Calculus II

Notes from TAU Course with Additional Information
Lecturer: Daniel Tsodikovich

Gabriel Domingues

July 17, 2020

Contents

1 Functional Analysis

1.1 Sequence of Functions

Definition 1.1.1 (Supremum)**.** *For a set $A \subset \mathbb{R}$, if A is bounded from above, the supremum of A is the lowest upper bound (denoted* $\sup A$*). Otherwise* $\sup A = \infty$*. That is, $M = \sup(A)$ iff $\forall\, a \in A\,,\, a \leq M$.*

Definition 1.1.2 (Sequence of Functions)**.** *A sequence of function is a family $\{f_n\}_{n\in\mathbb{N}}$ where $\forall\, n \in \mathbb{N}\,,\, f_n : \mathcal{I} \subseteq \mathbb{R} \to \mathbb{R}$. Observe, the interval $\mathcal{I}$ is the same domain for all f_n in the sequence.*

Definition 1.1.3 (Pointwise Convergence)**.** *For a sequence $\{f_n\}_{n\in\mathbb{N}}$, we say f_n converges pointwise to f (denoted $f_n \to f$) if, $\forall\, x_0 \in \mathcal{I}\,,\, f_n(x_0) \to f(x_0)$. That is, if f is defined explicitly $\forall\, x_0 \in \mathcal{I}\,,\, f(x_0) := \lim\limits_{n\to\infty} f_n(x_0)$.*

Remark 1.1.4. *The pointwise limit is unique, since $\lim\limits_{n\to\infty} f_n(x_0)$ is unique.*

Remark 1.1.5. *That pointwise limit of continuous functions can be discontinuous. For illustration, take $f_n : [0,1] \to \mathbb{R}$ where $f_n(x) = x^n$. Then, $f_n(x) \to f(x)$ where:*

$$f(x) = \begin{cases} 0 & \text{if } x < 1 \\ 1 & \text{if } x = 1 \end{cases}$$

which is then discontinuous.

Definition 1.1.6 (Uniform Convergence)**.** *For a sequence $\{f_n\}_{n\in\mathbb{N}}$, we say f_n converges pointwise to f (denoted $f_n \xrightarrow{u} f$) if,*

$$\forall\, \epsilon > 0\,,\, \exists\, N \in \mathbb{N} : \forall\, n \geq N\,,\, \forall\, x \in \mathcal{I}\,,\, |f(x) - f_n(x)| < \epsilon$$

Lemma 1.1.7. *(UC $\Rightarrow$ PC) If $f_n \xrightarrow{u} f$, then $f_n \to f$.*

Proof. If $f_n \xrightarrow{u} f$, then $\forall\, x \in \mathcal{I}\,,\, f(x) = \lim\limits_{n\to\infty} f_n(x)$, by definition. $\square$

Definition 1.1.8 (Vector Space of Functions)**.** $\{f \mid f : \mathcal{I} \to \mathbb{R}\}$ *is a vector space over $\mathbb{R}$ with pointwise addition and scalar multiplication: $(f+g)(x) = f(x) + g(x)$ and $(\alpha \cdot f) = \alpha \cdot f(x)$*

Definition 1.1.9 (Uniform Norm)**.** *We define the following norm for functions $f : \mathcal{I} \to \mathbb{R}$:*

$$\|f\|_\infty = \sup_{x\in\mathcal{I}} |f(x)|$$

which we can check is a norm. Also, f is bounded iff $\|f\|_\infty < \infty$.

Remark 1.1.10. *The idea of using $\|\cdot\|_\infty$ is to bound independent of x, since $\|f-g\|_\infty$ only depends on f and g. We can substitute: $\|f-g\|_\infty \leq \epsilon \Leftrightarrow \forall\, x \in \mathcal{I}\,,\, |f(x)-g(x)| \leq \epsilon$ (cf. 1.1.1).*

Remark 1.1.11 (Banach Algebra)**.** $\forall\, f, g : \mathcal{I} \to \mathbb{R}\,,\, \|f\cdot g\|_\infty \leq \|f\|_\infty \cdot \|g\|_\infty$, *where $\cdot$ is pointwise multiplication.*

Lemma 1.1.12. *$f_n \xrightarrow{u} f$ iff $\|f-f_n\|_\infty \to 0$*

Proof. We prove each direction:

($\Rightarrow$) $\forall\, \epsilon > 0\,,\, \exists\, N \in \mathbb{N} : \forall\, n \geq N\,,\, \forall\, x \in \mathcal{I}\,,\, |f(x) - f_n(x)| < \epsilon/2$ (cf. 1.1.3). Taking the supremum on $x \in \mathcal{I}$, $\forall\, n \geq N\,,\, \|f-f_n\|_\infty \leq \epsilon/2 < \epsilon$. That is, $\|f-f_n\|_\infty \to 0$ by definition.

($\Leftarrow$) $\|f-f_n\|_\infty \to 0 \Leftrightarrow \forall\, \epsilon > 0\,,\, \exists\, n \in \mathbb{N} : \forall\, n \geq N\,,\, \|f-f_n\|_\infty < \epsilon$. Then, $\forall\, n \geq N\,,\, \forall\, x \in \mathcal{I}\,,\, |f(x)-f_n(x)| < \epsilon$ (cf. 1.1.1, 1.1.10).

Hence, $\|\cdot\|_\infty$ is the norm that defines uniform continuity. ☐

Lemma 1.1.13. *If $f_n \to f$, then $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly iff $f_n \xrightarrow{u} f$.*

Proof. We prove each direction:

($\Rightarrow$) If $f_n \xrightarrow{u} g$, by 1.1.7 $f_n \to g$ and by 1.1.4, $g = f$.

($\Leftarrow$) If $f_n \xrightarrow{u} f$, then $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly.

Hence, if $f_n \to f$, $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly iff $\lim\limits_{n\to\infty} \|f-f_n\|_\infty = 0$. ☐

Remark 1.1.14. *If we change our domain, we way have UC. Going back to the example of $f_n(x) = x^n$, we get $\{f_n\}_{n\in\mathbb{N}} \xrightarrow{u} f \equiv 0$ in $\mathcal{I} = [0,t]$ for $t < 1$ since $\|f-f_n\|_\infty = \sup\limits_{x\in\mathcal{I}} |x|^n = t^n \to 0$.*

Lemma 1.1.15 (Bounded Limit)**.** *If $f_n \xrightarrow{u} f$ and $\forall\, n \in \mathbb{N}\,,\, f_n$ is bounded, then f is bounded.*

Proof. By definiton of uniform limit (cf. 1.1.6) $\exists\, N \in \mathbb{N} : \|f-f_N\|_\infty < 1$. By the triangle inequality: $\|f\|_\infty \leq \|f-f_N\|_\infty + \|f_N\|_\infty < 1 + \|f_N\|_\infty < \infty$ ☐

Theorem 1.1.16 (Uniform Limit)**.** *Every uniformly convergent sequence of continuous, the limit is continuous.*

Proof. Let $f_n \xrightarrow{u} f$. For any $\epsilon > 0$, let $N \in \mathbb{N}$ s.t. $\|f - f_N\|_\infty < \epsilon/3$, that is, $\forall\, x \in \mathcal{I}\,,\ |f(x) - f_N(x)| < \epsilon/3$. Since f_N is continuous, $\forall\, a \in \mathcal{I}\,,\ \exists\, \delta > 0 : \forall\, x \in (a - \delta, a + \delta) \subseteq \mathcal{I}\,,\ |f_N(x) - f_N(a)| < \epsilon/3$. Putting all the terms together, and using triangle inequality:

$$|f(x) - f(a)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| < \epsilon$$

Hence, $\forall\, a \in \mathcal{I}\,,\ f$ is continuous at a. Therefore, f is a continuous on $\mathcal{I}$. $\square$

Remark 1.1.17. *Defining the set of:*

- *Bounded Functions on $\mathcal{I}$, $B(\mathcal{I})$*
- *Continuous Functions on $\mathcal{I}$, $C(\mathcal{I})$*

Then, 1.1.16 and 1.1.15 imply $B(\mathcal{I})$ and $C(\mathcal{I})$ are closed under limits.

Theorem 1.1.18 ($B(\mathcal{I})$ and $C(\mathcal{I})$ are complete)**.** *If $\{f_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{I}$, that is,*

$$\forall\, \epsilon > 0\,,\ \exists\, N \in \mathbb{N} : \forall\, m, n > N\,,\ \|f_m - f_n\|_\infty < \epsilon$$

then, $\exists\, f : \mathcal{I} \to \mathbb{R} : f_n \xrightarrow{u} f$.

Proof. Let $\{f_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence. As in 1.1.10,

$$\forall\, \epsilon > 0\,,\ \exists\, N \in \mathbb{N} : \forall\, m, n > N\,,\ \forall\, x \in \mathcal{I}\,,\ |f_m(x) - f_n(x)| < \epsilon$$

Therefore, for each $x \in \mathcal{I}$, the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$. Hence, since $\mathbb{R}$ is complete, each sequence converges to some $f(x)$. We define the pointwise limit $f(x) := \lim\limits_{n\to\infty} f_n(x)$, so it converges pointwise, which is necessary. Lastly, we need to prove $f_n \xrightarrow{u} f$. By the continuity of absolute value, we have $|f(x) - f_n(x)| = \lim\limits_{m\to\infty} |f_m(x) - f_n(x)|$. Since $\forall\, \epsilon > 0\,,\ \exists\, N \in \mathbb{N} : \forall\, m, n > N\,,\ \forall\, x \in \mathcal{I}\,,\ |f_m(x) - f_n(x)| < \epsilon/2$, we may take $m \to \infty$:

$$\forall\, \epsilon > 0\,,\ \exists\, N \in \mathbb{N} : \forall\, n > N\,,\ \forall\, x \in \mathcal{I}\,,\ |f(x) - f_n(x)| \leq \epsilon/2$$

So that, $\|f - f_n\|_\infty \leq \epsilon/2 < \epsilon$ (cf. 1.1.1, 1.1.10), hence, it converges uniformly. $\square$

Theorem 1.1.19 (Convergence of Integral)**.** *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of continuous functions in $[a,b]$. Suppose $f_n \xrightarrow{u} f$ in $[a,b]$. then*

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\int_a^b f_n(x)\,dx$$

which is defined, cf. 1.1.16.

Proof. Calculating: $\left|\int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx\right| = \left|\int_a^b \Big[f(x)-f_n(x)\Big]\,dx\right| \leq \int_a^b \big|f(x)-f_n(x)\big|\,dx \leq \|f-f_n\|_\infty \cdot (b-a) \to 0$ □

Remark 1.1.20. *Uniform limit in 1.1.19 is necessary. For example, take $f_n : [0,1] \to \mathbb{R}$ where: $f_n(x) = \begin{cases} n & \textit{if } 0 < x < \frac{1}{n} \\ 0 & \textit{otherwise} \end{cases}$. We get: $f_n \to f \equiv 0$, but $\int_0^1 f_n(x)\,dx = 1 \not\to 0$.*

Theorem 1.1.21 (UC of Derivative)**.** *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of differentiable functions in $\mathcal{I}$. Suppose $f_n \to f$ (pointwise) and $f_n' \xrightarrow{u} g$ in $\mathcal{I}$. then f is differentiable and $f' = g$.*

Proof. By FTC II (Newton-Leibnitz), $f_n(x) = f_n(a) + \int_a^x f_n'(t)\,dt$ for some $a \in \mathcal{I}$, taking limit of both sides, for a fixed $x \in \mathcal{I}$, we get:

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} f_n(a) + \lim_{n\to\infty}\int_a^x f_n(t)\,dt = f(a) + \int_a^x g(t)\,dt$$

the last equality by 1.1.19. Hence, by FTC I, $f' = g$. □

1.2 Series of Functions

Definition 1.2.1. *Let $f_n : \mathcal{I} \to \mathbb{R}$, we define:*

1. *$\sum_{n=1}^{\infty} f_n$ converges pointwise iff $\{\sum_{k=1}^{n} f_k\}_{n\in\mathbb{N}}$ converges pointwise, i.e. $\forall\, x_0 \in \mathcal{I}\,,\ \sum_{n=1}^{\infty} f_n(x_0)$ converges (cf. 1.1.3).*

2. *$\sum_{n=1}^{\infty} f_n$ converges uniformly iff $\{\sum_{k=1}^{n} f_k\}_{n\in\mathbb{N}}$ converges uniformly (cf. 1.1.6).*

Lemma 1.2.2. *A series $\sum_{n=1}^{\infty} f_n$ converges uniformly in $\mathcal{I}$ iff it converges pointwise and $\lim\limits_{n\to\infty} \sup\limits_{x\in\mathcal{I}} |\sum_{k=n}^{\infty} f_k(x)| = 0$*

Proof. Let $S_n = \sum_{k=1}^{n} f_k$, the partial sums. By definition (cf. 1.2.1), $\sum_{n=1}^{\infty} f_n$ converges uniformly iff $\{S_n\}_{n\in\mathbb{N}}$ converges uniformly. It converges uniformly to S, if it converges pointwise to S and $\|S - S_n\|_\infty \to 0$ (cf. 1.1.12,1.1.7). Then, $S_n \to S : x \mapsto \sum_{k=1}^{\infty} f_k(x)$. It is N&S $\|S - S_n\|_\infty \to 0$, that is, $\lim\limits_{n\to\infty} \|S - S_n\|_\infty = \lim\limits_{n\to\infty} \sup\limits_{x\in\mathcal{I}} \left|\sum_{k=n+1}^{\infty} f_k(x)\right| = 0$. □

Theorem 1.2.3 (AbsC ⇒ UC of Series)**.** *If $\sum_{n=1}^{\infty} \|f_n\|_\infty$ converges (absolutely), then $\sum_{n=1}^{\infty} f_n$ converges uniformly.*

Proof. Let $S_n = \sum_{k=1}^{n} f_k$. Let $\epsilon > 0$. Since $\sum_{n=1}^{\infty} \|f_n\|_\infty$ converges, $\exists\, N \in \mathbb{N} : \forall\, m > n \geq N\,,\ \sum_{k=n+1}^{m} \|f_k\|_\infty < \epsilon$. Then, we get directly by triangle inequality: $\forall\, m > n \geq N\,,\ \forall\, x \in \mathcal{I}\,,$

$$|S_m(x) - S_n(x)| = \left|\sum_{k=n+1}^{m} f_k(x)\right| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} \|f_k\|_\infty < \epsilon$$

Hence, $\sum_{n=1}^{\infty} f_n$ converges uniformly by Cauchy (cf. 1.1.18). □

Corollary 1.2.4 (Weierstrass M-test)**.** *Let $f_n : \mathcal{I} \to \mathbb{R}$ be a sequence of functions. Suppose there is a (non-negative) sequence $\{M_n\}_{n\in\mathbb{N}}$ such that:*

1. *$\forall\, n \in \mathbb{N}\,,\ \forall\, x \in \mathcal{I}\,,\ |f_n(x)| \leq M_n$, that is, $\forall\, n \in \mathbb{N}\,,\ \|f_n\|_\infty \leq M_n$*

2. *$\sum_{n=1}^{\infty} M_n$ converges (absolutely).*

Then $\sum_{n=1}^{\infty} f_n$ converges uniformly.

