t \leq 1\} \subseteq \mathbb{R}^n$. The set $\text{Co}(\vec{x}, \vec{y})$ is called the **line segment** in \mathbb{R}^n with endpoints at \vec{x} and \vec{y} .
- (2) A set $A \subseteq \mathbb{R}^n$ is said to be **polygonally connected** if for every $\vec{a}, \vec{b} \in A$, it is possible to find points $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_k \in A$ such that $\vec{x}_0 = \vec{a}$, $\vec{x}_k = \vec{b}$, and such that $\text{Co}(\vec{x}_{j-1}, \vec{x}_j) \subseteq A$ for every $1 \leq j \leq k$.

Definition (Strict contraction). Let $A \subseteq \mathbb{R}^n$ and let $T : A \rightarrow A$ be a function. If there exists a constant c with $0 < c < 1$ such that

$$\|T(\vec{x}) - T(\vec{y})\| \leq c \cdot \|\vec{x} - \vec{y}\| \quad \forall \vec{x}, \vec{y} \in A$$

then we say that T is a **strict contraction**.

Definition (Differentiable path, velocity vector, C^1 -path). Let $I \subseteq \mathbb{R}$ be an open interval and let $\gamma : I \rightarrow \mathbb{R}^n$ be a path. For every $1 \leq j \leq n$, let $\gamma_j : I \rightarrow \mathbb{R}$ denote the j -th component of γ . That is, we have

$$\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t)), \quad t \in I$$

We define the following:

- If all n component functions of γ are differentiable functions from I to \mathbb{R} (in the 1-dimensional sense), then we say that γ is a **differentiable path**. If this is the case, then for every $t \in I$, we can consider the vector

$$\gamma'(t) := (\gamma'_1(t), \dots, \gamma'_n(t)) \in \mathbb{R}^n$$

which is called the **velocity vector** of the path γ at time t .

- If all n component functions of γ are C^1 -functions from I to \mathbb{R} (that is, the derivative $\gamma'_j : I \rightarrow \mathbb{R}$ exists and is continuous for every $1 \leq j \leq n$), then we say that γ is a C^1 -path.

Definition (Level set). Let A be a subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a function. Let $c \in \mathbb{R}$ be one of the values attained by f on A . The **level set** of f corresponding to the value c is defined as

$$L := \{\vec{x} \in A \mid f(\vec{x}) = c\}$$

Definition (Normal vector, tangent vector). Let A be an open subset of \mathbb{R}^n and let $f : A \rightarrow \mathbb{R}$ be a C^1 -function. Let $c \in \mathbb{R}$ be one of the values attained by f on A , and let L be the level set of f corresponding to the value c . We fix a point $\vec{a} \in L$ which is regular for f (that is, $f(\vec{a}) = c$ and $\nabla f(\vec{a}) \neq \vec{0}$).

- Any non-zero vector \vec{w} which is a scalar multiple of $\nabla f(\vec{a})$ is said to be a **normal vector** to L at the point \vec{a} .
- Let \vec{v} be a vector in \mathbb{R}^n . We say that \vec{v} is a **tangent vector** to L at the point \vec{a} to mean that it is possible to find an open interval $I \subseteq \mathbb{R}$ and a C^1 -path $\gamma : I \rightarrow \mathbb{R}^n$ such that the following conditions hold:
 - (i) $\gamma(t) \in L$ for all $t \in I$.
 - (ii) $0 \in I$, and we have $\gamma(0) = \vec{a}$ and $\gamma'(0) = \vec{v}$.