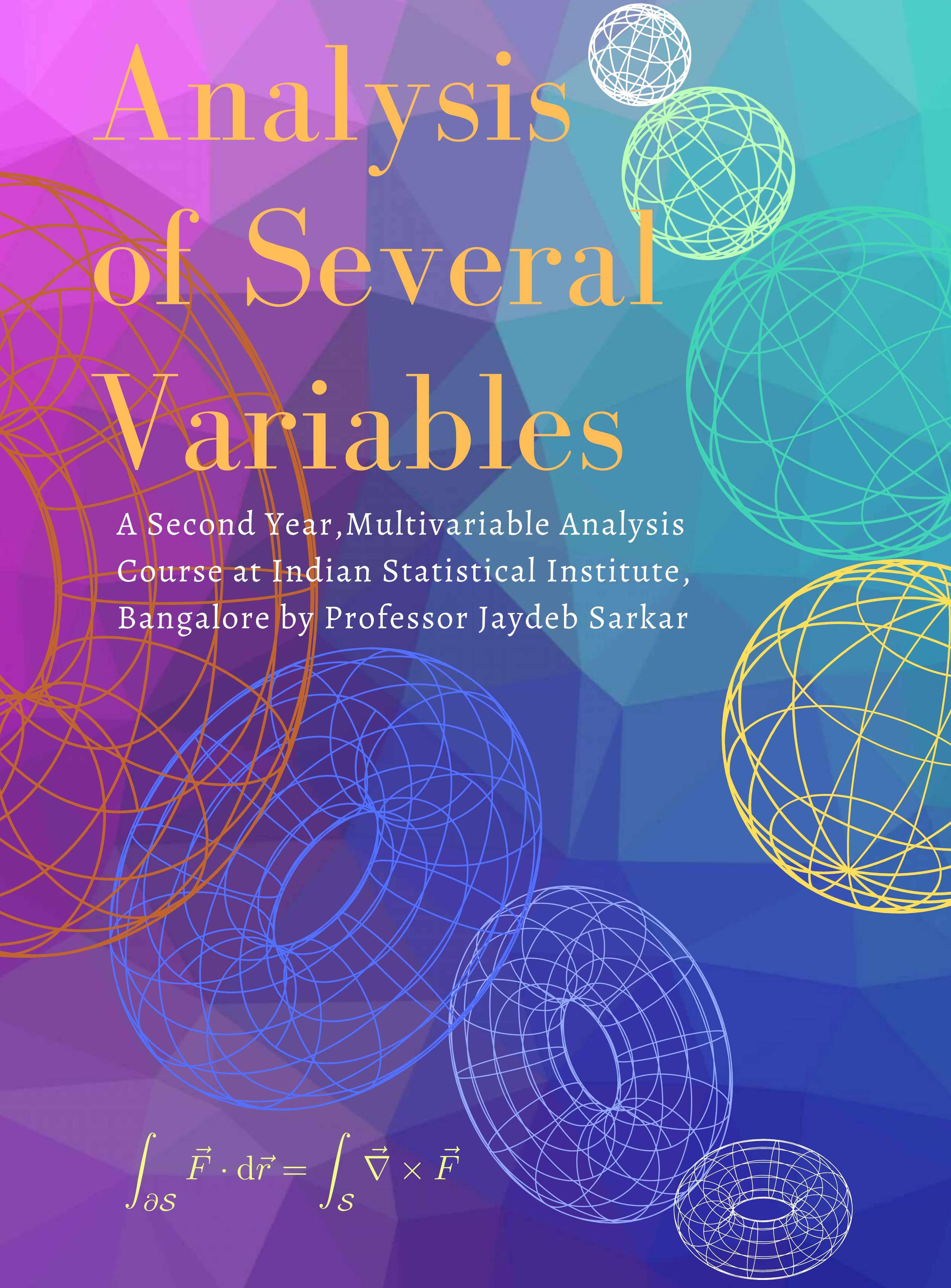
Analysis of Several Variables

A Second Year,Multivariable Analysis
Course at Indian Statistical Institute,
Bangalore by Professor Jaydeb Sarkar



$$\int_{\partial S} \vec{F} \cdot \mathrm{d}\vec{r} = \int_{S} \vec{\nabla} \times \vec{F}$$

Analysis of Several Variables

Jaydeb Sarkar

Statistics and Mathematics Unit
Indian Statistical Institute, Bangalore

LaTeXed by

Aaratrick Basu
Bikram Halder
Deepta Basak
Priyatosh Jana
Saraswata Sensarma
Soumya Dasgupta
Trishan Mondal

B. Math (Hons.) (2021-'24)
Indian Statistical Institute, Bangalore

Availability

The PDF copies of these documents are freely available at the following locations:

- **Main Link:** `https://awnathan1893.github.io/Analysis3_Notes/`
  This link is regularly updated to reflect changes in the original source code.
- **Mirror Link:** `https://bikramhalder.github.io/Analysis3_Notes/`
  The mirror link is updated every Wednesday and Saturday.

You can access the source code for all these documents on GitHub at: `https://github.com/AWNathan1893/Analysis3_Notes`

Contents

List of Notation

$\ell(\gamma)$	Arclength of a rectifiable curve γ
$B_r(a)$	Open ball centered at $a \in \mathbb{R}^n$ of radius r
∂S	Boundary of a set $S \subseteq \mathbb{R}^n$
$\mathscr{B}(S)$	Space of all bounded functions defined over S
$B, B^n = \prod_{i=1}^{n}[a_i, b_i]$	Closed box in $\mathbb{R}^n$
$C^k(S)$	Class of all functions defined on S for which all partial derivatives on S upto order k exist and are continuous
$\overline{S}$	Closure of $S \subseteq \mathbb{R}^n$
$\boldsymbol{\nabla} \times \vec{F}$	Curl of $\vec{F}$
$D_r(a)$	Deleted ball of radius r around $a \in \mathbb{R}^n$
$d(A)$	Diameter of the set $A \subseteq \mathbb{R}^n$
$(\boldsymbol{\nabla} \cdot h)^m = \sum_{\lvert\alpha\rvert = m} \frac{m!}{\alpha!} h^\alpha \partial^\alpha$	Differential order of order m and coefficients h
$(D_u f)(a)$	Directional derivative of f along u at a
$\boldsymbol{\nabla} \cdot \vec{F}$	Divergence of $\vec{F}$
$\mathrm{Ext}(S)$	Exterior of a set $S \subseteq \mathbb{R}^n$
$\boldsymbol{\nabla} f$	Gradient of f
$\boldsymbol{\nabla} f(x)$	Gradient of f at x
$\mathcal{G}(f)$	Graph of a $f : \mathbb{R}^2 \to \mathbb{R}$
$H_f(x)$	Hessian of f at x
$f_{x_i x_j}(a), \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}, \frac{\partial^2 f}{\partial x_i \partial x_j}$	Partial derivative of f with respect to the i^{th} and j^{th} coordinates at a
χ_S	Indicator function of S
$\langle x, y \rangle$	Euclidean inner product of x and y
$\mathscr{R}(S)$	Space of all Riemann-Darboux integrable functions over S
$\mathrm{Int}(S), S^\circ$	Interior of a set $S \subseteq \mathbb{R}^n$
$\int_S f \mathrm{d}V$	Riemann-Darboux integral of f over S
$J_f(a)$	Jacobian matrix of f at a
$\Delta f(x)$	Laplacian of f at x