Proof. By comparision test of $\{M_n\}_{n\in\mathbb{N}}$ with $\{\|f_n\|_\infty\}_{n\in\mathbb{N}}$, if $\sum_{n=1}^{\infty} M_n$ converges, $\sum_{n=1}^{\infty} \|f_n\|_\infty$ converges. By 1.2.3, $\sum_{n=1}^{\infty} f_n$ converges uniformly. □

Lemma 1.2.5. *Let* $\{f_n\}_{n\in\mathbb{N}}$ *be a sequence of continuous functions in* $[a, b]$*. If* $\sum_{n=1}^{\infty} f_n$ *converges uniformly, then*

$$\int_a^b \sum_{n=1}^{\infty} f_n(x)\,dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,dx$$

Proof. By linearity of the integral, $\int_a^b \sum_{k=1}^{n} f_k(x)\,dx = \sum_{k=1}^{n} \int_a^b f_k(x)\,dx$. Taking the limit of both sides, it follows from 1.1.19 and the definition (cf. 1.2.1). □

Lemma 1.2.6. *Let* $\{f_n\}_{n\in\mathbb{N}}$ *be a sequence of differentiable functions in* $[a, b]$*. If* $\sum_{n=1}^{\infty} f_n$ *converges pointwise and* $\sum_{n=1}^{\infty} f_n'$ *converges uniformly, then*

$$\left(\sum_{n=1}^{\infty} f_n\right)' = \sum_{n=1}^{\infty} f_n'$$

Proof. By linearity of the derivative, $(\sum_{k=1}^{n} f_k)' = \sum_{k=1}^{n} f_k'$. Taking the limit of both sides, it follows from 1.1.21 and the definition (cf. 1.2.1). □

1.3 Power Series

Definition 1.3.1 (Power Series). *Given a sequence of real numbers* $\{a_n\}_{n\in\mathbb{N}}$ *and* $a \in \mathbb{R}$*, its power series is the series of functions* $f_n(x) = a_n(x-a)^n$ *for* $n \in \mathbb{N}_0$*. That is, the power series is:*

$$f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n$$

where the left hand side converges uniformly on some interval $\mathcal{I}$*. Of course, it converges pointwise at* $x = a$*.*

Lemma 1.3.2. *If* $\sum_{n=0}^{\infty} a_n(x-a)^n$ *converges at* $x = x_0$*, then, it converges uniformly in* $(a-r, a+r)$ *for any* $r < |x_0 - a|$*.*

Proof. Let $\mathcal{I} = (a-r, a+r)$. We calculate: $\|f_n\|_\infty = \sup\limits_{x\in\mathcal{I}} |a_n(x-a)^n| = |a_n|r^n$. Since $\sum_{n=0}^{\infty} a_n(x_0-a)^n$ converges, $\{a_n(x_0-a)^n\}_{n\in\mathbb{N}}$ is bounded (by M). Hence, $\|f_n\|_\infty \leq M\left(\frac{r}{|x_0-a|}\right)^n$. Since $\frac{r}{|x_0-a|} < 1$, it follows the series of f_n converges uniformly by Weierstrass M-test (cf. 1.2.4). □

Corollary 1.3.3. *If* $\sum_{n=0}^{\infty} a_n(x-a)^n$ *converges at* x_0*, the pointwise limit (which exists by 1.3.2 and 1.1.7 taking* $|x-a| < r < |x_0-a|$*) is continuous at* $(a - |x_0-a|, |x_0-a|)$*.*

Definition 1.3.4 (Radius of Convergence). *R is a radius of convergence of* $\sum_{n=0}^{\infty} a_n(x-a)^n$ *iff, for any given* $x \in \mathbb{R}$

$\forall\, x \in (a-R, a+R)\,,\ \sum_{n=0}^{\infty} a_n(x-a)^n$ *converges*
$\forall\, x \notin [a-R, a+R]\,,\ \sum_{n=0}^{\infty} a_n(x-a)^n$ *diverges*

Lemma 1.3.5 (Cauchy Hadamard Formula). *Given a sequence of real numbers* $\{a_n\}_{n\in\mathbb{N}}$*, the radius of convergence (cf. 1.3.4) satisfies:*

$$\frac{1}{R} = \limsup \sqrt[n]{|a_n|}$$

if $\limsup \sqrt[n]{|a_n|} = 0$*, then* $R = \infty$
if $\limsup \sqrt[n]{|a_n|} = \infty$*, then* $R = 0$

Proof. It is a direct result of Cauchy's Criteria (Root Test), we get the formula: $|x-a| \cdot \dfrac{1}{R} < 1$. The second proposition is the contrapositive of the divergence criteria. □

Remark 1.3.6. *The radius of convergence only shows pointwise convergence. Moreover, we have to check the endpoints* $x = a \pm R$ *separately.*

Corollary 1.3.7. *By 1.3.2, for any integral* $\mathcal{I} \subsetneqq (a-R, a+R)$*, the power series* $\sum_{n=0}^{\infty} a_n(x-a)^n$ *converges uniformly in* $\mathcal{I}$*.*

Remark 1.3.8. *In general, nothing can be said about uniform convergence on* $(a-R, a+R)$*.*

Lemma 1.3.9. *Differentiation and Integration term-by-term (cf. 1.1.19 and 1.1.21) is valid for power series on the interval of convergence.*

Proof. Since both are local properties, we can take an arbitrarly interval (cf. 1.3.2) to prove differentiability/continuity on every point in $(a-R, a+R)$ and integration on every interval in $(a-R, a+R)$ (cf. 1.3.7). $\square$

Corollary 1.3.10 (Taylor Series)**.** *If* $f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n$ *with a positive radius of convergence, then* f *is infinetly differentiable in* $(a-R, a+R)$ *and* $\forall\, n \in \mathbb{N}_0\,,\ a_n = \dfrac{f^{(n)}(a)}{n!}$

Remark 1.3.11 (Analytic Functions)**.** *Let* T_n *be the* n*-th Taylor Polynomial of* f*. It is not necessarily true that* $T_n \xrightarrow{u} f$*. If it is true, we say* $f \in C^{\omega}$*.*

2 Multivariable Analysis

2.1 Multivariable Geometry

Definition 2.1.1 (Euclidean Space). $\mathbb{R}^n = \{(a_1, \cdots, a_n) \mid a_1, \cdots, a_n \in \mathbb{R}\}$.
We have the following operations:

Addition: $(a_1, \cdots, a_n) + (b_1, \cdots, b_n) = (a_1 + b_1, \cdots, a_n + b_n)$
Scalar multiplication: $\lambda \cdot (a_1, \cdots, a_n) = (\lambda \cdot a_1, \cdots, \lambda \cdot a_n)$
Norm: $\|(a_1, \cdots, a_n)\| = \sqrt{\sum_{i=1}^n a_i^2}$
Scalar product: $(a_1, \cdots, a_n) \bullet (b_1, \cdots, b_n) = \sum_{i=1}^n a_i \cdot b_i$
Basis $e_i = (0, \cdots, 1, \cdots, 0)$ *at the* i*-th place.*

With those operations, $\mathbb{R}^n$ *is an Euclidean Space (cf. Linear Algebra).*

Lemma 2.1.2. *The angle between two vectors is* $\arccos\left(\dfrac{\vec{a} \bullet \vec{b}}{\|\vec{a}\| \cdot \|\vec{b}\|}\right)$.

Corollary 2.1.3 (Perpendicularity). $\vec{a} \perp \vec{b} \Leftrightarrow \vec{a} \bullet \vec{b} = (0, 0, 0)$

Definition 2.1.4 (Vector Product). *In* $\mathbb{R}^3$*, we define the following operation* $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ *as:*

$$(a_1, a_2, a_3) \times (b_1, b_2, b_3) = (a_2 \cdot b_3 - a_3 \cdot b_2, a_3 \cdot b_1 - a_1 \cdot b_3, a_1 \cdot b_2 - a_2 \cdot b_1)$$

further, we can use the short hand using determinants, by formally expanding Laplace's formula (cf. Linear Algebra) on the first row:

$$\vec{a} \times \vec{b} = (a_1, a_2, a_3) \times (b_1, b_2, b_3) = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

where $\vec{e}_1 = (1, 0, 0)$*,* $\vec{e}_2 = (0, 1, 0)$ *and* $\vec{e}_3 = (0, 0, 1)$*, the standard basis.*

Lemma 2.1.5. *The cross product obeys:*

Antisymmetry: $\forall\, \vec{a}, \vec{b} \in \mathbb{R}^3\,,\ \vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$
Linearity: $\forall\, \alpha, \beta \in \mathbb{R}\,,\ \forall\, \vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3\,,$
$(\alpha \cdot \vec{a} + \beta \cdot \vec{b}) \times \vec{c} = \alpha \cdot \vec{a} \times \vec{c} + \beta \cdot \vec{b} \times \vec{c}$
$\vec{c} \times (\alpha \cdot \vec{a} + \beta \cdot \vec{b}) = \alpha \cdot \vec{c} \times \vec{a} + \beta \cdot \vec{c} \times \vec{b}$
Perpendicularity: $\vec{a} \times \vec{b} \perp \vec{a}, \vec{b}$.

Proof. Antisymmetry and linearity follow directly from the definition with determinants. For perpendicularity, we only need to check $\vec{a} \cdot (\vec{a} \times \vec{b}) = 0$ and $\vec{b} \cdot (\vec{a} \times \vec{b}) = 0$, by explicit definition (cf. 2.1.4). □

Corollary 2.1.6. $\vec{a} \times \vec{b} = (0, 0, 0) \Leftrightarrow \vec{a}, \vec{b}$ *are linearly dependent.*

Definition 2.1.7 (Right Handed)**.** *A basis* $(\vec{b}_1, \vec{b}_2, \vec{b}_3)$ *of* $\mathbb{R}^3$ *is right handed iff* $\vec{b}_1 \times \vec{b}_2 = \vec{b}_3$*. The standard basis* $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ *is right handed (direct calculatation with 2.1.4).*

Definition 2.1.8 (Lines and Planes)**.** *We define the following geometrical objects: For* $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$

- *Line:* $\left\{ \vec{a} + t \cdot \vec{ab} \;\middle|\; t \in \mathbb{R} \right\} = \vec{a} + \mathrm{Span}(\vec{u})$
- *Segment:* $[\vec{a}, \vec{b}] = \left\{ \vec{a} + t \cdot \vec{ab} \;\middle|\; t \in [0, 1] \right\}$
- *Hyperplane:* $\left\{ \vec{a} + t \cdot \vec{ab} + s \cdot \vec{ac} \;\middle|\; t, s \in \mathbb{R} \right\} = \vec{a} + \mathrm{Span}(\vec{ab}, \vec{ac})$

where $\vec{ab} = \vec{b} - \vec{a}$ *and* $\vec{ac} = \vec{c} - \vec{a}$

Lemma 2.1.9. *For a plane equation* $ax + by + cz = d$*, we can convert into* $\vec{a} + \mathrm{Span}(\vec{ab}, \vec{ac})$*.*

- *If* $a, b, c \neq 0$*, then* $\vec{a} = (d/a, 0, 0), \vec{b} = (0, d/b, 0), \vec{c} = (0, 0, d/c)$*.*
- *If any of those are zero, change the corresponding vector entry to* 1*.*

To reverse, let $(a, b, c) = \vec{n} = \vec{ab} \times \vec{ac}$*, then* $\vec{n} \cdot (\vec{x} - \vec{a}) = 0$ *is the plane equation.*

Definition 2.1.10 (Affine Map)**.** *Let* $A \in M_{n\times k}(\mathbb{R})$ *(cf. Linear Algebra) and* $\vec{w} \in \mathbb{R}^n$*. Then an affine map is:*

$$\begin{aligned} \Phi : \mathbb{R}^k &\to \mathbb{R}^n \\ \vec{x} &\mapsto A\vec{x} + \vec{w} \end{aligned}$$

That is, an affine map is a linear map composed with a translation. Moreover, S is an affine transformation iff $T(\vec{x}) = \Phi(\vec{x}) - \Phi(\vec{0})$ *is a linear transformation.*

Lemma 2.1.11. *An affine map preserves lines and planes.*

Proof. Let $\Phi(\vec{x}) = T(\vec{x}) + \vec{w}$, T linear. In general,

$$\forall\, S \in (\mathbb{R}^n)^k,\ \Phi(\vec{a} + \mathrm{Span}(S)) = \Phi(\vec{a}) + \mathrm{Span}(T(S))$$

Moreover, $T(\vec{ab}) = \vec{a'b'}$ where $\vec{a'} = S(\vec{a})$ and $\vec{b'} = S(\vec{b})$ □

Definition 2.1.12 (Quadratic Curves)**.** *A quadratic curve/conic section is a set defined by* $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$*. Those are three categories:*

- *Ellipse:* $\dfrac{(x - x_0)^2}{a^2} + \dfrac{(y - y_0)^2}{b^2} = 1$
- *Hiperbola:* $\dfrac{(x - x_0)^2}{a^2} - \dfrac{(y - y_0)^2}{b^2} = \pm 1$
- *Parabola:* $y = a(x - x_0)^2$ *or* $x = a(y - y_0)^2$

Under correct translation and rotation (cf. 2.1.10), every quadratic curve is either one of these three or is degenerate (one or two lines, one point, or $\varnothing$*).*

Remark 2.1.13. *Those curves are given by the intersection of a plane with the double cone* $z^2 = x^2 + y^2$*.*

2.2 Metric Topology

Definition 2.2.1 (Open and Closed Sets)**.** *Let* $B_r(x) = \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$ *and* $K_r(x) = \{y \in \mathbb{R}^n \mid \|x - y\| \leq r\}$ *be, respectively, the open and closed balls in* $\mathbb{R}^n$*. A set* $A \subseteq \mathbb{R}^n$ *is:*

- *Open if* $\forall\, x \in A\,,\, \exists\, \epsilon > 0 : B_\epsilon(x) \subseteq A$
- *Closed if* $\mathbb{R}^n \setminus A$ *is open.*

Lemma 2.2.2. *The open ball is open and the closed ball is closed.*

Proof. We prove each claim separately.

(i) For $x \in B_r(x_0)$, let $\epsilon = r - \|x - x_0\| > 0$ and $\forall\, y \in B_\epsilon(x)\,,\, \|y - x_0\| \leq \|y - x\| + \|x - x_0\| < \epsilon + \|x - x_0\| = r$, by triangle inequality, $\Rightarrow y \in B_r(x_0)$. Therefore, $B_\epsilon(x) \subseteq B_r(x_0)$.

(ii) $\forall\, x \in \mathbb{R}^n \setminus K_r(x_0)\,,$ let $\epsilon = \|x - x_0\| - r > 0$ and $\forall\, y \in B_\epsilon(x)\,,\, \|y - x_0\| \geq \|x - x_0\| - \|y - x\| > \|x - x_0\| - \epsilon = r$, by reverse triangle inequality, $\Rightarrow y \in \mathbb{R}^n \setminus K_r(x_0)$. Therefore, $B_\epsilon(x) \subseteq \mathbb{R}^n \setminus K_r(x_0)$.

$\square$

Definition 2.2.3 (Interior and Boundary)**.** *For a subset* $A \subseteq \mathbb{R}^n$*, we define:*

- *Interior:* $A^\circ = \{x \in A \mid \exists\, \epsilon > 0 : B_\epsilon(x) \subseteq A\}$
- *Closure:* $\overline{A} := \{x \in \mathbb{R}^n \mid \forall\, r > 0\,,\, B_r(x) \cap A \neq \varnothing\}$
- *Boundary:* $\partial A = \{x \in \mathbb{R}^n \mid \forall\, r > 0\,,\, \exists\, y \in A, z \notin A : y, z \in B_r(x)\} = \{x \in \mathbb{R}^n \mid \forall\, r > 0\,,\, B_r(x) \cap A \neq \varnothing$ *and* $B_r(x) \cap (\mathbb{R}^n \setminus A) \neq \varnothing\}$
- *Derived:* $A' = \{x \in \mathbb{R}^n \mid \forall\, \epsilon > 0\,,\, \exists\, y \in A : 0 < \|x - y\| < \epsilon\}$
- *Isolated:* $A^i = \{x \in A \mid \exists\, r > 0 : B_r(x) \cap A = \{x\}\}$

We name $x \in A'$ *a limit point,* $x \in A^i$ *an isolated point,* $x \in \partial A$ *a boundary point and* $x \in A^\circ$ *an interior point.*

Remark 2.2.4. *A is open iff $A^\circ = A$ (by definition).*

Lemma 2.2.5 (Interior)**.** *A° is open.*

Proof. Let $x \in A^\circ$, then $\exists\, r > 0 : B_r(x) \subseteq A$. Let $y \in B_r(x)$, then $\exists\, s > 0 : B_s(y) \subseteq B_r(x) \subseteq A$ (cf. 2.2.2) $\Rightarrow y \in A^\circ$. Hence $\forall\, x \in A^\circ\,,\, B_r(x) \subseteq A^\circ$ for some $r > 0$. So, A° is open. $\square$

Remark 2.2.6. *Notice* $\big[B_r(x) \cap A\big] \setminus \{x\} = \{y \in A \mid 0 < \|x - y\| < r\}$. *Hence,* $A' = \big\{x \in \mathbb{R}^n \bigm| \forall\, r > 0\,,\ \big[B_r(x) \cap A\big] \setminus \{x\} \neq \varnothing\big\}$.

Theorem 2.2.7 (Closure)**.** $\overline{A} = A^i \sqcup A' = A \cup \partial A = A \cup A' = \mathbb{R}^n \setminus (\mathbb{R}^n \setminus A)^\circ$

Proof. Observe that $A' \subseteq \overline{A}$ and $A^i \cap A' = \varnothing$ (cf. 2.2.3). Therefore (cf. 2.2.6): $\overline{A} \setminus A' = \big\{x \in \mathbb{R}^n \bigm| \exists\, r > 0 : B_r(x) \cap A \neq \varnothing \text{ and } \big[B_r(x) \cap A\big] \setminus \{x\} = \varnothing\big\}$. Let $x \in \mathbb{R}^n \setminus A$, if $\varnothing = \big[B_r(x) \cap A\big] \setminus \{x\} \Rightarrow B_r(x) \cap A \subseteq \{x\} \Rightarrow B_r(x) \cap A = \varnothing$. Then, $x \notin \overline{A} \setminus A'$. Hence, $\overline{A} \setminus A' = \{x \in A \mid \exists\, r > 0 : B_r(x) \cap A = \{x\}\} = A^i$. We now prove each term is equal to $\mathbb{R}^n \setminus (\mathbb{R}^n \setminus A)^\circ$.

- ($\overline{A}$) Notice $B_r(x) \subseteq \mathbb{R}^n \setminus A \Leftrightarrow B_r(x) \cap A = \varnothing$. By definition: $\mathbb{R}^n \setminus \overline{A} = \{x \in \mathbb{R}^n \mid \exists\, r > 0 : B_r(x) \cap A = \varnothing\}$ and if $x \in A \Rightarrow B_r(x) \cap A = \{x\}$. Then, $\mathbb{R}^n \setminus \overline{A} = \{x \in \mathbb{R}^n \setminus A \mid \exists\, r > 0 : B_r(x) \subseteq \mathbb{R}^n \setminus A\} = (\mathbb{R}^n \setminus A)^\circ$.
- (∂A) Notice $B_r(x) \subseteq \mathbb{R}^n \setminus A \Leftrightarrow B_r(x) \cap A = \varnothing$. Moreover, by definition: $\mathbb{R}^n \setminus \partial A = \{x \in \mathbb{R}^n \mid \exists\, r > 0 : B_r(x) \cap A = \varnothing \text{ or } B_r(x) \cap (\mathbb{R}^n \setminus A) = \varnothing\}$ and if $x \in \mathbb{R}^n \setminus A \Rightarrow B_r(x) \cap (\mathbb{R}^n \setminus A) = \{x\}$. So, $(\mathbb{R}^n \setminus A) \cap (\mathbb{R}^n \setminus \partial A) = \{x \in \mathbb{R}^n \mid \exists\, r > 0 : B_r(x) \cap A = \varnothing\} = (\mathbb{R}^n \setminus A)^\circ$.
- (A') Observe $B_r(x) \subseteq \mathbb{R}^n \setminus A \Leftrightarrow \forall\, y \in A\,,\ \|x - y\| \geq r$. By definition, $\mathbb{R}^n \setminus A' = \{x \in \mathbb{R}^n \mid \exists\, r > 0 : \forall\, y \in A\,,\ x = y \text{ or } \|x - y\| \geq r\}$. Hence, $(\mathbb{R}^n \setminus A) \cap (\mathbb{R}^n \setminus A') = \{x \in \mathbb{R}^n \setminus A \mid \exists\, r > 0 : B_r(x) \subseteq \mathbb{R}^n \setminus A\} = (\mathbb{R}^n \setminus A)^\circ$.

□

Corollary 2.2.8. $\overline{A}$ *is closed (due to* $\mathbb{R}^n \setminus \overline{A} = (\mathbb{R}^n \setminus A)^\circ$ *open).*

Corollary 2.2.9. $\partial A = \overline{A} \cap \overline{(\mathbb{R}^n \setminus A)}$

Theorem 2.2.10. *The following are equivalent:*
(a) A *is closed*
(b) $\partial A \subseteq A$
(c) $A' \subseteq A$
(d) $\overline{A} = A$

Proof. We prove each one separately.

($a \Leftrightarrow b$) $\mathbb{R}^n \setminus A$ is open $\Leftrightarrow$ $\mathbb{R}^n \setminus A = (\mathbb{R}^n \setminus A)^\circ = (\mathbb{R}^n \setminus A) \cap (\mathbb{R}^n \setminus \partial A) \Leftrightarrow \mathbb{R}^n \setminus A \subseteq \mathbb{R}^n \setminus \partial A$ (cf. 2.2.7)

($a \Leftrightarrow c$) $\mathbb{R}^n \setminus A$ is open $\Leftrightarrow$ $\mathbb{R}^n \setminus A = (\mathbb{R}^n \setminus A)^\circ = (\mathbb{R}^n \setminus A) \cap (\mathbb{R}^n \setminus A') \Leftrightarrow \mathbb{R}^n \setminus A \subseteq \mathbb{R}^n \setminus A'$ (cf. 2.2.7)

($b \Rightarrow d$) $\partial A \subseteq A \Rightarrow A \subseteq A \cup \partial A = \overline{A} \subseteq A \Rightarrow A = \overline{A}$

($d \Rightarrow b$) $A = \overline{A} = A \cup \partial A \Rightarrow \partial A \subseteq A$

□

Lemma 2.2.11. *Maximality of the interior and minimality of the closure:*
(a) U *open and* $A^\circ \subseteq U \subseteq A \Rightarrow U = A^\circ$
(b) V *closed and* $A \subseteq V \subseteq \overline{A} \Rightarrow V = \overline{A}$

Proof. We prove each one separately.
(a) U is open, then let $x \in U$, so $\exists\, r > 0 : B_r(x) \subseteq U \subseteq A \Rightarrow x \in A^\circ$. Hence, $U \subseteq A^\circ$ and $A^\circ \subseteq U$ (given), therefore, $A^\circ = U$.
(b) $(\mathbb{R}^n \setminus A)^\circ = \mathbb{R}^n \setminus \overline{A} \subseteq \mathbb{R}^n \setminus V \subseteq \mathbb{R}^n \setminus A$ (cf. 2.2.7), since $\mathbb{R}^n \setminus V$ is open, by (a), $\Rightarrow (\mathbb{R}^n \setminus A)^\circ = \mathbb{R}^n \setminus V \Rightarrow V = \overline{A}$.

□

Definition 2.2.12 (Topology)**.** *An (open) topology* $\mathcal{T} \subseteq \mathcal{P}(\mathbb{R}^n)$ *is the set of all open sets. It obeys the following requirements:*
(T1) $\varnothing, \mathbb{R}^n \in \mathcal{T}$
(T2) Finite Intersection: $U, V \in \mathcal{T} \Rightarrow U \cap V \in \mathcal{T}$
(T3) Arbitrary Union: $\mathcal{C} \subseteq \mathcal{T} \Rightarrow \displaystyle\bigcup_{U \in \mathcal{C}} U \in \mathcal{T}$
A closed topology $\mathcal{K} \subseteq \mathcal{P}(\mathbb{R}^n)$ *is the set of all closed sets. It obeys the following requirements:*
(K1) $\varnothing, \mathbb{R}^n \in \mathcal{K}$
(K2) Finite Union: $U, V \in \mathcal{K} \Rightarrow U \cup V \in \mathcal{K}$
(K3) Arbitrary Intersection: $\mathcal{C} \subseteq \mathcal{K} \Rightarrow \displaystyle\bigcap_{U \in \mathcal{C}} U \in \mathcal{K}$

Theorem 2.2.13. *The definition of open (cf. 2.2.1) obeys the open topology (cf. 2.2.12), denoted* $\mathcal{T}^n$.

Proof. We trivially obey (T1). For (T2), we need to find $r > 0$ such that $B_r(x) \subseteq U \cap V$. We pick $r = min\{\alpha, \beta\}$ where $B_\alpha(x) \subseteq U$ and $B_\beta(x) \subseteq V$, with the observation that $B_\alpha(x) \cap B_\beta(x) = B_{\min\{\alpha,\beta\}}(x)$. For (T3), we note that any open set U can be written as: $U = \bigcup_{x \in U} B_{r(x)}(x)$. Since the open balls are open (cf. 2.2.2), every union of open sets can be rewritten as another union of open ball neighbourhoods. □

Corollary 2.2.14. *The closed topology follows by taking the complement of the previous relations and using deMorgan's Law.*

Lemma 2.2.15. *Let* $A \in \mathbb{R}^n$*, define* $\mathcal{T}_A = \{U \cap A \mid U \in \mathcal{T}^n\}$*. Then,* $\mathcal{T}_A$ *is a topology on* A.

Proof. For all statements (T1-3), intersection is distributive. □

2.3 Limits

Definition 2.3.1 (Limit of Sequences)**.** *A sequence* $\{a_n \in \mathbb{R}^k\}_{n\in\mathbb{N}}$ *converges to* $L \in \mathbb{R}^k$ *iff:*

$$\forall\, \epsilon > 0\,,\, \exists\, N \in \mathbb{N} : \forall\, n \geq N\,,\, \|a_n - L\| < \epsilon$$

equivalently, for $a_n = (a_{1,n}, a_{2,n}, \cdots, a_{k,n})$ *and* $L = (L_1, L_2, \cdots, L_k)$ *then* $a_n \to L$ *iff:*

$$\forall\, m \in \{1, 2 \cdots, k\}\,,\, a_{m,n} \to L_m$$

Remark 2.3.2. *By using convergence on* $\mathbb{R}$*, it is immediate to see* $\mathbb{R}^n$ *is complete. That is, every Cauchy sequence is convergent. Further, those lemmas are immediately valid:*
(a) If a sequence $\{a_n \in \mathbb{R}^k\}_{n\in\mathbb{N}}$ *has a limit* L*, then it is unique.*
(b) $\lim\limits_{n\to\infty} a_n = L \Leftrightarrow \lim\limits_{n\to\infty} \|a_n - L\| = 0$
(c) Every convergent sequence is bounded.
(d) $\lim\limits_{n\to\infty} \lambda = \lambda$
(e) $\lim\limits_{n\to\infty} (\lambda \cdot a_n) = \lambda \cdot (\lim\limits_{n\to\infty} a_n)$
(f) $\lim\limits_{n\to\infty} (a_n \pm b_n) = (\lim\limits_{n\to\infty} a_n) \pm (\lim\limits_{n\to\infty} b_n)$
(g) $\{b_n \in \mathbb{R}^k\}_{n\in\mathbb{N}}$ *be bounded and* $\lim\limits_{n\to\infty} a_n = 0$*, then:* $\lim\limits_{n\to\infty} (a_n \bullet b_n) = 0$
(h) $\lim\limits_{n\to\infty} (a_n \bullet b_n) = (\lim\limits_{n\to\infty} a_n) \bullet (\lim\limits_{n\to\infty} b_n)$

Lemma 2.3.3. $x \in A'$ *iff* $\exists\, \{x_n \in A \setminus \{x\}\}_{n\in\mathbb{N}} : x_n \to x$.

Proof. We prove each direction.

($\Rightarrow$) Then, $\forall\, n \in \mathbb{N}\,,\, \exists\, y \in A : 0 < \|x - y\| < 2^{-n}$. By Axiom of Countable Choice, we may choose $\{x_n\}_{n\in\mathbb{N}}$ such that $\|x - x_n\| < 2^{-n} \to 0$. Therefore, $x_n \to x$.

($\Leftarrow$) By contrapositive, $x \notin A' :\Leftrightarrow \exists\, \epsilon > 0 : \forall\, y \in A\,,\, x = y$ or $\|x - y\| \geq \epsilon$. For any sequence $\{x_n \in A \setminus \{x\}\}_{n\in\mathbb{N}}$, $\exists\, \epsilon > 0 : \forall\, n \in \mathbb{N}\,,\, \|x - x_n\| \geq \epsilon$. Hence $x_n \not\to x$.

□

Corollary 2.3.4. $x \in \overline{A}$ *iff* $\exists\, \{x_n \in A\}_{n\in\mathbb{N}} : x_n \to x$.

Proof. This follows from 2.3.3 and 2.2.7 ($\overline{A} = A' \sqcup A^i$) by taking the constant sequence $x_n = x$ for $x \in A^i$. □

Definition 2.3.5. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ for $a \in \overline{A}$, then, $L \in \mathbb{R}^n$ is called the limit of f at a if:*

(Heine) $\forall \{x_n \in A\}_{n\in\mathbb{N}}\,,\ x_n \to a \Rightarrow f(x_n) \to L$

(Cauchy) $\forall \epsilon > 0\,,\ \exists \delta > 0 : \forall x \in A\,,\ \|x - a\| < \delta \Rightarrow \|f(x) - L\| < \epsilon$

Lemma 2.3.6 (H&C)**.** *The Heine definition and the Cauchy definition of the limit are equivalent.*

Proof. We prove each direction:

($\Rightarrow$) By contrapositive, suppose $\exists \epsilon > 0 : \forall \delta > 0\,,\ \exists x \in A : 0 < \|x - a\| < \delta \Rightarrow \|f(x) - L\| \geq \epsilon$. By Axiom of Countable Choice, define a sequence $\{x_n \in A\}_{n\in\mathbb{N}}$ such that $\|x_n - a\| < 2^{-n}$ and $\|f(x_n) - L\| \geq \epsilon$. Hence $x_n \to a$. Therefore $f(x_n) \not\to L$, by definition, so Heine does not hold.

($\Leftarrow$) By contrapositive, suppose $\exists \{x_n \in A\}_{n\in\mathbb{N}} : x_n \to a$ (cf. 2.3.4), but $f(x_n) \not\to L$. By definition of the limits:

$$\forall \delta > 0\,,\ \exists N \in \mathbb{N} : \forall n \geq N\,,\ \|x_n - a\| < \delta$$
$$\exists \epsilon > 0 : \forall N \in \mathbb{N}\,,\ \exists n \geq N : \|f(x_n) - L\| \geq \epsilon$$

Hence $\exists \epsilon > 0 : \forall \delta > 0\,,\ \|x_n - a\| < \delta \not\Rightarrow \|f(x_n) - L\| < \epsilon$, so Cauchy does not hold.

Hence, both definitons can be used interchangebly. ☐

Lemma 2.3.7 (Calculating Limits)**.** *For some $\epsilon, \delta > 0$, let $g : (-\epsilon, \epsilon) \to \mathbb{R}$ be a function such that $\forall x \in B_\delta(a) \cap A\,,\ \|f(x) - L\| \leq g(\|x - a\|)$. Then $f(x) \to L$ as $x \to a$ iff $\lim\limits_{r\to 0} g(r) = 0$.*

Proof. Follows directly from Heine (cf. 2.3.5). ☐

2.4 Continuity

Definition 2.4.1 (Continuity)**.** *Let f be defined on $A \ni a$. We say that f is continuous at point a if: $\lim_{x \to a} f(x) = f(a)$. If $\forall\, a \in A\,,\ f$ is continuous at a, then f is continuous in A.*

Example 2.4.2. *The norm is continuous.*

Definition 2.4.3 (Image and Preimage)**.** *For $f : A \to B$, we write:*

- *For $S \subseteq A$, $f(S) = \{f(a) \in B \mid a \in S\}$*
- *For $R \subseteq B$, $f^{-1}(R) = \{a \in A \mid f(a) \in R\}$*

Remark 2.4.4. *Cauchy can be rephrased as follows: Given $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$, then $f(x) \to L$ as $x \to a$ iff:*

$$\forall\, \epsilon > 0\,,\ \exists\, \delta > 0 : \forall\, x \in A\,,\ x \in B_\delta(a) \Rightarrow f(x) \in B_\epsilon(L)$$

$$\forall\, \epsilon > 0\,,\ \exists\, \delta > 0 : f(B_\delta(a) \cap A) \subseteq B_\epsilon(L)$$

Theorem 2.4.5. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is continuous iff*

$$\forall\, U \in \mathcal{T}^k\,,\ f^{-1}(U) \in \mathcal{T}_A$$

cf. 2.2.13, 2.2.15.

Proof. We prove both directions:

- ($\Rightarrow$) Let $U \in \mathcal{T}^k$ and $f(a) \in U$. By definition, $\exists\, \epsilon > 0 : B_\epsilon(f(a)) \subseteq U$. Then (cf. 2.4.4) $\exists\, \delta > 0 : f(B_\delta(a) \cap A) \subseteq B_\epsilon(f(a)) \subseteq U$. Hence, $\forall\, a \in f^{-1}(U)\,,\ \exists\, \epsilon > 0 : \exists\, \delta > 0 : B_\delta(a) \cap A \subseteq f^{-1}(U)$. Therefore, $f^{-1}(U) \in \mathcal{T}_A$.
- ($\Leftarrow$) Take $B_\epsilon(f(a)) \in \mathcal{T}^k$, then $f^{-1}(B_\epsilon(f(a))) = A \cap U$ for $U \in \mathcal{T}^n$. By definition, $\exists\, \delta > 0 : B_\delta(a) \cap A \subseteq f^{-1}(B_\epsilon(f(a)))$. Hence (cf. 2.4.4), $\forall\, a \in A\,,\ \forall\, \epsilon > 0\,,\ \exists\, \delta > 0 : f(B_\delta(a) \cap A) \subseteq B_\epsilon(f(a))$.

$\square$

Corollary 2.4.6. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is continuous iff*

$$\forall\, V \in \mathcal{K}^k\,,\ f^{-1}(V) \in \mathcal{K}_A$$

where $\mathcal{K}_A = \{V \cap A \mid V \in \mathcal{K}^n\}$.

Definition 2.4.7 (Level Set)**.** *For $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$. Define the level set*

$$N(\alpha) = f^{-1}(\{\alpha\}) = \{a \in A \mid f(a) = \alpha\}$$

Moreover, if f is continuous, $N(\alpha)$ is closed.