$\lim\limits_{x\to a} f$	Limit of f as x tends to a
S'	Set of limit points of $S \subseteq \mathbb{R}^n$
$\int_\gamma f$, $\int_{\mathcal{C}} f$	Line integral of f over $\mathcal{C} = \operatorname{ran}(\gamma)$
$\int_\gamma \vec{F}\cdot \mathrm{d}\vec{r}$, $\int_{\mathcal{C}} \vec{F}\cdot \mathrm{d}\vec{r}$	Line integral of $\vec{F}$ over $\mathcal{C} = \operatorname{ran}(\gamma)$
$\underline{\int}_B f$	Lower Riemann integral of f over B
$L(f,P)$	Lower Riemann sum of f with respect to the partition P
$\lVert P\rVert$	Mesh of the partition P
$\partial^\alpha = \dfrac{\partial^{\lvert\alpha\rvert}}{\partial {x_1}^{\alpha_1}\cdots\partial {x_n}^{\alpha_n}}$	α^{th} derivative, for $\alpha\in\mathbb{Z}^n_{\geq 0}$
$\alpha! = \prod\limits_{i=1}^{n} \alpha_i!$	Factorial of $\alpha\in\mathbb{Z}^n_{\geq 0}$
$\lvert\alpha\rvert = \sum\limits_{i=1}^{n} \alpha_i$	Modulus of $\alpha\in\mathbb{Z}^n_{\geq 0}$
$\lVert x\rVert$	Euclidean norm of $x\in\mathbb{R}^n$
$\mathcal{O}_n$, $\mathcal{O}$	Open set in $\mathbb{R}^n$
$\mathscr{P}(B^n)$	Set of all partitions of the box B^n
$\dfrac{\partial f}{\partial x_i}$, $f_{x_i}(a)$	Partial derivative of f with respect to the i^{th} coordinate at a
$\mathbb{R}[x_1,\ldots,x_n]$	Space of polynomials with real coefficients in n variables
$\mathcal{P}(S)$	Power set of the set S
Π_i	Projection onto i^{th} coordinate
$Q_A(x) = x^tAx$	Quadratic form associated to the matrix A evaluated at x
$\mathbb{R}^n$	$n-$dimensional Euclidean space
$\int_{\mathcal{S}} f\,\mathrm{d}S$	Surface integral of f over $\mathcal{S}$
$\int_{\mathcal{S}}\vec{F}\cdot \mathrm{d}\vec{S} = \int_{\mathcal{S}}\vec{F}\cdot\vec{n}\,\mathrm{d}S$	Surface integral of $\vec{F}$ over $\mathcal{S}$
$T_P\mathcal{S}$	Tangent plane of $\mathcal{S}$ through P
$p_{a,k}(f,h) = \sum\limits_{\lvert\alpha\rvert\leq k}\dfrac{1}{\alpha!}(\partial^\alpha f)(a)h^\alpha$	Taylor polynomial of f at a, of order k and difference h
$Df(a)$	Derivative of f at a
$\overline{\int}_B f$	Upper Riemann integral of f over B
$U(f,P)$	Upper Riemann sum of f with respect to the partition P
$\mathrm{Vol}(S)$	Volume of the (measurable) set S

Lecture 1

1.1 Introduction

We will talk about n-variable calculus, that is, calculus on $\mathbb{R}^n$. Recall the following:

- The setting is,

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \cdots \times \mathbb{R}}_{n \text{ times}} = \{x = (x_1, \ldots, x_n) \mid x_i \in \mathbb{R} \, \forall \, i = 1, 2, \ldots, n\}$$

- Analysis on $\mathbb{R}$ consisted of ideas like open sets, compact sets, convergence, limits, differentiability, integrability etc.
- $\mathbb{R}^n$ is an n-dimensional inner product space over $\mathbb{R}$, with the standard orthonormal basis $\{e_j\}_{j=1}^n$

Extending the analytic ideas to $\mathbb{R}^n$ exploiting the algebraic structure is the matter of this course, which further gives way to differential geometry.

1.2 Review: $\mathbb{R}^n$ as a vector space

(i) The standard orthonormal basis of $\mathbb{R}^n$ is $\{e_i\}_{i=1}^n$.

(ii) For all $x \in \mathbb{R}^n$, there is a unique representation

$$x = \sum_{i=1}^{n} x_i e_i, \; x_i \in \mathbb{R}$$

Thus we identify x with the *coordinates* $(x_1, x_2, \ldots, x_n)$.

(iii) Euclidean inner product on $\mathbb{R}^n$:
For $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, we define

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

1.3 Linear Functions

In doing analysis on $\mathbb{R}$, the main motive was to study functions $f : \mathbb{R} \to \mathbb{R}$ and their properties, namely continuity, differentiability, integrability etc. We now wish to do the same for functions $f : \mathbb{R}^n \to \mathbb{R}^m$ for arbitrary natural numbers n, m.
Two easy examples of such functions are:

(i) Constant maps

$$f : \mathbb{R}^n \to \mathbb{R}^m$$
$$\forall\, x \in \mathbb{R}^n, f(x) = a,\ a \in \mathbb{R}^m$$

(ii) Linear maps

A function $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear if for all $\alpha \in \mathbb{R}, x, y \in \mathbb{R}^n$, $L(\alpha x + y) = \alpha L(x) + L(y)$.

It turns out that linear maps are useful in understanding most other '*nice*' functions, and so we now look at these in more detail.

Let L be a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. Consider the domain $\{\alpha x + y \mid \alpha \in \mathbb{R}\}$, that is, the line through y in the direction of x. The image under L is,

$$\{\alpha Lx + Ly \mid \alpha \in \mathbb{R}\}$$

which is the line through Ly in the direction of Lx. Hence, L maps lines to lines.

Exercise. Is the converse also true?