Lemma 2.4.8. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is continuous on A^i.*

Proof. For $a \in A^i$, let $\delta > 0$ such that $B_\delta(a) \cap A = \{a\}$ (cf. 2.2.3). Then, $\forall \epsilon > 0\,,\ f(B_\delta(a) \cap A) = \{f(a)\} \subseteq B_\epsilon(f(a))$. Hence, by 2.4.4, f is continuous at a. □

Definition 2.4.9 (Lipschitz)**.** *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is Lipschitz continuous if:*

$$\exists\, K > 0 : \forall\, x, y \in A\,,\ \|f(x) - f(y)\| \leq K \cdot \|x - y\|$$

Remark 2.4.10. *A Lipschitz continuous function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is continuous on A since $\|f(x) - f(x_n)\| \leq K \cdot \|x - x_n\| \to 0$, for $x_n \to x$.*

Example 2.4.11. *An affine map (cf. 2.1.10) is Lipschitz continuous. Let $\Phi(x) = A\,x + w$ and the rows of A denoted r_i. $\|\Phi(x) - \Phi(y)\| = \|A\,(x - y)\| = \sqrt{\sum_{i=1}^{k} \|r_i \cdot (x - y)\|^2} \leq \sqrt{\sum_{i=1}^{k} \sum_{j=1}^{n} |a_{i,j}|^2} \cdot \|x - y\| = \sqrt{\operatorname{tr}(A\,A^t)} \cdot \|x - y\|$. We write $\|A\| = \sqrt{\operatorname{tr}(A\,A^t)}$.*

Theorem 2.4.12. *If $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ and $g : B \subseteq \mathbb{R}^k \to \mathbb{R}^m$, such that $f(A) \subseteq B$, then $g \circ f : A \to \mathbb{R}^m$ is continuous.*

Proof. By Heine (cf. 2.3.5), let $x_n \to x$ for $\{x_n \in A\}_{n \in \mathbb{N}}$. Since f is continuous, $f(x_n) \to f(x)$. Observe $\{f(x_n) \in B\}_{n \in \mathbb{N}}$. Since g is continuous, $g(f(x_n)) \to g(f(x))$. Therefore, $g \circ f$ is continuous. □

2.5 Compactness

Definition 2.5.1 (Boundedness)**.** *A set $A \subseteq \mathbb{R}^n$ is called a bounded set iff $\exists\, x \in \mathbb{R}^n, r > 0 : A \subseteq B_r(x)$. Equivalently if $A \subseteq K_r(x)$.*

Definition 2.5.2. *A sequence $\{x_n \in \mathbb{R}^k\}_{n\in\mathbb{N}}$ is bounded if $\{x_n \mid n \in \mathbb{N}\}$ is bounded (cf. 2.5.1). Equivalently, if $\exists\, M > 0 : \forall\, n \in \mathbb{N}\,,\ \|x_n\| \leq M$.*

Theorem 2.5.3 (Bolzano-Weierstrass)**.** *Any bounded sequence $\{a_n \in \mathbb{R}^m\}_{n\in\mathbb{N}}$ (cf. 2.5.2), there exists a convergent subsequence. That is, exist a sequence $\{b_k \in \mathbb{R}^m\}_{k\in\mathbb{N}}$ is such that: $b_k = a_{n_k}$ where $\{n_k\}_{k\in\mathbb{N}}$ is a stricly increasing sequence of natural numbers.*

Proof. We'll work with induction on m. Base case is Bolzano-Weierstrass on $\mathbb{R}$ (cf. Calculus I). Let $a_n = (a_{1,n}, \cdots, a_{m,n})$ and $b_n = (a_{1,n}, \cdots, a_{m-1,n})$. By induction hypothesis, there is a convergent subsequence $\{b_{n_k}\}_{k\in\mathbb{N}}$. Further, $\{a_{m,n_k}\}_{k\in\mathbb{N}}$ has a converging subsequence $\{a_{m,n_{k_j}}\}_{j\in\mathbb{N}}$. Then $\{a_{n_{k_j}}\}_{j\in\mathbb{N}}$ is a converging subsequence of a_n. $\square$

Definition 2.5.4 (Compactness in $\mathbb{R}^n$)**.** *A set $A \subseteq \mathbb{R}^n$ is compact iff it closed and bounded (cf. 2.5.1).*

Remark 2.5.5. *There is a broader definition of (topological) compactness, but the Heine-Borel theorem guarantees 2.5.4 is necessary and sufficient in $\mathbb{R}^n$. Also, due to 2.5.3, there is a definition of sequentially compact, that is, for any bounded sequence $\{a_n \in K\}_{n\in\mathbb{N}}$, there exists a convergent subsequence in K.*

Remark 2.5.6. *$K_r(x)$ is compact (cf. 2.2.2).*

Theorem 2.5.7 (Weierstrass Theorem)**.** *Let $f : K \to \mathbb{R}^k$ be a continuous function, where K is compact (cf. 2.5.4), then $f(K)$ is compact (cf. 2.4.3).*

Proof. We prove $f(K)$ is bounded and closed (cf. 2.5.4).

- By contrary, suppose $f(K)$ is not bounded. Then, $\forall\, n \in \mathbb{N}\,,\ \exists\, x_n \in K : \|f(x_n)\| > n$. By Axiom of Countable Choice, this defines a sequence $\{x_n\}_{n\in\mathbb{N}}$. Because K is bounded, by 2.5.3, there is a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ converging to $x_0 \in K$ (cf. 2.3.4). Since f is continuous, $f(x_{n_k}) \to f(x_0)$. But $\forall\, k \in \mathbb{N}\,,\ \|f(x_{n_k})\| > n_k \geq k$ which implies that $\|f(x_{n_k})\| \to \infty$, contradiction since $\|f(x_{n_k})\| \to \|f(x_0)\|$ (cf. 2.4.2). Therefore, $f(K)$ is bounded.

- Take a sequence $\{f(x_n) \in f(K)\}_{n\in\mathbb{N}}$ converging to y_0. Then $\{x_n \in K\}_{n\in\mathbb{N}}$ is bounded. Hence, by 2.5.3, there is a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ converging to $x_0 \in K$. Since f is continuous, $f(x_{n_k}) \to f(x_0) = y_0$ (cf. 2.3.5 Heine). Hence, $f(K)$ is closed (cf. 2.3.4, 2.2.10).

□

Corollary 2.5.8. *Let $f : K \subseteq \mathbb{R}^n \to \mathbb{R}$ be a continuous function, where K is compact (cf. 2.5.4), then f attains it's maximum and minimum.*

Definition 2.5.9. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is uniformly continuous if:*

$$\forall \epsilon > 0 \,, \exists \delta > 0 : \forall x, y \in A \,, \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon$$

Theorem 2.5.10 (Heine-Cantor)**.** *Let $f : K \to \mathbb{R}^k$ be a continuous function, where K is compact (cf. 2.5.4), then f is uniformly continuous.*

Proof. The proof is simply to the real line (cf. Calculus I). By contrary, suppose

$$\exists \epsilon > 0 : \forall \delta > 0 \,, \exists x, y \in K : \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| \geq \epsilon$$

Define (by Axiom of Countable Choice) $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ such that $\forall n \in \mathbb{N} \,, \|x_n - y_n\| < 2^{-n}$ and $\|f(x_n) - f(y_n)\| \geq \epsilon$. Since x_n is bounded, by 2.5.3, there is a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ converging to $x_0 \in K$. By Triangle Inequality, $\|y_{n_k} - x_0\| \leq \|x_{n_k} - x_0\| + \|x_{n_k} - y_{n_k}\| \leq \|x_{n_k} - x_0\| + 2^{-n_k}$ hence $y_{n_k} \to x_0$. Since f is continuous, $\lim\limits_{k\to\infty} f(x_{n_k}) = f(x_0) = \lim\limits_{k\to\infty} f(y_{n_k})$. But $\forall k \in \mathbb{N} \,, \|f(x_{n_k}) - f(y_{n_k})\| \geq \epsilon > 0$. By taking $k \to \infty$, there is a contradiction. □

2.6 Connected Sets

Definition 2.6.1 (Path/Curve)**.** *A path/curve in $A \subseteq \mathbb{R}^n$ is a continuous function (cf. 2.4.1) $\gamma : [a,b] \to A$. Further, $\gamma(a)$ and $\gamma(b)$ are called the endpoints. Let $\gamma(t) = (\gamma_1(t), \gamma_2(t), \cdots, \gamma_n(t))$. Then, it is much simpler to check continuity, γ is continuous iff $\forall\, i \in \{1, 2, \cdots, n\}$, $\gamma_i : [a,b] \to \mathbb{R}$ is continuous.*

Remark 2.6.2. *If we have $\delta : [a,b] \to A$, we may write $\gamma(x) = \delta\left(\frac{x-a}{b-a}\right)$ so that $\gamma : [0,1] \to A$ and $\gamma([0,1]) = \delta([a,b])$, also matching the endpoints.*

Remark 2.6.3. *By 2.5.7, $\Gamma = \gamma([a,b])$ is compact. Further, we may refer to Γ as the curve and γ as the parametrization.*

Lemma 2.6.4 (Path in Limits)**.** *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ for $a \in \overline{A}$, then, $L \in \mathbb{R}^n$ is called the limit of f at a iff:*

$$\forall\, path\ \gamma : [0,1] \to A \ with\ \gamma(0) = a\,,\ \lim_{t\to 0} f(\gamma(t)) = L$$

Proof. ($\Rightarrow$) By definition,

$$\begin{aligned}
&\forall\, \epsilon > 0\,,\ \exists\, \delta > 0 : \forall\, x \in A\,,\ \|x - a\| < \delta \Rightarrow \|f(x) - L\| < \epsilon \\
&\forall\, \delta > 0\,,\ \exists\, \theta > 0 : \forall\, t \in [0,1]\,,\ 0 < t < \theta \Rightarrow \|\gamma(t) - a\| < \delta \\
\Rightarrow\ :\ &\forall\, \epsilon > 0\,,\ \exists\, \theta > 0 : \forall\, t \in [0,1]\,,\ 0 < t < \theta \Rightarrow \|f(\gamma(t)) - L\| < \epsilon
\end{aligned}$$

($\Leftarrow$) By contrary, suppose

$$\begin{aligned}
&\exists\, \epsilon > 0 : \forall\, \delta > 0\,,\ \exists\, x \in A : \|x - a\| < \delta \Rightarrow \|f(x) - L\| \geq \epsilon \\
&\exists\, \epsilon > 0 : \forall\, n \in \mathbb{N}\,,\ \exists\, x_n \in A : \|x_n - a\| < 2^{-n} \Rightarrow \|f(x) - L\| \geq \epsilon
\end{aligned}$$

For each $n \in \mathbb{N}$, by Axiom of Countable Choice, take a curve γ connecting x_n linearly (clearly continuous).

$$\begin{aligned}
&\forall\, \delta > 0\,,\ \exists\, \theta > 0 : \forall\, t \in [0,1]\,,\ 0 < t < \theta \Rightarrow \|\gamma(t) - a\| < \delta \\
&\forall\, n \in \mathbb{N}\,,\ \exists\, \theta > 0 : \forall\, t \in [0,1]\,,\ 0 < t < \theta \Rightarrow \|\gamma(t) - a\| < 2^{-n} \\
\Rightarrow\ :\ &\exists\, \epsilon > 0 : \forall\, \theta > 0\,,\ \exists\, t \in [0,1] : 0 < t < \theta \Rightarrow \|f(\gamma(t)) - L\| \geq \epsilon
\end{aligned}$$

□

Lemma 2.6.5 (Joining Paths)**.** *Let $\gamma : [a,b] \to A$ and $\delta : [b,c] \to A$ and let $\gamma(b) = \delta(b)$. Then $\beta : [a,c] \to A$ where*

$$\beta(t) = (\gamma\#\delta)(t) = \begin{cases} \gamma(t) & if\ t \in [a,b] \\ \delta(t) & if\ t \in (b,c] \end{cases}$$

is path from $\gamma(a)$ to $\delta(c)$. Further, observe: $\beta([a,c]) = \gamma([a,b]) \cup \delta([c,d])$

Proof. The only we need to prove is that β is continuous, which follows directly from $\gamma(b) = \beta(b) = \delta(b) = \lim_{t \to b^+} \delta(t)$. $\square$

Definition 2.6.6 (Path-Connected Sets)**.** *A subset $A \subset \mathbb{R}^n$ is path connected iff $\forall\, x, y \in A$, $\exists$ path from x to y (cf. 2.6.1).*

Definition 2.6.7 (Connected Set)**.** *A subset $A \subseteq \mathbb{R}^n$ is (topologically) connected iff $\nexists\, R, S \in \mathcal{T}_A \setminus \{\varnothing\} : R \sqcup S = A$, that is, $R \cup S = A$ and $R \cap S = \varnothing$.*

Lemma 2.6.8 (Clopen)**.** *A subset $A \subseteq \mathbb{R}^n$ is a connected subset (cf. 2.6.7) iff $\mathcal{T}_A \cap \mathcal{K}_A = \{A, \varnothing\}$.*

Proof. $\mathcal{T}_A \cap \mathcal{K}_A \supseteq \{A, \varnothing\}$ is trivially given. We prove both directions:

($\Rightarrow$) By contrary, suppose $\mathcal{T}_A \cap \mathcal{K}_A \supsetneq \{A, \varnothing\}$ and let $R = A \cap U = A \cap V$ be in the difference with $U \in \mathcal{T}^n$ and $V \in \mathcal{K}^n$. Let $S = A \cap (\mathbb{R}^n \setminus V) \in \mathcal{T}_A$. Observe $R \cup S = A \cap (V \cup (\mathbb{R}^n \setminus V)) = A$ and $R \cap S = A \cap (V \cap (\mathbb{R}^n \setminus V)) = \varnothing$. Hence, A is not connected.

($\Leftarrow$) By contrary, suppose $\exists\, R, S \in \mathcal{T}_A \setminus \{\varnothing\} : R \cup S = A$ and $R \cap S = \varnothing$. Observe $R, S \neq A$. Since $R = A \setminus S = A \cap (\mathbb{R}^n \setminus V)$, where $S = A \cap U$ and $U \in \mathcal{T}^n$ (cf. 2.2.15), then R is closed in A. Hence, $R \in \mathcal{T}_A \cap \mathcal{K}_A \setminus \{A, \varnothing\}$.

$\square$

Definition 2.6.9 (Interval)**.** *An interval is a subset $\mathcal{I} \subseteq \mathbb{R}$ on the real line, iff $\forall\, a < b \in \mathcal{I}$, $a < c < b \Rightarrow c \in \mathcal{I}$.*

Lemma 2.6.10. *A subset $A \subseteq \mathbb{R}$ is connected iff it is either a singleton $\{x\}$ or an interval (cf. 2.6.9).*

Proof. We prove that every connected subset must be either a singleton or an interval. Then, we prove those are connected.

($\Rightarrow$) By contrary, if A is not an interval and has more than one point, by definition (cf. 2.6.9), $\exists\, a < b \in A : \exists\, c \notin A : a < c < b$. Then, let $R = A \cap (-\infty, c)$ and $S = A \cap (c, \infty)$, which are in $\mathcal{T}_A \setminus \{\varnothing\}$ (cf. 2.2.2,2.2.15), would satisfy 2.6.7.

($\Leftarrow$) Cleary $\{x\}$ is connected since if both R and S are non-empty, $R \sqcup S$ shall have at least two elements. Now, by contrary, suppose an interval $\mathcal{I}$ is not connected and $\mathcal{I} = R \sqcup S$. Then, $\exists\, a \in R, b \in S : a < b$ or $b < a$, wlog, $a < b$. Let $c = \sup\{x \in \mathbb{R} \mid [a, x) \subseteq R\}$. Hence, $c \leq b \Rightarrow c \in \mathcal{I}$ (cf. 2.6.9). Since R is closed in $\mathcal{I}$ (cf. 2.6.8), $c \in R$. Further, $R = \mathcal{I} \cap U$

for $U \in \mathcal{T}^n$, then $\exists\, \delta > 0 : (c - \delta, c + \delta) \cap \mathcal{I} \subseteq R$ which contradicts the maximality of c. Contradiction.

$\square$

Theorem 2.6.11. *A subset $A \subseteq \mathbb{R}^n$ is path-connected then it is connected. Conversly, if A is connected and open, then it is path-connected.*

Proof. We prove both directions:

($\Rightarrow$) Suppose $\exists\, R, S \in \mathcal{T}_A : R \sqcup S = A$, take $r \in R$ and $s \in S$. Since A is path-connected, there is a path γ from r to s. Since γ is continuous, $\gamma^{-1}(R)$ and $\gamma^{-1}(S)$ are open in $\mathcal{T}_{[0,1]}$, and also non-empty ($r \in \gamma^{-1}(R)$ and $s \in \gamma^{-1}(S)$). Therefore, $[0, 1] = \gamma^{-1}(A) = \gamma^{-1}(R \sqcup S) = \gamma^{-1}(R) \sqcup \gamma^{-1}(S)$. However, by 2.6.10, $[0, 1]$ is connected. Contradiction.

($\Leftarrow$) Given $a \in A$, let $P \subseteq A$ be the subset of points in A which can be joined to a by a path in A. For $x \in P$, $\exists\, \epsilon > 0 : B_\epsilon(x) \subseteq A$, since A is open. For any $y \in B_\epsilon(x)$, there us a path from x to y by a straightline. Hence, by 2.6.5, $y \in P$. Therefore, $B_\epsilon(x) \subseteq P$, so P is open. For $x \in Q = A \setminus P$, $\exists\, \epsilon > 0 : B_\epsilon(x) \subseteq A$, since A is open. If $B_\epsilon(x) \cap P \neq \varnothing$, then, by 2.6.5, $x \in P$. Hence, $B_\epsilon(x) \subseteq Q$, so Q is open. Moreover, $P \cap Q = \varnothing$ and $P \cup Q = A$. Since A is connected, and P is not empty ($a \in P$), then $P = A$ and $Q = \varnothing$. This is valid for any $a \in A$.

$\square$

Remark 2.6.12. *The converse of 2.6.11 is only valid on $\mathbb{R}^n$.*

Theorem 2.6.13 (Intermediate Value Theorem)**.** *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ be a continuous function.*

1. *If A is path-connected (cf. 2.6.6), then so is $f(A)$*
2. *If A is connected (cf. 2.6.7), then so is $f(A)$*

Proof. We prove each one:

1. Let $x, y \in f(A)$. Then $\exists\, a, b \in A : f(a) = x$ and $f(b) = y$. Since A is path-connected, there is a path $\gamma : [0, 1] \to A$ from a to b. Further, $f \circ \gamma : [0, 1] \to f(A)$ is continuous (cf. 2.4.12), it is a path from x to y.
2. By contrary, suppose $f(A)$ is disconnected. Hence, $f(A) = R \sqcup S$ with $R, S \in \mathcal{T}_A$. Therefore, $A = f^{-1}(R) \sqcup f^{-1}(S)$. Then, A is disconnected.

$\square$

Lemma 2.6.14. *If $\mathfrak{D}$ is an open connected domain, and $x, y \in \mathfrak{D}$, then there is a differentiable curve connectiong x and y.*

3 Differentiation

3.1 Differentiability

Definition 3.1.1 (Derivative). *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is differentiable at the point $a \in A^\circ$ if $\exists\, Df(a) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ (cf. Linear Algebra), called the derivative at a, such that:*

$$\lim_{x \to a} \frac{\|f(x) - f(a) - Df(a)[x-a]\|}{\|x-a\|} = \lim_{h \to \vec{0}} \frac{\|f(a+h) - f(a) - Df(a)[h]\|}{\|h\|} = 0$$

Equivalently, $\exists\, \epsilon_a : A \to \mathbb{R}^k : f(a+h) = f(a) + Df(a)[h] + \epsilon_a(h) \cdot \|h\|$ and $\lim\limits_{h \to \vec{0}} \|\epsilon_a(h)\| = 0$. Moreover, the derivative can be expressed by its (standard) matrix representative $[Df(a)]$ and we define the derivative funcion $Df : A \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ at each differentiable point.