Matrix representation of a linear map

Let $L : \mathbb{R} \to \mathbb{R}$ be a linear map. Then,

$$L(x) = xL(1) \quad \forall\, x \in \mathbb{R}$$

Therefore,

$$\mathcal{L}(\mathbb{R}, \mathbb{R}) := \{\text{set of all linear maps from } \mathbb{R} \text{ to } \mathbb{R}\} \leftrightarrow \mathbb{R}$$

Now consider the general case; let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. If we fix the bases $\{e_j\}_{j=1}^n$ of $\mathbb{R}^n$ and $\{e_i\}_{i=1}^m$ of $\mathbb{R}^m$, L is determined uniquely by the equations

$$Le_j = \sum_{i=1}^m a_{ij} e_i$$

and hence,

$$L \leftrightarrow (a_{ij})_{m\times n} \in M_{m,n}(\mathbb{R})$$

1.4 Analytic ideas in $\mathbb{R}^n$

We have the Euclidean norm on $\mathbb{R}^n$ defined by,

$$\|x\| = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}} \qquad \forall\, x \in \mathbb{R}^n$$

This induces the metric given as,

$$d(x, y) = \|x - y\| = \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{\frac{1}{2}} \qquad \forall\, x, y \in \mathbb{R}^n$$

Theorem 1.4.1 (Cauchy-Schwarz Inequality)

For all $x, y \in \mathbb{R}^n$,

$$\langle x, y\rangle \le \|x\|\|y\|$$

Proof. Consider $x, y \in \mathbb{R}^n$. We have,

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i y_j - x_j y_i)^2 \geq 0$$

But the left-hand side is, after expanding,

$$\sum_{i,j=1}^{n} x_i^2 y_j^2 + \sum_{i,j=1}^{n} x_j^2 y_i^2 - 2\sum_{i,j=1}^{n} x_i x_j y_i y_j = 2\|x\|^2\|y\|^2 - 2\langle x, y\rangle^2$$

which gives the desired inequality. ☐

Note: The proof shows that equality holds only if there is $\lambda \in \mathbb{R}$ such that for all i, either $x_i = \lambda y_i$ or $y_i = \lambda x_i$.

Recall the triangle inequality for $\mathbb{R}$, for all $x, y \in \mathbb{R}$

$$|x + y| \leq |x| + |y|$$

Theorem 1.4.2 (Triangle inequality for $\mathbb{R}^n$)

For all $x, y \in \mathbb{R}^n$,

$$\|x + y\| \leq \|x\| + \|y\|$$

Proof. We have,

$$\begin{aligned} \|x+y\|^2 &= \langle x + y, x + y\rangle \\ &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| && \text{(Cauchy Schwarz inequality)} \\ \implies \|x + y\| &\leq \|x\| + \|y\| \end{aligned}$$

which is the desired inequality. ☐

The following is a technical result, which can be thought of as an analogue of the Lipschitz condition for linear maps on $\mathbb{R}^n$. It hints towards continuity of linear maps, and we will see that it is indeed so later.

Theorem 1.4.3

Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. There is $M > 0$ such that

$$\|Lx\| \leq M\|x\| \quad \forall x \in \mathbb{R}^n$$

Proof. We have, for $x = \sum_{i=1}^{n} x_i e_i$,

$$\begin{aligned} \|Lx\| &= \left\|\sum_{i=1}^{n} x_i L e_i\right\| \\ &\leq \sum_{i=1}^{n} |x_i|\|Le_i\| && \text{(Triangle inequality)} \\ \implies \|Lx\| &\leq \|x\|\left(\sum_{i=1}^{n}\|Le_i\|^2\right)^{\frac{1}{2}} && \text{(Cauchy Schwarz inequality)} \end{aligned}$$

Taking $M = \left(\sum_{i=1}^{n}\|Le_i\|^2\right)^{\frac{1}{2}}$, we get the result. ☐

Lecture 2

2.1 Distance and Topology in $\mathbb{R}^n$

Using the inner product on $\mathbb{R}^n$, we get the Euclidean distance

$$d(x,y) = \|x-y\| = \sqrt{\langle x-y, x-y\rangle} = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}$$

We wish to extend the notions of $(\mathbb{R}, |\cdot|)$ to $(\mathbb{R}^n, \|\cdot\|)$. We already know from Lecture 1 that the triangle inequality holds,

$$\|x+y\| \leq \|x\| + \|y\|$$

Definition 2.1.1 ▶ Open balls

The open ball centered at $a \in \mathbb{R}^n$ of radius r is,

$$B_r(a) = \{x \in \mathbb{R}^n \mid \|x-a\| < r\}$$

Exercise. Show that open balls are convex sets.

Definition 2.1.2 ▶ Open sets

A set $\mathcal{O} \subseteq \mathbb{R}^n$ is open if $\mathcal{O} = \phi$ or for all $x \in \mathcal{O}$, there is $r > 0$ such that $B_r(x) \subseteq \mathcal{O}$.

Example 2.1.1

(i) Any open ball is open.

(ii) We define open boxes in $\mathbb{R}^n$ to be the subsets of the form $\prod_{i=1}^{n}(a_i, b_i)$. Any open box is open.

Definition 2.1.3 ▶ Convergence of Sequences

Let $\{x_m\}_{m\in\mathbb{N}} \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$. We say $x_m \longrightarrow x$ if for all $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that

$$\begin{aligned} &\|x_m - x\| < \varepsilon \,\forall\, m \geq N \\ \iff &d(x_m, x) < \varepsilon \quad \forall\, m \geq N \\ \iff &x_m \in B_\varepsilon(x) \quad \forall\, m \geq N \end{aligned}$$

Exercise. Show that the limit of a sequence in $\mathbb{R}^n$ is unique whenever it exists.

Definition 2.1.4 ▶ Limit points

We define the deleted ε-neighbourhood of $a \in \mathbb{R}^n$ to be $D_\varepsilon(a) = B_\varepsilon(a) \setminus \{a\}$. The point a is a limit point of $S \subseteq \mathbb{R}^n$ if for all $\varepsilon > 0, D_\varepsilon(a) \cap S \neq \phi$. If we do not delete a, we get isolated points.

Definition 2.1.5 ▶ Projections

For all $i \in \{1, 2, \ldots, n\}$ we define the maps

$$\begin{aligned} \Pi_i : \mathbb{R}^n &\to \mathbb{R} \\ x = (x_1, \ldots, x_n) &\mapsto x_i \end{aligned}$$

Π_i is called the projection onto the ith coordinate.

Theorem 2.1.1

Let $\{x_m\}_{m\in\mathbb{N}} \cup \{x\} \subseteq \mathbb{R}^n$. Then,

$$\begin{aligned} & x_m \longrightarrow x \\ \iff & \Pi_i(x_m) \longrightarrow \Pi_i(x) \, \forall \, i \in \{1, 2, \ldots, n\} \end{aligned}$$

Proof. Assume $x_m \longrightarrow x$. Now, for all $j \in \{1, 2, \ldots, n\}$,

$$\begin{aligned} \|x_m - x\|^2 = \sum_{i=1}^{n} |\Pi_i(x_m) - \Pi_i(x)|^2 &\geq |\Pi_j(x_m) - \Pi_j(x)|^2 \\ \implies |\Pi_j(x_m) - \Pi_j(x)| &\longrightarrow 0 \\ \implies \Pi_j(x_m) &\longrightarrow \Pi_j(x) \end{aligned}$$

Now assume $\Pi_j(x_m) \longrightarrow \Pi_j(x)$ for all $j \in \{1, 2, \ldots, n\}$. Then,

$$\begin{aligned} |\Pi_j(x_m) - \Pi_j(x)| &\longrightarrow 0, \, \forall \, j \in \{1, 2, \ldots, n\} \\ \implies \sum_{i=1}^{n} |\Pi_j(x_m) - \Pi_j(x)|^2 &\longrightarrow 0 \\ \implies \|x_m - x\|^2 &\longrightarrow 0 \\ \implies x_m &\longrightarrow x \end{aligned}$$

☐

Definition 2.1.6 ▶ Closed sets

A set $C \subseteq \mathbb{R}^n$ is closed if $\mathbb{R}^n \setminus C$ is open.

Exercise. Show that a set $C \subseteq \mathbb{R}^n$ is closed iff $\forall \, \{x_m\}_{m\in\mathbb{N}} \subseteq C$ with $x_m \longrightarrow x$ for some $x \in \mathbb{R}^n$, we have $x \in C$.

Exercise. Show that:

(1) Arbitrary union of open sets is open.

(2) Finite intersection of open sets is open.

(3) Arbitrary intersection of closed sets is closed.

(4) Finite union of closed sets is closed.

(5) Any finite subset of $\mathbb{R}^n$ is closed.

Definition 2.1.7 ▶ Interior of a set

Let $\phi \neq S \subseteq \mathbb{R}^n$. The interior of S is,

$$\operatorname{Int}(S) = \{a \in S \mid \exists\, r > 0, B_r(a) \subseteq S\}$$

Exercise. Show that:

(1) For any nonempty set $S \subseteq \mathbb{R}^n$, $\operatorname{Int}(S)$ is open.

(2) A set S is open iff $\operatorname{Int}(S) = S$.

Definition 2.1.8 ▶ Exterior of a set

Let $\phi \neq S \subseteq \mathbb{R}^n$. The exterior of S is,

$$\operatorname{Ext}(S) = \{a \in \mathbb{R}^n \mid \exists\, r > 0, B_r(a) \cap S = \phi\}$$

Exercise. Show that $\operatorname{Ext}(S) = \operatorname{Int}(\mathbb{R}^n \setminus S)$.

Example 2.1.2

For $S = [0,2] \setminus \{1\} = [0,1) \cup (1,2]$, $1 \notin \operatorname{Ext}(S)$.

Definition 2.1.9 ▶ Boundary of a set

Let $\phi \neq S \subseteq \mathbb{R}^n$. The boundary of S is,

$$\partial S = \{a \in \mathbb{R}^n \mid \forall\, r > 0, B_r(a) \cap S \neq \phi \text{ and } B_r(a) \cap (\mathbb{R}^n \setminus S) \neq \phi\}$$

Example 2.1.3

For $S = [0,1) \cup (1,2] \cup \{5\}$, $\partial S = \{0,1,2,5\}$ but the set of limit points is $\{0,1,2\}$.

Exercise. Show that:

(1) S is open iff $S \cap \partial S = \phi$.

(2) S is closed iff $S \supseteq \partial S$.

(3) S is closed iff $S = \overline{S} =: S \cup \partial S = S \cup \{\text{Limit points of } S\}$

(4) $\overline{S} = \operatorname{Int}(S) \sqcup \partial S$. This gives the partition $\mathbb{R}^n = \operatorname{Int}(S) \sqcup \partial S \sqcup \operatorname{Ext}(S)$

(5) ∂S is closed.

(6) Let $\{\mathcal{O}_i\}_{i=1}^n \subseteq \mathcal{P}(\mathbb{R})$ and define $\mathcal{O} = \prod_{i=1}^n \mathcal{O}_i$. If $\mathcal{O}_i$'s are open (closed), $\mathcal{O}$ is open (closed).

2.2 Limits and Continuity

Recall the notion of limit in $\mathbb{R}$:
Suppose $f : (a,b) \setminus \{c\} \to \mathbb{R}$ is a function. We say $\lim\limits_{x\to c} f$ exists if there is $\alpha \in \mathbb{R}$ such that $\forall\, \varepsilon > 0, \exists\, \delta > 0$ such that $x \in D_\delta(c) \implies |f(x) - \alpha| < \varepsilon$, that is, $f(D_\delta(c)) \subseteq B_\varepsilon(\alpha)$; in such a case we say that $\lim\limits_{x\to c} f = \alpha$.
We now extend this to $\mathbb{R}^n$.

Definition 2.2.1 ▶ Limits in $\mathbb{R}^n$

Let $S \subseteq \mathbb{R}^n, a \in \{\text{Limit points of } S\}$ and, $f : S \setminus \{a\} \to \mathbb{R}^m$. We say that $\lim\limits_{x\to a} f = b$ if for all $\varepsilon > 0$, there is $\delta > 0$ such that $f(x) \in B_\varepsilon(b)$ for all $x \in D_\delta(a) \cap S$.
In other words, $\lim\limits_{x\to a} f = b$ if for all $\varepsilon > 0$, there is $\delta > 0$ such that

$$\|f(x) - b\| < \varepsilon \quad \forall\, x \in S, 0 < \|x - a\| < \delta$$

Definition 2.2.2 ▶ Continuity in $\mathbb{R}^n$

Let $S \subseteq \mathbb{R}^n, a \in S$ and, $f : S \to \mathbb{R}^m$. We say that f is continuous at a if for all $\varepsilon > 0$, there is $\delta > 0$ such that $f(x) \in B_\varepsilon(f(a))$ for all $x \in B_\delta(a) \cap S$, that is, $\|f(x) - f(a)\| < \varepsilon$ for all $x \in S$ with $\|x - a\| < \delta$.

Note: Any function defined on S is vacuously continuous at an isolated point a by our definition.

Lecture 3

3.1 Introduction

We denote the set of limit points of $S \subseteq \mathbb{R}^n$ by S'. Let $f : S \to \mathbb{R}^m$ and $a \in S$. f is continuous at a iff for all $\varepsilon > 0$, there is $\delta > 0$ such that

$$\|f(x) - f(a)\| < \varepsilon \quad \forall \|x - a\| < \delta, x \in S$$
$$\Longleftrightarrow$$
$$f(B_\delta(a) \cap S) \subseteq B_\varepsilon(f(a))$$

For $a \in S'$, we then get that f is continuous at a iff

$$\lim_{\|x-a\| \to 0} \|f(x) - f(a)\| = 0 \iff \lim_{\|h\| \to 0} \|f(a+h) - f(a)\| = 0$$

But $\|h\| \to 0 \iff h \to 0$, and so we have f is continuous at $a \in S'$ iff

$$\lim_{h \to 0} \|f(a+h) - f(a)\| = 0$$

The proof of the following theorem is left as an exercise.