Lemma 3.1.2. *If f is differentiable at $a \in A^\circ$, then $Df(a)$ is uniquely determined by:*

$$\forall\, v \in \mathbb{R}^n\,,\ Df(a)[v] = \lim_{t \to 0} \frac{f(a + t\,v) - f(a)}{t}$$

*Further, $D_v f(a) := Df(a)[v]$ is the **directional derivative** of f in the v direction.*

Proof. We rewrite $h = t\,v$ and $f(a+t\,v) = f(a) + Df(a)[t\,v] + \epsilon_a(t\,v)\,|t|\,\|v\|$. Since $Df(a)$ is linear, we get: $\forall\, t \in \mathbb{R}\,,$

$$Df(a)[v] = \frac{f(a+t\,v) - f(a)}{t} - \|v\|\,\frac{|t|}{t} \cdot \epsilon_a(tv)$$

Since $\|v\|\,\dfrac{|t|}{t}$ is bounded and $\lim\limits_{h \to \vec{0}} \|\epsilon_a(h)\| = \lim\limits_{t \to 0} \|\epsilon_a(t\,v)\| = 0$, by sandwich (cf 2.3.2), the result follows taking $t \to 0$. $\square$

Example 3.1.3. *An affine map $\Phi(x) = A\,x + w$ (cf. 2.1.10) is differentiable in $\mathbb{R}^n$ and $\forall\, a \in \mathbb{R}^n\,,\ [Df(a)] = A$. We get: $\Phi(a+h) - \Phi(a) = A\,h$ so* $\lim\limits_{h \to 0} \dfrac{\|(A - Df(a))[h]\|}{\|h\|} = 0$

Remark 3.1.4. *This shows that the derivative will calculate the best linear approximation of f near a.*

Theorem 3.1.5. *If f is differentiable at a, then f is continuous at a.*

Proof. By Heine 2.3.5, let $a_n \to a$ and set $h = a_n - a$. Then, $f(a_n) = f(a+h) = f(a) + Df(a)[h] + \epsilon_a(h) \cdot \|h\|$. Taking $n \to \infty$, we get $h \to 0$. Since every linear function is continuous 2.4.11 and $\lim_{h \to \vec{0}} \|\epsilon_a(h)\| = 0$, we get $f(a_n) \to f(a)$. $\square$

Definition 3.1.6 (Partial Derivative)**.** *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$, given $f(x) = (f_1(x), f_2(t), \cdots, f_k(x))$, for each f_i and $a \in A^\circ$, define:*

$$\frac{\partial f_i}{\partial x_j}(a) = \lim_{t \to 0} \frac{f_i(a + t\, e_j) - f_i(a)}{t}$$

where e_i is a basis vector (cf. 2.1.1). If all the partial derivatives exist, then f is partially differentiable (at a). Also write

$$\frac{\partial f_i}{\partial u}(a) = \lim_{t \to 0} \frac{f_i(a + t\, u) - f_i(a)}{t}$$

for the directional derivative.

Lemma 3.1.7. *If f is differentiable at $a \in A^\circ$, then f is partially differentiable at a and*

$$[Df(a)] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(a) & \frac{\partial f_k}{\partial x_2}(a) & \cdots & \frac{\partial f_k}{\partial x_n}(a) \end{bmatrix}$$

That is, $[Df(a)]_{i,j} = \dfrac{\partial f_i}{\partial x_j}(a)$.

Proof. First, by definition, $f(x) = \sum_{i=1}^{n} f_i(x)\, e_i$, and, by 3.1.2, $Df(a)[e_j] = \lim_{t \to 0} \frac{f(a + t\, e_j) - f(a)}{t} = \sum_{i=1}^{n} \left(\lim_{t \to 0} \frac{f_i(a + t\, e_j) - f_i(a)}{t} \right) e_i = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(a)\, e_i$. Hence, by definition of $[Df(a)]_{i,j}$, it is as given on the standard basis. $\square$

Lemma 3.1.8 (Hadamard)**.** *$f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is differentiable at $a \in A^\circ$ iff there exists a map $\phi_a : A \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ continuous at a such that:*

$$f(x) = f(a) + \phi_a(x)[x - a]$$

Proof. We prove both directions.

($\Rightarrow$) Define:

$$\phi_a(x) = \begin{cases} Df(a) & \text{if } x = a \\ Df(a) + \dfrac{1}{\|x-a\|} \cdot \epsilon_a(x-a) \cdot (x-a)^t & \text{otherwise} \end{cases}$$

and $\|\phi_a(x) - \phi_a(a)\| = \dfrac{1}{\|x-a\|} \cdot \|\epsilon_a(x-a)\| \cdot \|x-a\| = \|\epsilon_a(x-a)\| \to 0$, where the norm taken was in 2.4.11, so it is continuous at a.

($\Leftarrow$) Calculating:

$$\begin{aligned} \frac{\|f(x) - f(a) - \phi_a(a)[x-a]\|}{\|x-a\|} &= \frac{\|[\phi_a(x) - \phi_a(a)]\,[x-a]\|}{\|x-a\|} \\ &\leq \frac{\|\phi_a(x) - \phi_a(a)\| \cdot \|x-a\|}{\|x-a\|} = \|\phi_a(x) - \phi_a(a)\| \to 0 \end{aligned}$$

(cf. 2.4.11). Hence, $\phi_a(a) = Df(a)$.

□

Theorem 3.1.9 (Continuous Partials)**.** *If $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is partially differentiable on a neighbourhood of $a \in A^\circ$ and each partial derivative is continuous at a, then f is differentiable at a.*

Proof. Define $h^{(0)} = 0$ and $h^{(j)} = h^{(j-1)} + h_j\, e_j$. For $1 \leq i \leq k$:

$$f_i(a+h) - f_i(a) = \sum_{j=1}^{n} \left[f_i(a + h^{(j)}) - f_i(a + h^{(j-1)}) \right] = \sum_{j=1}^{n} \left[g_{i,j}(h_j) - g_{i,j}(0) \right]$$

where $g_{i,j}(t) = f_i(a + h^{(j-1)} + t\, e_j)$, then $g'_{i,j}(t) = \dfrac{\partial f_i}{\partial x_j}\,(a + h^{(j-1)} + t\, e_j)$. By Mean Value Theorem (cf. Calculus I)

$$\exists\, \xi_j \text{ between } 0 \text{ and } h_j : g_{i,j}(h_j) - g_{i,j}(0) = \frac{\partial f_i}{\partial x_j}\,(a + h^{(j-1)} + \xi_j\, e_j) \cdot h_j$$

Then, let $[\phi_a(a+h)]_{i,j} = \dfrac{\partial f_i}{\partial x_j}\,(a + h^{(j-1)} + \xi_j\, e_j)$, which is continuous on a, taking $h \to 0$. The rest follows from 3.1.8. □

Definition 3.1.10. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is called C^1 if it is partially differentiable on A and each partial derivative is continuous in A. Further, we write $f \in C^1(A, \mathbb{R}^k)$*

Corollary 3.1.11. *$f \in C^1(A, \mathbb{R}^k) \Rightarrow f$ is differentiable on A.*

Corollary 3.1.12. *$f \in C^1(A, \mathbb{R}^k) \Rightarrow Df : A \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^k)$ is continuous on A.*

Theorem 3.1.13 (Chain Rule)**.** *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ and $g : B \subseteq \mathbb{R}^k \to \mathbb{R}^m$ such that $f(a) \in B$. If f is differentiable at a and g is differentiable at $f(a)$, then $g \circ f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a and:*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$$

Proof. By 3.1.8, let:

$$\begin{gathered} f(x) - f(a) = \phi_a(x)\,[x - a] \\ g(y) - g(f(a)) = \psi_{f(a)}(y)\,[y - f(a)] \\ \hline \end{gathered}$$

$$\begin{aligned} g(f(x)) - g(f(a)) &= \psi_{f(a)}(f(x))\,\Big[\phi_a(x)\,[x - a]\Big] \\ &= \varphi_a(x)\,[x - a] \end{aligned}$$

Hence, $\varphi_a(x) = \psi_{f(a)}(f(x)) \circ \phi_a(x)$. Taking $x \to a$ and using f, ϕ, ψ are continuous: $D(g\circ f)(a) = \varphi_a(a) = \psi_{f(a)}(f(a))\circ\phi_a(a) = Dg(f(a))\circ Df(a)$ ☐

Definition 3.1.14 (Gradient)**.** *For $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ differentiable define the gradient as: $\nabla f : A^\circ \subseteq \mathbb{R}^n \to \mathbb{R}^n$ so that $\forall\, x \in \mathbb{R}^n$, $Df(a)[x] = \big(\nabla f(a)\big) \cdot x$*

Remark 3.1.15 (Gradient Boost)**.** *The gradient of f at a point is in the direction of greatest increast for f at that point.*

Corollary 3.1.16. *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ differentiable and $\gamma : [0, 1] \to A$ a differentiable curve. Then: $\dfrac{d}{dt}\big(f(\gamma(t))\big) = \nabla f(\gamma(t)) \cdot \gamma'(t)$*

Lemma 3.1.17. *If $\nabla f \equiv 0$ on an open connected domain, then $f \equiv$ const.*

Proof. By 2.6.14, for $x, y \in \mathfrak{D}$, there is a differentiable path $\gamma : [a, b] \to \mathfrak{D}$ connecting the two. Then: $\frac{d}{dt}\big(f(\gamma(t))\big) = \nabla f(\gamma(t)) \cdot \gamma'(t) = 0 \Rightarrow f(\gamma(t)) \equiv$ const. $\Rightarrow f(x) = f(y)$. ☐

Lemma 3.1.18. *Let* $a \in N(\alpha)$ *(cf. 2.4.7) then,* $\nabla f(a)$ *is perpendicular to* $N(\alpha)$*. That is, it is perpendicular to the tangent line/plane/hyperplane.*

Proof. Let $\gamma : [0,1] \to N(\alpha)$, hence, $\forall\, t \in [0,1]\,,\ f(\gamma(t)) = \alpha \Rightarrow$ (by 3.1.16) $\nabla f(\gamma(t)) \cdot \gamma'(t) = \dfrac{d}{dt}\big(f(\gamma(t))\big) = 0$, hence $\nabla f(\gamma(t)) \perp \gamma'(t)$. $\square$

3.2 Higher Order Derivatives and Taylor

Definition 3.2.1 (Higher Partial Derivatives)**.** *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, for $a \in A^\circ$, define:*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right), \text{ in general, } \frac{\partial^n f}{\partial x_{i_1} \cdots \partial x_{i_n}} = \frac{\partial}{\partial x_{i_1}}\left(\frac{\partial^{n-1} f}{\partial x_{i_2} \cdots \partial x_{i_n}}\right)$$

when the previous partial derivative exists in a neighbourhood of a.

Example 3.2.2. *$f(x,y) = x^2 y$, then $f_x(x,y) = 2xy$ and $f_y(x,y) = x^2$, and so: $f_{xx}(x,y) = 2y$, $f_{yx}(x,y) = 2x$, $f_{xy}(x,y) = 2x$, $f_{yy}(x,y) = 0$.*

Theorem 3.2.3 (Schwarz-Clairut)**.** *If both $\partial_{x_i}\partial_{x_j} f$ and $\partial_{x_j}\partial_{x_i} f$ are continuous at $a \in A^\circ$:*

$$\partial_{x_i}\partial_{x_j} f\,(a) = \partial_{x_j}\partial_{x_i} f\,(a)$$

Proof. By definition (cf. 3.1.6):

$$\begin{aligned}
\partial_{x_i}\partial_{x_j} f\,(a) &= \lim_{s\to 0} \frac{\partial_{x_j} f\,(a + s\,e_i) - \partial_{x_j} f\,(a)}{s} = \lim_{s\to 0}\lim_{t\to 0} g(s,t) \\
\partial_{x_j}\partial_{x_i} f\,(a) &= \lim_{t\to 0} \frac{\partial_{x_i} f\,(a + t\,e_j) - \partial_{x_i} f\,(a)}{t} = \lim_{t\to 0}\lim_{s\to 0} g(s,t)
\end{aligned}$$

where $g(s,t) = \dfrac{f(a + s\,e_i + t\,e_j) - f(a + s\,e_i) - f(a + t\,e_j) + f(a)}{st}$. By Mean Value Theorem (cf. Calculus I):

$$\begin{aligned}
g(s,t) &= \frac{\partial_{x_j} f\,(a + s\,e_i + \tau_{1,t}) - \partial_{x_j} f\,(a + \tau_{1,t})}{s} = \partial_{x_i}\partial_{x_j} f\,(a + \tau_{1,s} + \tau_{1,t}) \\
&= \frac{\partial_{x_i} f\,(a + t\,e_j + \tau_{2,s}) - \partial_{x_i} f\,(a + \tau_{2,s})}{t} = \partial_{x_j}\partial_{x_i} f\,(a + \tau_{2,s} + \tau_{2,t})
\end{aligned}$$

where $\tau_{k,s}$ between $0, s$ and $\tau_{k,t}$ between $0, t$. Since $\partial_{x_i}\partial_{x_j} f$ and $\partial_{x_j}\partial_{x_i} f$ are continuous at a, the result follows taking $t \to 0, s \to 0$ and $s \to 0, t \to 0$. □

Definition 3.2.4. *A function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ is called C^r if it is partially differentiable on A r times and each partial derivative is continuous in A (including at the r-th derivative). Further, we write $f \in C^r(A, \mathbb{R}^k)$*

Corollary 3.2.5. *If $f \in C^2(A, \mathbb{R}^k)$ then,*

$$\forall\, a \in A\,,\ \forall\, i,j \in \{1, \cdots, n\}\,,\ \partial_{x_i}\partial_{x_j} f\,(a) = \partial_{x_j}\partial_{x_i} f\,(a)$$

Corollary 3.2.6. *If $f \in C^r(A, \mathbb{R}^k)$ then all mixed partial derivatives are equal at every point in A.*

Definition 3.2.7 (Multi-index Notation)**.** *Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k) \in \mathbb{N}_0^k$, we define: For $h \in \mathbb{R}^k$ and $f \in C^r(A, \mathbb{R})$*

(i) $|\alpha| = \alpha_1 + \alpha \cdots + \alpha_k$

(ii) $h^\alpha = h_1^{\alpha_1} \cdot h_2^{\alpha_2} \cdots h_k^{\alpha_k}$

(iii) $\alpha! = \alpha_1! \cdot \alpha_2! \cdots \alpha_k!$

(iv) $\partial^\alpha f = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_k}^{\alpha_k} f = \dfrac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$ *where* $|\alpha| \leq r$

Theorem 3.2.8 (Taylor)**.** *Let $f \in C^{r+1}(A, \mathbb{R})$, then $\exists\, R_r(a) : \mathbb{R}^n \to \mathbb{R}$: $\forall\, h \in \mathbb{R}^n$, for $c \in (0, 1)$:*

$$f(a+h) = \sum_{|\alpha| \leq r} \frac{h^\alpha}{\alpha!} \partial^\alpha f(a) + R_r(a,h) \quad \textit{and } R_r(a,h) = \sum_{|\alpha|=r+1} \frac{h^\alpha}{\alpha!} \partial^\alpha f(a + c\,h)$$

Proof. Let $g(t) = f(a + t\,h)$, by the multinomial theorem:

$$g^{(j)}(t) = \big(h_1\, \partial_{x_1} + \cdots + h_n\, \partial_{x_n}\big)^j f(a + t\,h) = \sum_{|\alpha|=j} \frac{j!}{\alpha!} h^\alpha\, \partial^\alpha f(a+th)$$

By Taylor's Theorem (cf. Calculus I):

$$f(a+h) = g(1) = \sum_{j=0}^{r} \frac{g^{(j)}(0)}{j!} + \frac{g^{(r+1)}(c)}{(r+1)!} = \sum_{|\alpha| \leq r} \frac{h^\alpha}{\alpha!} \partial^\alpha f(a) + R_r(a,h)$$

as given in the formula, $R_r(a,h) = \dfrac{g^{(r+1)}(c)}{(r+1)!} = \sum_{|\alpha|=r+1} \dfrac{h^\alpha}{\alpha!} \partial^\alpha f(a + c\,h)$. $\square$

3.3 Optimization and Critical Points

Definition 3.3.1 (Extremum Point). *We say that $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ has a local minimum (eq. maximum) at a if exists an open ball $B_\epsilon(a) \subseteq A$ about a such that:*

$$\forall\, x \in B_\epsilon(a)\,,\ f(x) \geq f(a)\ \Big(\textit{eq. } f(x) \leq f(a)\Big)$$

The point a which is either a local minimum or local maximum point is called a local extremum point.

Definition 3.3.2 (Critical Point). *We say that a is a critical point of f if either $\nabla f(a) = 0$ or $\nexists\, \nabla f(a)$. Further, a critical point that is not an extremum is a saddle point.*

Theorem 3.3.3 (Fermat). *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ be partially differentiable at $a \in A^\circ$. If f gets its maximal (or minimal) value at a, then: $\nabla f(a) = \vec{0}$*

Proof. Let $g_i(t) = f(a + t\, e_i) \Rightarrow g'(t) = \partial_{x_i} f(a + t\, e_i) \Rightarrow g'(0) = \partial_{x_i} f(a)$. By Fermat (cf. Calculus I), $t = 0$ is an extremum iff $\partial_{x_i} f(a) = 0$. □

Lemma 3.3.4 (NC for Local Extremum). *a is a local extremum $\Rightarrow$ a is a critical point.*

Proof. If f is differentiable at a, we use 3.3.3. Otherwise, $\nexists\, \nabla f(a)$. Either way, a is a critical point. □

Definition 3.3.5 (Hessian). *For $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ twice differentiable, let:*

$$Hf(a) = [D^2 f(a)] = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(a) \end{bmatrix} = [\partial_{x_i} \partial_{x_j} f(a)]_{i,j}$$

Further, if $f \in C^2$ at a, then $Hf(a)^t = Hf(a)$.

Lemma 3.3.6. *For $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, C^2 at a, $\forall\, h \in \mathbb{R}^n$,*

$$\sum_{|\alpha|=2} \frac{h^\alpha}{\alpha!}\, \partial^\alpha f(a) = \frac{1}{2}\, h \cdot \Big(Hf(a)\, h\Big)$$

Proof. Follows from the definition of the Hessian and from 3.2.3. $\square$

Corollary 3.3.7. *By 3.2.8,* $\forall\, h \in \mathbb{R}^n\,,\ \exists\, c \in (0,1)$:

$$f(a+h) = f(a) + \big(\nabla f(a)\big) \cdot h + \frac{1}{2}\, h \cdot \Big(Hf(a + c\,h)\, h\Big)$$

Definition 3.3.8 (Definiteness)**.** *A matrix* $A \in \mathrm{Sym}_n(\mathbb{R})$ *(symmetric* $n \times n$ *matrix) is positive definite iff:* $\forall\, x \in \mathbb{R}^n \setminus \{0\}\,,\ x \cdot (A\,x) > 0$*. Moreover, it is negative definite iff* $-A$ *is positive definite. Otherwise, it is indefinite.*

Theorem 3.3.9 (Second Derivative Test)**.** *For* $f \in C^2(A, \mathbb{R})$*. Then,* a *is a critical point and* $Hf(a)$ *is positive definite* $\Rightarrow$ a *is a local minimum of* f*.*

Proof. Since a is a critical point of f, then $\nabla f(a) = \vec{0}$, we get, for $h \in B_\epsilon(\vec{0})\ \exists\, c \in (0,1)\ :\ f(a+h) = f(a) + \frac{1}{2}\, h \cdot \Big(Hf(a + c\,h)\, h\Big)$. Since f is C^2, $Hf(a + c\,h)$ is still positive definite for some $\|h\| = \epsilon > 0$. Hence, $\forall\, h \in B_\epsilon(\vec{0}) \setminus \{\vec{0}\}\,,\ f(a+h) - f(a) > 0$. $\square$

Corollary 3.3.10. *For* $f \in C^2(A, \mathbb{R})$*. Then,* a *is a critical point and* $Hf(a)$ *is negative definite* $\Rightarrow$ a *is a local maximum of* f*.*

Lemma 3.3.11 (Saddle Point)**.** *For* $f \in C^2(A, \mathbb{R})$*, if* a *is a critical point and* $Hf(a)$ *is indefinite* $\Rightarrow$ a *is a saddle point of* f*.*

Proof. Then, $\exists\, u, v \in B_\epsilon(\vec{0}) : u \cdot (Hf(a)\, u) > 0$ and $v \cdot (Hf(a)\, v) < 0$. Since f is C^2, $Hf(a + c\,h)$ is still positive or negative for some $\|h\| = \epsilon > 0$. So, $f(a+u) - f(a) = \frac{1}{2}\, u \cdot \Big(Hf(a + c\,u)\, u\Big) > 0$ and $f(a+v) - f(a) = \frac{1}{2}\, v \cdot \Big(Hf(a + c\,v)\, v\Big) < 0$. $\square$

Lemma 3.3.12. *For* $A \in \mathrm{Sym}_n(\mathbb{R})$*,* A *is positive definite iff all eigenvalues of* A *are positive.*

Proof. We prove each one:

($\Rightarrow$) Let $x \neq 0$ be an eigenvector with eigenvalue λ, then A is positive definite $\Rightarrow x \cdot (A\,x) = \lambda \cdot \|x\|^2 > 0 \Rightarrow \lambda > 0$.

($\Leftarrow$) By Spectral Theorem (cf. Linear Algebra) we get: $A = Q^t\, \Lambda\, Q$, for Q orthogonal and $\Lambda = \mathrm{diag}\{\lambda_i\}$, the diagonal of eigenvalues. Since all eigenvalues are positive, $\Lambda = D^2$, where $D = \mathrm{diag}\{\sqrt{\lambda_i}\}$. Then, $x \cdot (A\,x) = x^t\, Q^t\, D^2\, Q\, x = \|D\, Q\, x\|^2 > 0$

$\square$

Theorem 3.3.13 (Sylvester's Criterion)**.** *For $A \in \mathrm{Sym}_n(\mathbb{R})$, A is positive definite iff every leading principal minors (the determinant obtained by removing the last $n-k$ rows and last $n-k$ columns, for $k \in \{1, \cdots, n\}$, which we denote Δ_k) are positive.*

Proof. We prove each one:

($\Rightarrow$) Since all eigenvalues are positive (cf. 3.3.12), Δ_n, which is the product of all eigenvalues (cf. Linear Algebra), is positive. Let $A^{(k)}$ be the leading principal matrix. Take x with last $n-k$ values zero and y the vector of first k entries of x, then $x \cdot (A\,x) = y \cdot (A^{(k)}\,y) > 0$. Then, $\forall\, k \in \{1, \cdots, n\}$, $A^{(k)}$ is positive definite. Hence, all Δ_k are positive.

($\Leftarrow$) By induction on n:

- Base Case: $A \in \mathrm{Sym}_1(\mathbb{R}) = \mathbb{R}$, then $\Delta_1 = A > 0$.
- Inductive Step: By induction, $A^{(n-1)}$ is positive definite. First, we prove A only has one negative eigenvalue. By contradiction, say there are two negative eigenvalues, hence two independent eigenvectors u, v such that $u \cdot (A\,u) < 0$ and $v \cdot (A\,v) < 0$. Define $w = v_n \cdot u - u_n \cdot v$, so that $w_n = 0$. Then, $w \cdot (A^{(n-1)}\,w) = w \cdot (A\,w) = v_n^2 \cdot u \cdot (A\,u) + u_n^2 \cdot v \cdot (A\,v) < 0$, which is a contradiction since $A^{(n-1)}$ is positive definite. But $\Delta_n > 0$, so there are no negative eigenvalues. Also, no zero eigenvalues.

$\square$

Definition 3.3.14 (Global Extremum)**.** *We say that $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ has a global minimum (eq. maximum) at a iff:*

$$\forall\, x \in A\,,\ f(x) \geq f(a)\ \Big(eq.\ f(x) \leq f(a)\Big)$$

The point a which is either a global minimum or global maximum point is called a global extremum point.

Lemma 3.3.15. *A global extremum of $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ is either a local extremum or a boundary point of A.*

Proof. By 2.2.7, $a \in A \subseteq \overline{A}$ either $a \in A^\circ$ or $a \in \partial A$. If $a \in A^\circ$, then $\exists\, \epsilon > 0 : B_\epsilon(a) \subseteq A^\circ$, hence, if a is a global extremum and interior, it is a local extremum. $\square$

3.4 Lagrange, Implicit and Inverse Theorems

Theorem 3.4.1 (Inverse Function)**.** *Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^k$ be C^1 and $Df(a) \in \operatorname{Aut}(\mathbb{R}^n)$, i.e. is invertible, (cf. Linear Algebra). Then,*

$$\exists\, U \ni a \text{ open } \subseteq A : U, f(U) \text{ open and } f : U \to f(U) \text{ is a bijection}$$

Proof. Outside the scope. □

Remark 3.4.2. *For a funcion $F : E \subseteq \mathbb{R}^{n+m} \to \mathbb{R}^k$, define $F^* : \pi_x(E) \to \mathbb{R}^k$ (π_x is the projection of first n components) such that: $F^*(x) = F(x, y_0)$, we write $D_x F(x_0, y_0) = DF^*(x_0)$. Moreover,*

$$[D_x F(x_0, y_0)] = \left[\frac{\partial F_i}{\partial x_j}(x_0, y_0)\right]_{1 \leq i \leq k\,;\, 1 \leq j \leq n}$$

Theorem 3.4.3 (Implicit Function)**.** *Let $F : E \subseteq \mathbb{R}^{n+m} \to \mathbb{R}^k$ be C^1 and $F(x_0, y_0) = 0$ and $D_x F(x_0, y_0) \in \operatorname{Aut}(\mathbb{R}^n)$ for $(x_0, y_0) \in E \subseteq \mathbb{R}^n \times \mathbb{R}^m$. Then,*

$$\exists\, U \ni a \text{ open } \subseteq \pi_x(E) \text{ and } f \in C^1(U, \mathbb{R}^m) : F(x, y) = 0 \Leftrightarrow y = f(x)$$

Proof. Outside the scope. □

Theorem 3.4.4 (Lagrangian Multiplier)**.** *Let a be a global extremum of $f \in C^1(A, \mathbb{R})$ over the constraint $g(x) = 0$, where $g \in C^1(B, \mathbb{R})$. That is, it is a global extremum over the set $A \cap g^{-1}(0)$. Then,*

$$\exists\, \lambda \in \mathbb{R} : \nabla f(a) = \lambda \cdot \nabla g(a)$$

we call λ the Lagragian multiplier.

Proof. Let γ be a differentiable curve in $A \cap g^{-1}(0)$ starting at a, then, by chain rule, since a is a global extremum: $\nabla f(a) \bullet \gamma'(0) = 0$, hence $\nabla f(a)$ is perpendicular to $A \cap g^{-1}(0)$. Therefore $\exists\, \lambda \in \mathbb{R} : \nabla f(a) = \lambda \cdot \nabla g(a)$, due to 3.1.18. Important to remark if $A, B \subseteq \mathbb{R}^n$ a g is not the constant zero (on some ball), then $g^{-1}(0)$ is at most $n-1$ dimensional. □

Corollary 3.4.5. *Let the Lagragian be: $\mathcal{L} : \left[A \cap g^{-1}(0)\right] \times \mathbb{R} \to \mathbb{R}$ where $\mathcal{L}(x, \lambda) = f(x) - \lambda \cdot g(x)$, then the Lagragian condition is: $\nabla_{x,\lambda} \mathcal{L}(a) = 0$.*

Theorem 3.4.6 (Generalized Lagrangian Multiplier)**.** *Let a be a global extremum of $f \in C^1(A, \mathbb{R})$ over the constraint $g(x) = \vec{0}$, where $g \in C^1(B, \mathbb{R}^m)$ such that $\nabla g_i(a)$ are linearly independent ($[Dg(a)]$ has full rank). Then,*

$$\exists\, \vec{\lambda} \in \mathbb{R}^m : \nabla f(a) = \sum_{i=1}^{m} \lambda_i \cdot \nabla g_i(a)$$

Proof. Let γ be a differentiable curve in $A \cap g_i^{-1}(0)$ starting at a, then, by chain rule, since a is a global extremum: $\nabla f(a) \bullet \gamma'(0) = 0$, hence $\nabla f(a)$ is perpendicular to $A \cap g_i^{-1}(0)$, so $\nabla f(a)$ is perpendicular to $A \cap g^{-1}(\vec{0})$. Therefore $\exists\, \vec{\lambda} \in \mathbb{R}^m : \nabla f(a) = \sum_{i=1}^{m} \lambda_i \cdot \nabla g_i(a)$, due to 3.1.18. $\square$

Remark 3.4.7. *If we parametrize ∂A piecewise by a curve g, we still need to analyze the discontinuous points.*

4 Integration

4.1 Riemann and Darboux

Definition 4.1.1 (Hyperrectangle Partition). *For $H = \prod_{k=1}^{n}[a_k, b_k] \subseteq \mathbb{R}^n$ (a hyperrectangle) and $N \in \mathbb{N}$, we define a partition $T = \{H_i\}_{i=1}^{N^n}$ on H:*

$$\begin{aligned} T: \quad & a_1 = x_{1,0} < x_{1,1} < \cdots < x_{1,i-1} < x_{1,i} < \cdots < x_{1,N} = b_1 \\ & a_2 = x_{2,0} < x_{2,1} < \cdots < x_{2,i-1} < x_{2,i} < \cdots < x_{2,N} = b_2 \\ & \vdots \\ & a_k = x_{k,0} < x_{k,1} < \cdots < x_{k,i-1} < x_{k,i} < \cdots < x_{k,N} = b_k \\ & \vdots \\ & a_n = x_{n,0} < x_{n,1} < \cdots < x_{n,i-1} < x_{n,i} < \cdots < x_{n,N} = b_n \end{aligned}$$

so that $H_i = [x_{1,i-1}, x_{1,i}] \times \cdots \times [x_{n,i-1}, x_{n,i}]$. Denote the displacement $\Delta x_{k,i} = x_{k,i} - x_{k,i-1}$, hence $\mathrm{vol}(H_i) = \prod_{k=1}^{n} \Delta x_{k,i}$ and the norm of the partition:

$$\|T\| = \max\{|\Delta x_{k,i}| \mid i = 1, 2, \cdots, N^n \;\; \textit{and} \;\; k = 1, 2, \cdots, n\}$$

Definition 4.1.2 (Rectangle Integration). *Let $f : H \subseteq \mathbb{R}^n \to \mathbb{R}$ where H is a hyperrectangle. For $N \in \mathbb{N}$, defined a partition $T = \{H_i\}_{i=1}^{N^n}$. Choose a point $x_i^* \in H_i$. Then, the sum $R(T, \{x_i^*\}_{i=1}^{N^n}) = \displaystyle\sum_{i=1}^{N^n} f(x_i^*)\, \mathrm{vol}(H_i)$ is called the Riemann sum of T. A function f is called Riemann integrable over H if there exists the limit denoted $I = \displaystyle\int_H f(x) d^n x = \lim_{\|T\| \to 0} R(T, \{x_i^*\}_{i=1}^{N^n})$ and the limit is independent on the choice of T and $\{x_i^*\}_{i=1}^{N^n}$, which is called the definite integral (or the Riemann integral).*

Definition 4.1.3 (Darboux). *Let $f : H \subseteq \mathbb{R}^n \to \mathbb{R}$ where H is a hyperrectangle. For $N \in \mathbb{N}$, defined a partition $T = \{H_i\}_{i=1}^{N^n}$, define $m_i = \inf\limits_{x \in H_i} f(x)$ and $M_i = \sup\limits_{x \in H_i} f(x)$ Then, the sums $L(T) = \sum_{i=1}^{N^n} m_i\, \mathrm{vol}(H_i)$ and $U(T) = \sum_{i=1}^{N^n} M_i\, \mathrm{vol}(H_i)$ are called the lower and upper Darboux sum of T, respectively. We define: $\sup\limits_T L(T) = L$ and $\inf\limits_T U(T) = U$ the lower and upper Darboux integral, denoted $\underline{\displaystyle\int_H} f(x)\, d^n x = L$ and $U = \overline{\displaystyle\int_H} f(x)\, d^n x$. If $U = L$, the function is said to be Darboux integrable over H.*

Lemma 4.1.4 (N&SC for Darboux Integrability)**.** *f is Darboux integrable over H iff $\forall \epsilon > 0\,, \exists T : U(T) - L(T) < \epsilon$*

Proof. That condition is equivalent to $\lim\limits_{\|T\|\to 0} \big(U(T) - L(T)\big) = 0$ $\square$

Lemma 4.1.5 (DI $\Leftrightarrow$ RI)**.** *f is Darboux integrable over H iff it is Riemann integrable over H.*

Proof. Any Riemann sum $R(T, \{x_i^*\}_{i=1}^{N^n})$, is between the Darboux sums: $U(T) \geq R(T, \{x_i^*\}_{i=1}^{N^n}) \geq L(T)$. If $\lim\limits_{\|T\|\to 0} U(T) = \lim\limits_{\|T\|\to 0} L(T) = I$, by Sandwich theorem, $\lim\limits_{\|T\|\to 0} R(T) = I \Rightarrow f$ is Riemann integrable over H. That is, $\overline{\int_H} f(x)\,d^n x = \underline{\int_H} f(x)\,d^n x = \int_H f(x)\,d^n x.$ $\square$

Lemma 4.1.6 (Fubini 2-dim)**.** *$f : [a,b] \times [c,d] = H \to \mathbb{R}$ bounded and integrable on H, we define $\varphi(x) = \int_c^d f(x,y)\,dy$ and $\psi(y) = \int_a^b f(x,y)\,dx$, then the double integral can be calculated as:*

$$\iint_H f\,dA = \int_a^b \varphi(x)\,dx = \int_c^d \psi(y)\,dy$$

Proof. Observe: $\varphi(x) = \sum_{j=1}^{N} \int_{y_{j-1}}^{y_j} f(x,y)\,dy$. Define: $\phi_i = \inf\limits_{x\in[x_{i-1},x_i]} \varphi(x)$ and $\Phi_i = \sup\limits_{x\in[x_{i-1},x_i]} \varphi(x)$. We get: $\sum_{j=1}^{N} m_{i,j}\,\Delta y_j \leq \phi_i \leq \Phi_i \leq \sum_{j=1}^{N} M_{i,j}\,\Delta y_j$. Therefore $L_f(T) \leq L_\varphi(T_x) \leq U_\varphi(T_x) \leq U_f(T)$, where T_x is the partition induced on x. Taking $\|T\| \to 0$ on both sides, we get the result, since f is integrable on H. $\square$

Definition 4.1.7. *For $f : \mathfrak{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ where $\mathfrak{D}$ is bounded, let $H \supseteq \mathfrak{D}$ be a hyperrectangle, define $\tilde{f} : H \to \mathbb{R}$ s.t.: $\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in \mathfrak{D} \\ 0 & \text{otherwise} \end{cases}$ and define $\int_{\mathfrak{D}} f(x)\,d^n x = \int_H \tilde{f}(x)\,d^n x$*

Definition 4.1.8. *A set $\mathfrak{D} \subseteq \mathbb{R}^n$ is measurable if we can assign a volume* $\mathrm{vol}(\mathfrak{D}) \in [0,\infty]$. *That is, iff the constant 1 is integrable over $\mathfrak{D}$.*

Lemma 4.1.9. *For* $\mathfrak{D} \subseteq \mathbb{R}^n$ *such that* $\mathrm{vol}(\mathfrak{D}) = 0$ *and* $f : \mathfrak{D} \to \mathbb{R}$ *is bounded, then* $\displaystyle\int_{\mathfrak{D}} f(x)\, d^n x = 0$.