Theorem 3.1.1

Let $S \subseteq \mathbb{R}^n, a \in S', b \in \mathbb{R}^m, f : S \to \mathbb{R}^m$. The following are equivalent:

(i) $\lim_{x \to a} f = b$

(ii) $\forall \{x_p\} \subseteq S \setminus \{a\}$ with $x_p \longrightarrow a$, we have $f(x_p) \longrightarrow b$

(iii) $\lim_{x \to a} \|f(x) - b\| = 0$

Note: If $a \in S$, we can take $b = f(a)$ and get analogous results for continuity at a.

3.2 Properties of Continuous functions

Definition 3.2.1 ▶ Continuity on sets

Let $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}^m$. f is continuous on S if it is continuous at all $a \in S$.

Note: It is convenient to take S to be open, as f is continuous at any isolated points of S vacuously.

Theorem 3.2.1

Let $S \subseteq \mathbb{R}^n, f : S \to \mathbb{R}^m$. The following are equivalent:

(1) f is continuous on S.

(2) $\forall \{x_p\} \subseteq S$ with $x_p \longrightarrow a \in S$, we have $f(x_p) \longrightarrow f(a)$

(3) (Assuming S is open) $f^{-1}(\mathcal{O})$ is open for all $\mathcal{O} \subseteq \mathbb{R}^m$ open.

(4) (Assuming S is open) $f^{-1}(C)$ is closed for all $C \subseteq \mathbb{R}^m$ closed.

Proof. We have the following cases.

- $(3) \iff (4)$
  This is true as if $g : X \to Y$ then for all $A \subseteq Y$, $g^{-1}(Y \setminus A) = X \setminus g^{-1}(A)$.

- $(1) \iff (2)$
  True by 3.2.1

- $(1) \implies (3)$
  Let $\mathcal{O} \subseteq \mathbb{R}^m$ be open, and without loss of generality, $f^{-1}(\mathcal{O}) \neq \phi$.

  Let $a \in f^{-1}(\mathcal{O})$ so that $f(a) \in \mathcal{O}$. Hence, there is $r > 0$ such that $B_r(f(a)) \subseteq \mathcal{O}$. By continuity, there is $\delta > 0$ such that

  $$\begin{aligned} &f(B_\delta(a)) \subseteq B_r(f(a)) \subseteq \mathcal{O} \\ \implies\ &B_\delta(a) \subseteq f^{-1}(\mathcal{O}) \end{aligned}$$

  Hence, $\mathcal{O}$ is open.

- $(3) \implies (1)$
  Fix $a \in S$ and let $\varepsilon > 0$. As $B_\varepsilon(f(a))$ is open in $\mathbb{R}^m$, we have $f^{-1}(B_\varepsilon(f(a)))$ is open in $\mathbb{R}^n$. But $a \in f^{-1}(B_\varepsilon(f(a)))$, and so, there is $\delta > 0$ such that

  $$\begin{aligned} &B_\delta(a) \subseteq f^{-1}(B_\varepsilon(f(a))) \\ \implies\ &f(B_\delta(a)) \subseteq B_\varepsilon(f(a)) \end{aligned}$$

  Hence, f is continuous at a for all $a \in S$.

$\square$

This theorem gives us a huge simplification. Recall that $x \longrightarrow y$ iff $\Pi_i(x) \longrightarrow \Pi_i(y)$ for all i. Now consider some $\{x_p\}$ with $x_p \longrightarrow a$. We have

$$f(x_p) \longrightarrow f(a) \iff \Pi_i(f(x_p)) \longrightarrow \Pi_i(f(a))\,\forall\, i$$

That is, f is continuous iff f is continuous coordinate wise! Hence, for talking about continuity, it is enough to discuss real valued functions rather than $f : \mathbb{R}^n \to \mathbb{R}^m$ for arbitrary m.

3.3 Examples

Example 3.3.1

Consider the function,

$$f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$$
$$f(x,y) = \frac{2xy}{x^2+y^2}$$

Consider the line L_1 defined by $y = 0$ and approach $(0,0)$ from the right $(x \longrightarrow 0^+)$. We have,

$$f\big|_{L_1} \equiv 0 \implies \lim_{(x,y)\to 0 \text{ along } L_1} f = \lim_{n\to\infty} f\left(0, \frac{1}{n}\right) = 0$$

Now consider the line L_2 defined by $x = y$. We have,

$$f\big|_{L_2} \equiv 1 \implies \lim_{(x,y)\to 0 \text{ along } L_2} f = \lim_{n\to\infty} f\left(\frac{1}{n}, \frac{1}{n}\right) = 1$$

Hence, $\lim\limits_{(x,y)\to(0,0)} f$ does not exist.

Note: The approach in the above example is often useful for showing non-existence of limits, or that a function is not continuous.

Example 3.3.2

We wish to compute $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^3}{x^2+y^2}$. We have, for all $(x,y) \neq (0,0)$,

$$\left|\frac{x^3}{x^2+y^2}\right| \leq \left|\frac{x^3}{x^2}\right| = |x| \leq \|(x,y)\|$$

and because $\|(x,y)\|$ goes to 0 as $(x,y) \longrightarrow (0,0)$, we get

$$\lim_{(x,y)\to(0,0)} \frac{x^3}{x^2+y^2} = 0$$

Hence, the function

$$f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2}, \ (x,y) \neq (0,0) \\ 0, \ (x,y) = (0,0) \end{cases}$$

is continuous at $(0,0)$.

Example 3.3.3

Consider $\lim\limits_{(x,y)\to(0,0)} \dfrac{\sin\left(x^2+y^2\right)}{x^2+y^2}$. We have $(x^2+y^2) \longrightarrow 0$ as $(x,y) \to (0,0)$, and hence, we get

$$\lim_{(x,y)\to(0,0)} \frac{\sin\left(x^2+y^2\right)}{x^2+y^2} = \lim_{t\to 0} \frac{\sin t}{t} = 1$$

Exercise. Let $S \subseteq \mathbb{R}^n, a \in S'$ and $f, g : S \to \mathbb{R}$. Suppose $\lim\limits_{x \to a} f = \alpha$, $\lim\limits_{x \to a} g = \beta$ exist. Show that:

(i) $\lim\limits_{x \to a} (rf + g) = r\alpha + \beta, \forall\, r \in \mathbb{R}$

(ii) $\lim_{x \to a} fg = \alpha\beta$

(iii) If $\beta \neq 0$, $\lim\limits_{x \to a} \dfrac{f}{g} = \dfrac{\alpha}{\beta}$

(iv) If $f \leq h \leq g$ for some $h : S \to \mathbb{R}$ and $\alpha = \beta$, then $\lim\limits_{x \to a} h = \alpha$

Note: Similar results hold for continuity as well, using which we get the next examples.

Example 3.3.4 (Some classes of continuous functions)

(1) The projection maps $\Pi_i : \mathbb{R}^n \to \mathbb{R}$ are continuous.

(2) $x_i \in \mathbb{R}[x_1, \ldots, x_n]$ is continuous for all i.

(3) $x_i^2 \in \mathbb{R}[x_1, \ldots, x_n]$ is continuous for all i.

(4) All monomials in $\mathbb{R}[x_1, \ldots, x_n]$ are continuous.

(5) Any $p \in \mathbb{R}[x_1, \ldots, x_n]$ is continuous.

(6) $\frac{p}{q}$ is continuous at $a \in \mathbb{R}^n$, where $p, q \in \mathbb{R}[x_1, \ldots, x_n]$ and $q(a) \neq 0$.

Lecture 4

4.1 More examples of Continuous maps

Example 4.1.1

Consider the function

$$f(x,y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}},\ (x,y) \neq (0,0) \\ 0,\ (x,y) = (0,0) \end{cases}$$

f is clearly continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$ so we need only check the limit at $(0,0)$. We have,

$$0 \leq \left| \frac{xy}{\sqrt{x^2+y^2}} \right| \leq \frac{1}{2} \frac{x^2+y^2}{\sqrt{x^2+y^2}} = \frac{1}{2} \|(x,y)\|$$

By the squeeze theorem, we get $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$, and hence, f is continuous on $\mathbb{R}^2$.

Exercise. Show that any linear map is continuous. (Hint: Use that the norm is continuous.)

Example 4.1.2

Let $D = \{(x,y) \in \mathbb{R}^2 \mid y \neq 0\} = \Pi_2^{-1}(\mathbb{R} \setminus \{0\})$. Consider the function

$$f : D \to \mathbb{R}$$
$$f(x,y) = x \sin \frac{1}{y}$$

Clearly f is continuous on D. We also have,

$$0 \leq \left| x \sin \frac{1}{y} \right| \leq \|(x,y)\|$$

and hence, by the Squeeze theorem, $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$
So, we can extend f to $(0,0)$ continuously by defining it to be 0.

4.2 Uniform Continuity

Definition 4.2.1 ▶ Uniform Continuity

Let $f : \mathcal{O}_n \to \mathbb{R}^m$, where $\mathcal{O}_n$ is open in $\mathbb{R}^n$. f is uniformly continuous if for all $\varepsilon > 0$, there is $\delta > 0$ such that

$$f(x) \in B_\varepsilon(f(a)) \quad \forall\, x \in B_\delta(a) \cap \mathcal{O}_n, \quad \forall\, a \in \mathcal{O}_n$$

$$\Updownarrow$$

$$\|f(x) - f(a)\| < \varepsilon \quad \forall\, \|x - a\| < \delta,\ a, x \in \mathcal{O}_n$$

Exercise. Show that uniform continuity implies continuity.

4.3 Derivatives

We will use the following notation for the sake of brevity:

(1) $\mathcal{O}_n$ denotes an open set in $\mathbb{R}^n$, and we omit the n if it is clear from context.

(2) A function $f : \mathbb{R}^n \to \mathbb{R}^m$ has components $f = (f_1, f_2, \ldots, f_m)$

Recall the notion of derivative in $\mathbb{R}$:
Let $f : \mathcal{O}_1 \to \mathbb{R}$ be a function and $a \in \mathcal{O}_1$. Then f is differentiable at a if there is a real number $\alpha(= f'(a))$ such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \alpha \tag{1}$$

This clearly cannot be carried over verbatim to functions of several variables; we can't divide by a vector! To get a reasonable definition, we note that (1) is equivalent to,

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

Now, $h \mapsto \alpha h$ is a linear map and we already know $\{\text{linear maps on } \mathbb{R}\} \longleftrightarrow \mathbb{R}$. So the derivative $f'(a)$ is really a linear map! This leads to the following definition.

Definition 4.3.1 ▶ Derivative in several variables

Let $f : \mathcal{O}_n \to \mathbb{R}^m$ and $a \in \mathcal{O}_n$. f is differentiable at a if there is a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Lh}{\|h\|} = 0$$

$$\Updownarrow$$

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - Lh\|}{\|h\|} = 0$$

$$\Updownarrow$$

$$\lim_{x \to a} \frac{\|f(x) - f(a) - L(x-a)\|}{\|x - a\|} = 0$$

If f is differentiable at $a \in \mathcal{O}_n$, its derivative is denoted as $Df(a)$. We say f is differentiable on $\mathcal{O}_n$ if it is differentiable at all $a \in \mathcal{O}_n$.

Theorem 4.3.1

Let $f : \mathcal{O}_n \to \mathbb{R}^m$ be differentiable at $a \in \mathcal{O}_n$. The derivative $Df(a)$ is unique.

Proof. Let $L = Df(a)$ and L_1 be any linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - L_1 h\|}{\|h\|} = 0$$

Suppose $L_1 \neq L$. Then, there is $h_0 \in \mathbb{R}^n$ such that $\|h_0\| = 1$ and $Lh_0 \neq L_1 h_0$. Consider the map $h : \mathbb{R} \to \mathbb{R}^n$ given by $h(t) = th_0$. Then, by the triangle inequality,

$$\frac{\|L(h(t)) - L_1(h(t))\|}{|t|} \leq \frac{\|f(a+h) - f(a) - L(h(t))\|}{\|h(t)\|} + \frac{\|f(a+h) - f(a) - L_1(h(t))\|}{\|h(t)\|}$$

Taking the limit as $t \to 0$, both terms on the right go to 0 by definition, and hence

$$\lim_{t\to 0} \frac{\|L(h(t)) - L_1(h(t))\|}{|t|} = 0 \implies \lim_{t\to 0} \frac{|t|\|Lh_0 - L_1 h_0\|}{|t|} = 0 \implies Lh_0 = L_1 h_1$$

which clearly contradicts the assumption. So, the derivative $Df(a)$ is unique. ☐

4.4 Examples

Example 4.4.1

Consider $f : \mathcal{O}_n \to \mathbb{R}^m$ defined by $f(x) = c$. For any $a \in \mathcal{O}_n$,

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - Oh}{\|h\|} = 0$$

where O denotes the zero linear map. Hence, f is differentiable on $\mathcal{O}_n$ and $Df(a) = O$ for any $a \in \mathcal{O}_n$.

Example 4.4.2

Consider a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$. For all $a \in \mathbb{R}^n$,

$$\lim_{h\to 0} \frac{L(a+h) - La - Lh}{\|h\|} = 0$$

Hence, L is differentiable everywhere and $DL(a) = L$ for all $a \in \mathbb{R}^n$. This is as expected, as the best linear approximation of a linear map is itself.

At this point, we are faced with the problem of actually computing derivatives of non-trivial maps. A priori, it is not even clear if functions that are made of various differentiable functions of one variable, say $f(x, y, z) = (x^2 e^y z, y^3 \sin(xy) \cos z)$, are differentiable! We will perform a series of reductions that will answer such basic questions about differentiability of functions and even provide techniques to compute the derivatives.