Proof. Let $M = \sup\limits_{x \in \mathfrak{D}} f(x)$ and $m = \inf\limits_{x \in \mathfrak{D}} f(x)$, then $m \cdot \mathrm{vol}(\mathfrak{D}) \leq \displaystyle\int_{\mathfrak{D}} f(x)\, d^n x \leq M \cdot \mathrm{vol}(\mathfrak{D})$, by definition of integration. □

Theorem 4.1.10 (Fubini)**.** *For* $f : \mathfrak{D} \to \mathbb{R}$, *let the domain be defined as*

- *Type I:* $\mathfrak{D}_{\mathrm{I}} = \left\{ (y, x_n) \in \mathbb{R}^{n-1} \times \mathbb{R} \;\middle|\; y \in \mathfrak{E} \text{ and } \alpha(y) \leq x_n \leq \beta(y) \right\}$
- *Type II:* $\mathfrak{D}_{\mathrm{II}} = \left\{ (y, x_n) \in \mathbb{R}^{n-1} \times \mathbb{R} \;\middle|\; y \in \mathfrak{E}(x_n) \text{ and } a \leq x_n \leq b \right\}$
- *Type III: Type I and Type II*

then the integral is:

$$\int_{\mathfrak{D}} f(x)\, d^n x = \int_{\mathfrak{E}} \left(\int_{\alpha(y)}^{\beta(y)} f(y, x_n)\, dx_n \right) d^{n-1} y = \int_a^b \left(\int_{\mathfrak{E}(x_n)} f(y, x_n)\, d^{n-1} y \right) dx_n$$

if $\mathfrak{D}$ *is type I or II, respectively.*

Proof. Follows from 4.1.6 and 4.1.7. □

Lemma 4.1.11 (Additivity)**.** *If* $\mathfrak{D} = \mathfrak{D}_1 \cup \mathfrak{D}_2$ *and* $\mathrm{vol}(\mathfrak{D}_1 \cap \mathfrak{D}_2) = 0$, *for* $f : \mathfrak{D} \to \mathbb{R}$ *integrable over* $\mathfrak{D}$,

$$\int_{\mathfrak{D}} f(x)\, d^n x = \int_{\mathfrak{D}_1} f(x)\, d^n x + \int_{\mathfrak{D}_2} f(x)\, d^n x$$

4.2 Change of Variables

Lemma 4.2.1 (Affine Transformation)**.** *Let $\Phi(x) = A\,x + w$ be a bijective affine map and $f : \mathfrak{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ integrable:*

$$\int_{\mathfrak{D}} f(x)\, d^n x = \int_{\Phi^{-1}(\mathfrak{D})} f(\Phi(y)) \cdot |\det(A)|\, d^n y$$

which is the substitution $x = \Phi(y)$.

Proof. For $H_i = \Phi(G_i) \Rightarrow \mathrm{vol}(H_i) = |\det(A)| \cdot \mathrm{vol}(G_i)$, for H_i, G_i hyper-rectangles. Notice $m_i = \inf\limits_{x \in H_i} f(x) = \inf\limits_{y \in G_i} f(\Phi(y))$ and $M_i = \sup\limits_{x \in H_i} f(x) = \sup\limits_{y \in G_i} f(\Phi(y))$, hence: $L(T) = \sum\limits_{i=1}^{N} m_i \; \mathrm{vol}(H_i) = \sum\limits_{i=1}^{N} m_i\, |\det(A)| \; \mathrm{vol}(G_i)$ and $U(T) = \sum\limits_{i=1}^{N} M_i\, |\det(A)| \; \mathrm{vol}(G_i)$. The result follows by definition of the Darboux integral. $\square$

Definition 4.2.2 (Jacobian)**.** *For $\Phi : \mathfrak{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ a differentiable map, define $J[\Phi](x) = \det([D\Phi(x)])$*

Definition 4.2.3 (Coordinates)**.** *A function $\Phi : \mathfrak{D}_\Phi \to R_\Phi$ with $\mathfrak{D}_\Phi, R_\Phi \subseteq \mathbb{R}^n$ is a homeomorphism if:*

- *Φ is bijective;*
- *Both Φ and Φ^{-1} are continuous.*

Further Φ is also called a coordinate map. If Φ is differentiable in $\mathfrak{D}_\Phi$, and $\forall\, a \in \mathfrak{D}_\Phi\,,\ D\Phi(a) \in \mathrm{Aut}(\mathbb{R}^n)$, it is called a diffeomorphism.

Theorem 4.2.4 (Change of Variables)**.** *Let $\Phi : \mathfrak{D}_\Phi \to R_\Phi$ be a diffeomorphism and $f : \mathfrak{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ integrable (and $\mathfrak{D} \subseteq R_\Phi$):*

$$\int_{\mathfrak{D}} f(x)\, d^n x = \int_{\Phi^{-1}(\mathfrak{D})} f(\Phi(y)) \cdot |J[\Phi](y)|\, d^n y$$

which is the substitution $x = \Phi(y)$.

Proof. For $y_0 \in H_i$, then $\Phi(y) = \Phi(y_0) + D\Phi(y_0)[y - y_0] + \epsilon(y - y_0) \cdot \|y - y_0\|$, hence Φ is approximated by an affine map on H_i. By definition of integration, we get the result by 4.2.1. $\square$

Corollary 4.2.5. *If* $\mathrm{vol}(\{J[\Phi](a) = 0 \mid a \in \mathfrak{D}_\Phi\}) = 0$*, the formula is also valid.*

Remark 4.2.6. *If* Φ *is a diffeomorphism,* $J[\Phi^{-1}](\Phi(x)) = \dfrac{1}{J[\Phi](x)}$.

Definition 4.2.7 (Common Coordinate Systems)**.** *Define the following coordinate systems*

Polar $\Phi_{\mathit{polar}} : \mathbb{R}^2 \setminus \{(0,0)\} \to (0,\infty) \times [0, 2\pi)$ *where* $(\rho, \varphi) \mapsto (\rho\cos\varphi, \rho\sin\varphi)$. *The inverse is given by:* $(x, y) \mapsto (\sqrt{x^2 + y^2}, \mathrm{atan2}(y, x))$ *where:*

$$\mathrm{atan2}(y, x) = \begin{cases} \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \text{if } y \geq 0 \\ 2\pi - \arccos\left(\frac{x}{\sqrt{x^2+y^2}}\right) & \text{if } y < 0 \end{cases}$$

Cylindrical $\Phi_{\mathit{cylindrical}} : \mathbb{R}^3 \setminus \Big(\{(0,0)\} \times \mathbb{R}\Big) \to (0,\infty) \times [0, 2\pi) \times \mathbb{R}$ *where the map is* $(\rho, \varphi, z) \mapsto (\rho\cos\varphi, \rho\sin\varphi, z)$ *and the inverse is given by:* $(x, y, z) \mapsto (\sqrt{x^2 + y^2}, \mathrm{atan2}(y, x), z)$

Spherical $\Phi_{\mathit{spherical}} : \mathbb{R}^3 \setminus \Big(\{(0,0)\} \times \mathbb{R}\Big) \to (0,\infty) \times [0, 2\pi) \times [0, \pi]$ *where the map is* $(r, \varphi, \theta) \mapsto (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)$ *and the inverse is given by:* $(x, y, z) \mapsto \left(\sqrt{x^2 + y^2 + z^2}, \mathrm{atan2}(y, x), \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right)\right)$

Remark 4.2.8. *The sets where these transformation at not defined have measure zero on their respective spaces, so they can be ignored on integration.*

4.3 Path Integrals

Definition 4.3.1 (Reparametrization)**.** *Two curves $\gamma : [a,b] \to \mathbb{R}^n$ and $\delta : [c,d] \to \mathbb{R}^n$ are reparametrization if there is a homeomorphism $\varphi : [a,b] \to [c,d]$ such that $\gamma = \delta \circ \varphi$. Then, $\gamma([a,b]) = \delta([c,d]) = \Gamma$, so these are both parametrizations of the same curve in $A \subseteq \mathbb{R}^n$. If φ is increasing, then the reparametrization is orientation-preserving, otherwise, it is orientation-reversing. Further, we denote $\neg\gamma : [a,b] \to \mathbb{R}^n$ the curve $\neg\gamma(t) = \gamma(a+b-t)$.*

Definition 4.3.2 (Arc Integral)**.** *For a function $f : \mathfrak{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ and a curve $\gamma : [a,b] \to \mathfrak{D}$. We define the arc integral $\displaystyle\int_\gamma f\,d\ell$ by taking the Riemann sums $\displaystyle\sum_{i=1}^{N} f(\gamma(c_i)) \cdot \|\gamma(t_i) - \gamma(t_{i-1})\|$ for a partition T of $[a,b]$.*

Lemma 4.3.3 (Additivity)**.** *For $\gamma : [a,b] \to \mathfrak{D}$ and $\delta : [b,c] \to \mathfrak{D}$ such that $\gamma(b) = \delta(b)$, for any $f \in C^1(\mathfrak{D})$,*

$$\int_{\gamma\#\delta} f\,d\ell = \int_\gamma f\,d\ell + \int_\delta f\,d\ell$$

Proof. Let $\beta = \gamma\#\delta$ and T a partition of $[a,c]$. Define M s.t. $b \in [t_M, t_{M+1}]$. So, by definition:

$$\int_\beta f\,d\ell = \lim_{\|T\|\to 0} \sum_{i=1}^{N} f(\beta(c_i))\|\beta(t_i) - \beta(t_{i-1})\|$$

$$= \lim_{\|T\|\to 0} \left\{ \sum_{i=1}^{M} f(\gamma(c_i))\|\gamma(t_i) - \gamma(t_{i-1})\| + f(\beta(c_{M+1}))\|\delta(t_{M+1}) - \gamma(t_M)\| \right.$$

$$\left. + \sum_{i=M+2}^{N} f(\delta(c_i))\|\delta(t_i) - \delta(t_{i-1})\| \right\} = \int_\gamma f\,d\ell + 0 + \int_\delta f\,d\ell$$

since we defined induced partitions of $[a,b]$ and $[b,c]$. $\square$

Definition 4.3.4 (Rectifiable)**.** *A curve is rectifiable if*

$$L_\gamma = \sup\left\{ \sum_{i=1}^{N} \|\gamma(t_i) - \gamma(t_{i-1})\| \;\middle|\; \textit{partition } T \textit{ of } [a,b] \right\} < \infty$$

in that case, we say L_γ is the length of the curve.

Remark 4.3.5. $L_\gamma = \displaystyle\int_\gamma d\ell$, *by triangle inequality on refinements of the partition.*

Lemma 4.3.6. *If $\gamma : [a, b] \to \mathbb{R}^n$ is C^1, then γ is rectifiable.*

Proof. Let $M = \sup_{t\in[a,b]} \|\gamma'(t)\|$. Taking Lagrange's MVT (cf. Calculus I):

$$\sum_{i=1}^{N} \|\gamma(t_i) - \gamma(t_{i-1})\| = \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{n} (\gamma_j(t_i) - \gamma_j(t_{i-1}))^2}$$

$$= \sum_{i=1}^{N} (t_i - t_{i-1}) \cdot \sqrt{\sum_{j=1}^{n} \gamma_j'(c_{i,j})^2} \le \sqrt{n}\, M \sum_{i=1}^{N} (t_i - t_{i-1}) = \sqrt{n}\, M\, (b - a)$$

Hence, $L_\gamma \le \sqrt{n}\, M\, (b - a)$. □

Theorem 4.3.7. *If $\gamma : [a, b] \to \mathbb{R}^n$ is C^1, then*

$$\int_\gamma f\, d\ell = \int_a^b f(\gamma(t)) \cdot \|\gamma'(t)\|\, dt$$

Proof. By the previous calculation and taking a equipartition Δt in the Riemann sum. The formula follows. □

Theorem 4.3.8. *If $\delta : [c, d] \to \mathbb{R}^n$ is a reparametrization of $\gamma : [a, b] \to \mathbb{R}^n$, then*

$$\int_\gamma f\, d\ell = \int_\delta f\, d\ell$$

Proof. First, $\gamma = \delta \circ \varphi \Rightarrow \gamma'(t) = \delta'(\varphi(t)) \cdot \varphi'(t)$. By 4.2.4 with $\varphi : [a, b] \to [c, d]$:

$$\begin{aligned}\int_\delta f\, d\ell &= \int_{[c,d]} f(\delta(s)) \cdot \|\delta'(s)\|\, ds = \int_{[a,b]} f(\delta(\varphi(t))) \cdot \|\delta'(\varphi)\| \cdot |\varphi'(t)|\, dt \\ &= \int_{[a,b]} f(\gamma(t)) \cdot \|\gamma'(t)\|\, dt = \int_\gamma f\, d\ell\end{aligned}$$

Further, it is valid regardless if φ is orientation-preserving or reversing. □

Definition 4.3.9 (Line Integral)**.** *For a function $F : \mathfrak{D} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and a curve $\gamma : [a,b] \to \mathfrak{D}$. We define the line integral $\displaystyle\int_\gamma F \cdot d\vec{r}$ by taking the Riemann sums $\displaystyle\sum_{i=1}^{N} F(\gamma(c_i)) \cdot \Big[\gamma(t_i) - \gamma(t_{i-1})\Big]$ for a partition T of $[a,b]$.*

Theorem 4.3.10. *If $\gamma : [a,b] \to \mathbb{R}^n$ is C^1, then*

$$\int_\gamma F \cdot d\vec{r} = \int_a^b F(\gamma(t)) \cdot \gamma'(t)\, dt$$

Proof. By Lagrange's MVt, similar to previous calculatation, and taking a equipartition Δt in the Riemann sum. The formula follows. $\square$

Theorem 4.3.11. *If $\delta : [c,d] \to \mathbb{R}^n$ is a reparametrization of $\gamma : [a,b] \to \mathbb{R}^n$, then:*

(i) δ is orientation-preserving: $\displaystyle\int_\gamma F \cdot d\vec{r} = \int_\delta F \cdot d\vec{r}$

(ii) δ is orientation-reversing, $\displaystyle\int_\gamma F \cdot d\vec{r} = -\int_\delta F \cdot d\vec{r}$

Proof. First, $\gamma = \delta \circ \varphi \Rightarrow \gamma'(t) = \delta'(\varphi(t)) \cdot \varphi'(t)$. By 4.2.4 with $\varphi : [a,b] \to [c,d]$:

$$\begin{aligned}\int_\delta F \cdot d\vec{r} = \int_{[c,d]} F(\delta(s)) \cdot \delta'(s)\, ds &= \int_{[a,b]} F(\delta(\varphi(t))) \cdot \delta'(\varphi) \cdot |\varphi'(t)|\, dt \\ &= \int_{[a,b]} F(\gamma(t)) \cdot \gamma'(t)\ \operatorname{sgn}(\varphi(t))\, dt = \operatorname{sgn}(\varphi') \int_\gamma F \cdot d\vec{r}\end{aligned}$$

And $\operatorname{sgn}(\varphi') = \begin{cases} 1 & \text{if } \delta \text{ is orientation-preserving} \\ -1 & \text{if } \delta \text{ is orientation-reversing} \end{cases}.$ $\square$

Corollary 4.3.12. $\displaystyle\int_{\neg\gamma} F \cdot d\vec{r} = -\int_\gamma F \cdot d\vec{r}$

Lemma 4.3.13 (Additivity)**.** *For $\gamma : [a,b] \to \mathfrak{D}$ and $\delta : [b,c] \to \mathfrak{D}$ such that $\gamma(b) = \delta(b)$, for any $F \in C^1(\mathfrak{D}, \mathbb{R}^n)$,*

$$\int_{\gamma \# \delta} F \cdot d\vec{r} = \int_\gamma F \cdot d\vec{r} + \int_\delta F \cdot d\vec{r}$$

Proof. Follows same calculatation as 4.3.3. ☐

Definition 4.3.14 (Winding number)**.** *For closed piecewise C^1 curve $\gamma : [a,b] \to \mathbb{R}^2$, for $(x_0, y_0) \notin \gamma$, define*

$$w_\gamma(x_0, y_0) = \frac{1}{2\pi} \oint_\gamma \frac{-(y-y_0)\,dx + (x-x_0)\,dy}{(x-x_0)^2 + (y-y_0)^2}$$

Lemma 4.3.15. $w_\gamma(x_0, y_0) \in \mathbb{Z}$

Proof. For $\gamma(t) = \big(x_0 + r(t)\cos\theta(t), y_0 + r(t)\sin\theta(t)\big)$, and we require θ it continuous. Expanding: $\theta(t) = 2\pi k(t) + \operatorname{atan2}(y(t) - y_0, x(t) - x_0)$, where $k(t) \in \mathbb{Z}$ is chosen so $\theta(t)$ is continuous. Then:

$$w_\gamma(x_0, y_0) = \frac{1}{2\pi} \int_a^b \dot{\theta}(t)\,dt = \frac{\theta(b) - \theta(a)}{2\pi} = k(b) - k(a) \in \mathbb{Z}$$

since $\gamma(b) = \gamma(a)$. ☐

Definition 4.3.16. *A closed curve $\gamma : [a,b] \to \mathbb{R}^2$ is said to be a Jordan curve if $\gamma\big|_{[a,b)}$ is injective.*

Lemma 4.3.17. *Let $\gamma : [a,b] \to \mathbb{R}^2$ be a piecewise C^1 Jordan curve, then $\forall\,(x_0, y_0) \in \mathbb{R}^2 \setminus \gamma\,,\ w_\gamma(x_0, y_0) \in \{-1, 0, 1\}$.*

Theorem 4.3.18 (Jordan Curve Theorem)**.** *For $\gamma : [a,b] \to \mathbb{R}^2$ a piecewise C^1 Jordan curve, then we can decompose $\mathbb{R}^2 \setminus \gamma = \operatorname{Int}(\gamma) \sqcup \operatorname{Ext}(\gamma)$ where:*

$$\operatorname{Int}(\gamma) = \Big\{(x_0, y_0) \in \mathbb{R}^2 \;\Big|\; |w_\gamma(x_0, y_0)| = 1\Big\}$$
$$\operatorname{Ext}(\gamma) = \Big\{(x_0, y_0) \in \mathbb{R}^2 \;\Big|\; w_\gamma(x_0, y_0) = 0\Big\}$$

Moreover, $\operatorname{Int}(\gamma)$ is bounded and $\operatorname{Ext}(\gamma)$ is unbounded and both are connected.