Lecture 5

We have figured out that the reasonable definition of the derivative of a map $f : \mathcal{O}_n \to \mathbb{R}^m$ at a point $a \in \mathcal{O}_n$ is $Df(a) = L$ where

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - Lh}{\|h\|} = 0$$

We now proceed to perform a series of reductions that will actually answer the question of differentiability of functions and also provide techniques to compute the derivative.

5.1 Matrix representation of the derivative

Theorem 5.1.1

Let $f : \mathcal{O}_n \to \mathbb{R}^m$ and $a \in \mathcal{O}_n$. Then, f is differentiable at a iff f_i is differentiable at a for all i. In that case, we have

$$Df(a) = \begin{pmatrix} Df_1(a) \\ Df_2(a) \\ \vdots \\ Df_m(a) \end{pmatrix}$$

Proof. Assume that f is differentiable at a and let $L = Df(a)$. For all $i = 1, 2, \ldots, m$ we set $L_i = \Pi_i \circ L$. Further, let $\tilde{f}_i(h) = f_i(a+h) - f_i(a) - L_i h$, so that

$$f(a+h) - f(a) - Lh = (\tilde{f}_1(h), \tilde{f}_2(h), \ldots, \tilde{f}_m(h))$$

But then, $\left|\tilde{f}_i(h)\right| \leq \|f(a+h) - f(a) - Lh\|$ which proves that f_i is differentiable at a and has derivative $Df_i(a) = L_i$, for all i.

Assume that f_i is differentiable at a for all i and define

$$L = \begin{pmatrix} Df_1(a) \\ Df_2(a) \\ \vdots \\ Df_m(a) \end{pmatrix}$$

Let $\Pi_i(f(a+h) - f(a) - Lh) = f_i(a+h) - f_i(a) - Df_i(a)h = \tilde{f}_i(h)$, so that

$$\lim_{h\to 0} \frac{\tilde{f}_i(h)}{\|h\|} = 0 \quad \forall\, i$$
$$\implies \lim_{h\to 0} \frac{f(a+h) - f(a) - Lh}{\|h\|} = 0$$

which proves that f is differentiable at a and has derivative $Df(a) = L$. ☐

Note the following:

- Because of this theorem, it is enough to study differentiability of maps $f : \mathcal{O}_n \to \mathbb{R}$.
- Let $\gamma : \mathcal{O}_1 \to \mathbb{R}^n$. Then, γ is differentiable at $a \in \mathcal{O}_1$ iff γ_i is differentiable at a for all i and in that case,

  $$D\gamma(a) = \gamma'(a) = \begin{pmatrix} \gamma_1'(a) \\ \gamma_2'(a) \\ \vdots \\ \gamma_n'(a) \end{pmatrix}$$

  This is the notion of velocity vector of a curve in $\mathbb{R}^n$ from elementary calculus.

5.2 Further properties of differentiable functions

Theorem 5.2.1 (Differentiability implies continuity)

Let $f : \mathcal{O}_n \to \mathbb{R}^m$. If f is differentiable at $a \in \mathcal{O}_n$, f is continuous at a.

Proof. Let f be differentiable at a. Then, for all $x \in \mathcal{O}_n$,

$$0 \leq \|f(x) - f(a)\| \leq \|f(x) - f(a) - Df(a)(x-a)\| + \|Df(a)(x-a)\|$$

Taking the limit as $x \to a$, the first term on the right goes to 0 by the definition of $Df(a)$ and the second term goes to zero as any linear map is continuous. Hence, by the squeeze theorem,

$$\lim_{x \to a} \|f(x) - f(a)\| = 0$$

which proves that f is continuous at a. □

Theorem 5.2.2 (Chain rule)

Consider maps f, g such that

$$\begin{tikzcd} \mathcal{O}_n \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & \mathcal{O}_m \arrow[r, "g"] & \mathbb{R}^p \end{tikzcd}$$

Assume that f is differentiable at $a \in \mathcal{O}_n$ and g is differentiable at $f(a) \in \mathcal{O}_m$. Then $g \circ f$ is differentiable at a and further

$$\underbrace{D(g\circ f)(a)}_{\mathbb{R}^n \to \mathbb{R}^p} = \underbrace{Dg(f(a))}_{\mathbb{R}^m \to \mathbb{R}^p} \cdot \underbrace{Df(a)}_{\mathbb{R}^n \to \mathbb{R}^m}$$

Proof. Let $A = Df(a)$ and $B = Dg(b)$ where $b = f(a)$. For x, y in sufficiently small neighbourhoods of a, b respectively, we consider the maps

$$\begin{aligned} r_f(x) &= f(x) - f(a) - A(x-a) \\ r_g(y) &= g(y) - g(b) - B(y-b) \\ r(x) &= g(f(x)) - g(b) - BA(x-a) \end{aligned}$$

By definition of the derivative,

$$\lim_{x \to a} \frac{r_f(x)}{\|x-a\|} = 0 \qquad \lim_{y \to b} \frac{r_g(y)}{\|y-b\|} = 0$$

We wish to prove that

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

We have,

$$\begin{aligned} r(x) &= g(f(x)) - g(b) - BA(x-a) \\ &= g(f(x)) - g(b) + B(r_f(x) - f(x) + f(a)) \\ &= Br_f(x) + g(f(x)) - g(b) - B(f(x) - f(a)) \\ \implies r(x) &= Br_f(x) + r_g(f(x)) \end{aligned}$$

Now,

$$\lim_{x \to a} \frac{Br_f(x)}{\|x-a\|} = B\left(\lim_{x \to a} \frac{Br_f(x)}{\|x-a\|}\right) = 0$$

The other term requires some more analysis. Fix $\varepsilon > 0$. There is $\delta > 0$ such that $\|r_g(y)\| < \varepsilon\|y-b\|$ for all y with $0 < \|y-b\| < \delta$. By continuity of f at a, there is $\delta_1 > 0$ such that $\|f(x) - f(a)\| < \delta$ for all x with $0 < \|x-a\| < \delta_1$. Hence,

$$0 < \|x-a\| < \delta_1 \implies \|r_g(f(x))\| < \varepsilon\|f(x) - f(a)\|$$

and so

$$\lim_{x \to a} \frac{r_g(f(x))}{\|f(x) - f(a)\|} = 0$$

We now note,

$$\|f(x) - f(a)\| = \|r_f(x) + A(x-a)\| \leq \|r_f(x)\| + M_A\|x-a\|$$

where we get $M_A > 0$ by Theorem 1.4.3. Hence, we finally have

$$\lim_{x \to a} \frac{r(x)}{\|x-a\|} = \lim_{x \to a} \frac{r_g(f(x))}{\|f(x) - f(a)\|} \frac{\|f(x) - f(a)\|}{\|x-a\|} = 0$$

which completes the proof. ☐

Lecture 6

6.1 Partial Derivatives

Now, we discuss the notion of partial derivatives. This tries to treat differentiation of functions with multiple arguments in the 1-variable setting. This construction ends up being an indispensable tool in the computation of the Total derivative in the standard basis.