Definition 4.3.19. *A closed curve $\gamma : [a,b] \to \mathbb{R}^2$ is positively oriented if $\forall\,(x_0, y_0) \in \mathbb{R}^2 \setminus \gamma\,,\ w_\gamma(x_0, y_0) \geq 0$, and negatively oriented if it is ≤ 0.*

4.4 Conservative Fields

Definition 4.4.1. *Let $\mathfrak{D} \subseteq \mathbb{R}^n$ be an open connected domain. A vector field $F : \mathfrak{D} \to \mathbb{R}^n$ is called conservative iff $\exists\, U \in C^1(\mathfrak{D}) : F = \nabla U$. Then U is called a potential of the vector field.*

Lemma 4.4.2. *The potential is unique up to a constant.*

Proof. Follows directly from linearity and 3.1.17. □

Theorem 4.4.3 (Gradient)**.** *Let $F : \mathfrak{D} \to \mathbb{R}^n$ be C^1 conservative field and U a potential function. Then, for any C^1 curve $\gamma : [a, b] \to \mathfrak{D}$,*

$$\int_\gamma F \cdot d\vec{r} = U(\gamma(b)) - U(\gamma(a))$$

Proof. Follows directly from 3.1.16, 4.3.10 and FTC I (cf. Calculus I). □

Theorem 4.4.4 (Converse Gradient)**.** *Let $F : \mathfrak{D} \to \mathbb{R}^n$ be C^1, then the following are equivalent:*

(i) F is conservative;

(ii) For any closed C^1 curve $\gamma \subset \mathfrak{D}$, $\oint_\gamma F \cdot d\vec{r} = 0$

(iii) For any C^1 curve $\gamma \subset \mathfrak{D}$, $\int_\gamma F \cdot d\vec{r}$ depends only on the endpoints of γ.

Proof. The directions $(i) \Rightarrow (ii)$ and $(i) \Rightarrow (iii)$ follow from the previous theorem.

$(ii) \Rightarrow (iii)$ Take two curves $\gamma_1, \gamma_2 : [a, b] \to \mathfrak{D}$ with same endpoints. Then: $\gamma_1 \#(\neg\gamma_2)$ is a closed curve, hence:

$$0 = \oint_{\gamma_1 \#(\neg\gamma_2)} F \cdot d\vec{r} = \int_{\gamma_1} F \cdot d\vec{r} - \int_{\gamma_2} F \cdot d\vec{r}$$

$(iii) \Rightarrow (i)$ For a fixed $x_0 \in \mathfrak{D}$, let $U(x) = \int_{x_0 \to x} F \cdot d\vec{r}$ where $x_0 \to x$ is any differentiable curve in $\mathfrak{D}$ connecting x_0 to x. We take a linear path:

$$\begin{aligned}\frac{\partial \phi}{\partial x_i} &= \lim_{\delta \to 0} \frac{1}{\delta}\left[\int_{x_0 \to x + \delta\, e_i} F \cdot d\vec{r} - \int_{x_0 \to x} F \cdot d\vec{r}\right] \\ &= \lim_{\delta \to 0} \frac{1}{\delta} \int_{x \to x + \delta\, e_i} F \cdot d\vec{r} = \lim_{\delta \to 0} \frac{1}{\delta} \int_0^\delta F(x + t\, e_i) \cdot e_i\, dt = F_i(x)\end{aligned}$$

$\square$

Theorem 4.4.5 (Irrotational $\Leftrightarrow$ Conservative)**.** *For $F = (F_1, \cdots, F_n) \in C^1(\mathfrak{D}, \mathbb{R}^n)$ where $\mathfrak{D} = \prod_{i=1}^n [a_i, b_i]$ is a hyperrectangle, if $[DF(a)]$ is symmetric (F is called irrotational), that is:*

$$\forall\, i, j \in \{1, \cdots, n\}\,,\ \frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$$

then, F is conservative.

Proof. Let α_i be:

$$\begin{aligned}
\alpha_1(t, x) &= (t, x_2, x_3, \cdots, x_{n-1}, x_n) \\
\alpha_2(t, x) &= (a_1, t, x_3, \cdots, x_{n-1}, x_n) \\
\vdots \quad & \quad \vdots \\
\alpha_n(t, x) &= (a_1, a_2, a_3, \cdots, a_{n-1}, t)
\end{aligned}$$

Then, for $U(x) = \displaystyle\sum_{i=1}^{n} \int_{a_i}^{x_i} F_i(\alpha_i(t, x))\, dt$, we show $F = \nabla U$. Notice the identity $\alpha_i(a_i, x) = \alpha_{i+1}(x_{i+1}, x)$. Then, we calculate:

$$\begin{aligned}
\frac{\partial U}{\partial x_j} &= F_j(\alpha_j(x_j, x)) + \sum_{i=1}^{j-1} \int_{a_i}^{x_i} \frac{\partial F_i}{\partial x_j}\,(\alpha_i(t, x))\, dt \\
&= F_j(\alpha_j(x_j, x)) + \sum_{i=1}^{j-1} \int_{a_i}^{x_i} \frac{\partial F_j}{\partial x_i}\,(\alpha_i(t, x))\, dt \\
&= F_j(\alpha_j(x_j, x)) + \sum_{i=1}^{j-1} F_j(\alpha_i(t, x))\Big|_{a_i}^{x_i} \\
&= F_j(\alpha_j(x_j, x)) + \sum_{i=1}^{j-1} \Big[F_j(\alpha_i(x_i, x)) - F_j(\alpha_{i+1}(x_{i+1}, x))\Big] \\
&= F_j(\alpha(x_1, x)) = F_j(x)
\end{aligned}$$

The result follows. $\square$

Remark 4.4.6. *The converse is true by 3.2.3.*

Remark 4.4.7 (Circle and Disk)**.** *Define the following sets.*

- $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$
- $D^2 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 1\}$

Notice $S^1 \subset D^2$.

Definition 4.4.8. *We say* $\gamma : [0,1] \to \mathfrak{D} \subseteq \mathbb{R}^n$ *is a loop if* $\gamma(0) = \gamma(1)$. *Then there is a function* $\zeta : S^1 \to \mathfrak{D}$ *defined* $\gamma(x) = \zeta\big(\cos(2\pi x), \sin(2\pi x)\big)$ *or* $\zeta(x,y) = \gamma\left(\dfrac{1}{2\pi}\operatorname{atan2}(y,x)\right)$ *(cf. 4.2.7).*

Definition 4.4.9 (Simply Connected Region)**.** *A subset* $A \subseteq \mathbb{R}^n$ *is simply connected if it is path-connected (cf. 2.6.6) and for any loop* $f : S^1 \to A$, *there is a continuous extension* $F : D^2 \to A$. *That is, we can "shrink" any closed curve to a point.*

Theorem 4.4.10 (Green)**.** *Let* $\mathfrak{D} \subseteq \mathbb{R}^2$ *be a simply connected domain and* $\partial D = \gamma$ *positively oriented. Then, for any* $F = (P,Q) \in C^1(\mathfrak{D}, \mathbb{R}^2)$*:*

$$\oint_\gamma F \cdot d\vec{r} = \oint_\gamma P\,dx + Q\,dy = \iint_{\mathfrak{D}} (Q_x - P_y)\,dA$$

Proof. We'll prove only for domains of the type III:

$$\begin{aligned}\mathfrak{D} &= \{(x,y) \in \mathbb{R}^2 \mid a \le x \le b \text{ and } g_1(x) \le y \le g_2(x)\} \\ &= \{(x,y) \in \mathbb{R}^2 \mid f_1(y) \le x \le f_2(y) \text{ and } \alpha \le y \le \beta\}\end{aligned}$$

We calculate $\displaystyle\oint_\gamma P\,\hat{x} \cdot d\vec{r}$. We split γ into four curves:

$$\gamma_1 = \{(x, g_1(x)) \mid a \le x \le b\} \Rightarrow \oint_{\gamma_1} P\,\hat{x} \cdot d\vec{r} = \int_a^b P(x, g_1(x)) \cdot dx$$

$$\gamma_2 = \{(a, y) \mid g_1(a) \le y \le g_2(a)\} \Rightarrow \oint_{\gamma_2} P\,\hat{x} \cdot d\vec{r} = 0$$

$$\gamma_3 = \{(x, g_2(x)) \mid a \le x \le b\} \Rightarrow \oint_{\gamma_3} P\,\hat{x} \cdot d\vec{r} = -\int_a^b P(x, g_2(x)) \cdot dx$$

$$\gamma_4 = \{(b, y) \mid g_1(b) \le y \le g_2(b)\} \Rightarrow \oint_{\gamma_4} P\,\hat{x} \cdot d\vec{r} = 0$$

since the curves γ_2 and γ_4 are perpendicular to the x-axis. Hence:

$$\oint_\gamma P\,\hat{x}\cdot d\vec{r} = \int_a^b \Big[P(x, g_1(x)) - P(x, g_2(x))\Big]\,dx = -\int_{x=a}^{b}\int_{y=g_1(x)}^{g_2(x)} P_y\,dy\,dx$$

A similar calculation holds for $\displaystyle\oint_\gamma Q\,\hat{y}\cdot d\vec{r} = \int_{y=\alpha}^{\beta}\int_{x=f_1(y)}^{f_2(y)} Q_x\,dx\,dy$. The result follows from linearity. Further, for type I and type II, where $Q = 0$ or $P = 0$, respectively, it is analogous. $\square$

Lemma 4.4.11 (Joining Domains)**.** *For two domains $\mathfrak{D}_1, \mathfrak{D}_2$ with* $\mathrm{vol}(\mathfrak{D}_1 \cap \mathfrak{D}_2) = 0$*, if Green's Theorem applies to both $\mathfrak{D}_1$ and $\mathfrak{D}_2$, then it applies for $\mathfrak{D}_1 \cup \mathfrak{D}_2$.*

Proof. Let $\gamma_1 = \partial A \setminus \partial(A \cap B)$, $\gamma_2 = \partial B \setminus \partial(A \cap B)$, $\gamma_3 = \partial(A \cup B)$. Hence: $\gamma_3 = \gamma_1 \# \gamma_2$ and the result follows from additivity of both the line integral and the double integral. $\square$

Remark 4.4.12. *A full proof for Green's Theorem for rectifiable Jordan curves γ can be given by chopping up the domain into rectangles and taking the limit together with the previous lemma.*

Corollary 4.4.13. *If $F \in C^1(\mathfrak{D}, \mathbb{R}^n)$ is irrotational, where $\mathfrak{D}$ is simply connected, then F is conservative.*

4.5 Flux Integrals

Definition 4.5.1 (Surface)**.** *A surface in $\mathbb{R}^3$ is a continuous function (cf. 2.4.1) $\sigma : S \subseteq \mathbb{R}^2 \to \mathbb{R}^3$, where S is simply connected. We refer to $\sigma(S) = \Sigma$ as the surface and σ as a parametrization.*

Definition 4.5.2 (Surface Integral)**.** *For a function $f : \Sigma \subseteq \mathbb{R}^3 \to \mathbb{R}$ defined on the surface Σ (parametrized by σ which we ask to be C^1). We define the surface integral*

$$\iint_\Sigma f\, dS = \iint_S f(\sigma(s,t)) \cdot \left\| \frac{\partial \sigma}{\partial s} \times \frac{\partial \sigma}{\partial t} \right\| ds\, dt$$

Lemma 4.5.3 (Additivity)**.** *For two surfaces Σ, Π with area$(\Sigma \cap \Pi) = 0$, for any $f \in C^1(\Sigma \cup \Pi)$,*

$$\iint_{\Sigma \cup \Pi} f\, dS = \iint_\Sigma f\, dS + \iint_\Pi f\, dS$$

Proof. Follows from additivity of double integrals. □

Theorem 4.5.4. *If $\pi : R \to \mathbb{R}^n$ is a reparametrization of $\sigma : S \to \mathbb{R}^n$ (that is, there is a homeomorphism $\varphi : S \to R$ such that $\sigma = \pi \circ \varphi$), then*

$$\iint_\Sigma f\, dS = \iint_\Pi f\, dS$$

Proof. First, $\sigma = \pi \circ \varphi \Rightarrow \sigma_s = \pi_\lambda \cdot \lambda_s + \pi_\mu \cdot \mu_s$ and $\sigma_t = \pi_\lambda \cdot \lambda_t + \pi_\mu \cdot \mu_t$, where $(\lambda, \mu) = \varphi(s,t)$. By 4.2.4:

$$\begin{aligned}
\iint_\Pi f\, dS &= \iint_R f(\pi(s,t)) \cdot \|\pi_\lambda(\lambda,\mu) \times \pi_\mu(\lambda,\mu)\|\, d\lambda\, d\mu \\
&= \iint_S f(\pi(\varphi(s,t))) \cdot \|\pi_\lambda(\varphi(s,t)) \times \pi_\mu(\varphi(s,t))\| \cdot \left\| \begin{vmatrix} \lambda_s & \lambda_t \\ \mu_s & \mu_t \end{vmatrix} \right\| ds\, dt \\
&= \iint_S f(\gamma(t)) \cdot \|\sigma_s(s,t) \times \sigma_t(s,t)\|\, ds\, dt = \iint_\Sigma f\, dS
\end{aligned}$$

where we calculated:

$$\sigma_s(s,t) \times \sigma_t(s,t) = \pi_\lambda(\varphi(s,t)) \times \pi_\mu(\varphi(s,t)) \cdot \begin{vmatrix} \lambda_s & \lambda_t \\ \mu_s & \mu_t \end{vmatrix}$$

□

Definition 4.5.5 (Flux Integral)**.** *For a function $F : \Sigma \to \mathbb{R}^3$ defined on the surface Σ (parametrized by σ which we ask to be C^1). We define the flux integral*

$$\iint_\Sigma F \cdot d\vec{S} = \iint_S F(\sigma(s,t)) \cdot \left(\frac{\partial \sigma}{\partial s} \times \frac{\partial \sigma}{\partial t} \right) ds\, dt$$

The normal to the surface is defined as: $\hat{n} = \dfrac{\sigma_s \times \sigma_t}{\|\sigma_s \times \sigma_t\|}$*, hence we can define:* $\displaystyle\iint_\Sigma F \cdot d\vec{S} = \iint_\Sigma F \cdot \hat{n}\, dS$*.*

Theorem 4.5.6 (Stoke's)**.** *For a surface Σ, parametrized by $\sigma : S \to \Sigma$, let $\Gamma = \sigma(\partial S)$, which we'll call the boundary of the surface, then for any $F \in C^1(\Sigma, \mathbb{R}^3)$:*

$$\int_\Gamma F \cdot d\vec{r} = \iint_\Sigma \big(\nabla \times F\big) \cdot d\vec{S}$$

Proof. We'll reduce it to Green's Theorem: Let $\sigma : S \to \Sigma$. Take

$$G(s,t) = (P(s,t), Q(s,t)) = (F(\sigma(s,t)) \cdot \sigma_s, F(\sigma(s,t)) \cdot \sigma_t)$$

Take the curve $\Delta = \partial S$, so that $\Gamma = \vec{\sigma}(\Delta)$, we get:

$$\oint_\Gamma F \cdot d\vec{r} = \oint_\Delta G \cdot d\vec{r} = \iint_S \big[Q_s - P_t\big]\, ds\, dt$$

by Green's Theorem. By direct calculation, we have:

$$Q_s - P_t = (\nabla \times F)(\sigma(s,t)) \cdot \big[\sigma_s \times \sigma_t\big]$$

The result follows by the definition of the flux integral. □

Theorem 4.5.7 (Gauß's)**.** *For a solid $\Omega \subseteq \mathbb{R}^3$ with boundary $\partial\Omega = \Sigma$, for any $F \in C^1(\Omega, \mathbb{R}^3)$:*

$$\iint_\Sigma F \cdot d\vec{S} = \iiint_\Omega \big(\nabla \cdot F\big)\, dV$$

Proof. The proof goes analogous to Green's Theorem, by separation into rectangles and connecting the domains. □