Consider a function $f : \mathcal{O}_n \to \mathbb{R}$ and a point $a \in \mathcal{O}_n$, and fix $1 \leq i \leq n$. Now, we define functions $\eta_i : (-\epsilon, \epsilon) \to \mathbb{R}$ such that $\eta(t) = f(a + te_i)$ for all $t \in (-\epsilon, \epsilon)$.

Definition 6.1.1 ▶ Partial Derivatives

For the given function $f : \mathcal{O}_n \to \mathbb{R}$, the partial derivatives of f with respect to the co-ordinate x_i is given by:

$$f_{x_i}(a) \equiv \frac{\partial f}{\partial x_i}(a) := \frac{d\eta_i}{dt}(a) = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t} \text{ if it exists.}$$

Considering the maps $h_i : \mathbb{R} \to \mathbb{R}^n$ such that $h_i(t) = a + te_i$, we have $\eta_i = f \circ h_i$. Thus, by Chain Rule, if f is differentiable, all its partial derivatives exist.

6.2 Geometric Meaning

The partial derivative measures the change of a function at a point due to a particular variable, keeping all others constant. The geometry of partial derivatives is best visualized in 3 dimensions. Taking $f : \mathcal{O}_2 \to \mathbb{R}$, we consider the surface $\mathcal{S} \subset \mathbb{R}^3$ defined by $z = f(x, y)$. Let $P_0 = (x_0, y_0, f(x_0, y_0))$ be a point on $\mathcal{S}$. Then the value $f_x(x_0, y_0)$ (if it exists) is the slope of the tangent to $\mathcal{S}$ at $(x_0, y_0, f(x_0, y_0))$ pointing in the positive x direction.

Another interpretation is to consider the plane $\mathcal{P} = \{(x, y, z) \mid y = y_0\}$, and the curve $\mathcal{C}$ on surface $\mathcal{S}$ given by $\mathcal{C} = \mathcal{S} \cap \mathcal{P}$. Then, $f_x(x_0, y_0)$ is the slope of the tangent to the curve $\mathcal{C}$, in the direction of increasing x co-ordinate.

For $f : \mathcal{O}_n \to \mathbb{R}$ with $n > 2$, although it becomes harder to visualize, the interpretation remains the same.

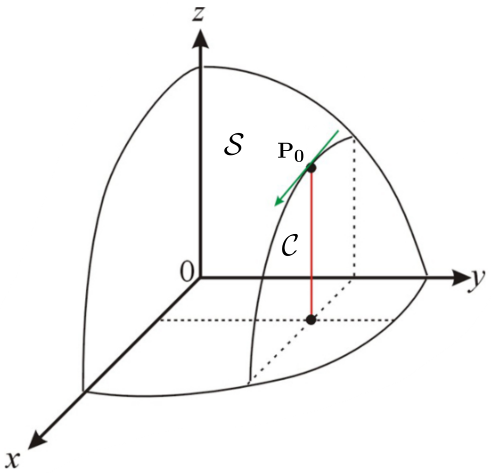


Figure 6.1: Geometric Meaning

6.3 Examples

We now lay out some interesting and instructive examples, which would illustrate some general results about partial derivatives.

Example 6.3.1

Consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^3 + y^4 + \sin(xy)$. Then,

$$\frac{\partial f}{\partial x} = 3x^2 + y \cos xy \text{ and } \frac{\partial f}{\partial y} = 4x^3 + x \cos xy$$

Example 6.3.2

We know that Differentiable functions $f : \mathcal{O}_n \to \mathbb{R}$ are continuous. However, even existence of all partial derivatives is too weak to ensure continuity.
Consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

Evidently, all the partial derivatives exist, but as shown previously, this function is discontinuous at $(0, 0)$.

Definition 6.3.1

A function $f : \mathcal{O} \to \mathbb{R}$ is called $C^k(\mathcal{O})$ if all the k^{th} order partial derivatives exist and are continuous.

6.4 Higher Order Partial Derivatives

Assume that partial derivatives of $f : \mathcal{O}_n \to \mathbb{R}$ exist in a neighbourhood of $a \in \mathcal{O}_n$. Then we can talk about the partial derivatives of $\frac{\partial f}{\partial x_i} : \mathcal{O}_n \to \mathbb{R}$ at a. We denote:

$$f_{x_i x_j}(a) \equiv \frac{\partial f_{x_i}}{\partial x_j} := \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{1}{h} \left(\frac{\partial f}{\partial x_i}(a + he_j) - \frac{\partial f}{\partial x_i}(a) \right)$$

We can define higher order partial derivatives similarly. Please note that the order of differentiation matters in general. For starters, of $f_{x_i x_j}$ and $f_{x_j x_i}$, one may exist while the other may not. Also, even if both exist, they may not necessarily be equal over the entire domain. We leave it to the reader to find an example of the former case, and provide an example for the latter.

Example 6.4.1

Consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

In this case, it is easy to show (Exercise!) that both f_{xy} and f_{yx} exist, but

$$f_{xy}(0, 0) = 1 \neq -1 = f_{yx}(0, 0)$$

Example 6.4.2

However, in many *well-behaved* cases, we will find $f_{xy} = f_{yx}$. For instance, consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = \sin(x) + e^y + xy$. Show that $f_{xx} = f_{yy} = 1$ over $\mathbb{R}^2$.

Often, the dependence of the partial derivative on the order of differentiation is the exception rather than the rule. We now develop a sufficient condition for $f_{xy} = f_{yx}$ to hold.

6.5 Clairaut's Theorem

Theorem 6.5.1 (Clairaut)

Let $(a, b) \in \mathcal{O}_2$ and $f : \mathcal{O}_2 \to \mathbb{R}$. Suppose f_x, f_y, f_{xy}, and f_{yx} all exist on $\mathcal{O}_2$. If f_{xy} and f_{yx} are continuous at (a, b), then $f_{xy}(a, b) = f_{yx}(a, b)$.

Proof. Without loss of generality, we take $(a, b) = (0, 0) \in \mathcal{O}^2$. As $\mathcal{O}^2$ is open, we choose a box $[h, 0] \times [0, k] \subset \mathcal{O}^2$. Now, we have

$$\begin{aligned} f_{xy}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \lim_{k \to 0} \frac{f_x(x, y + k) - f_x(x, y)}{k} \\ &= \lim_{k \to 0} \lim_{h \to 0} \frac{1}{hk}(f(x + h, y + k) - f(x, y + k) - f(x + h, y) + f(x, y)) \end{aligned}$$

We define

$$F(h, k) = \frac{1}{hk}(f(h, k) - f(0, k) - f(h, 0) + f(0, 0))$$

Thus, by the above result, we have

$$\begin{aligned} f_{xy}(x, y) &= \lim_{k \to 0} \lim_{h \to 0} F(h, k), \text{ and similarly} \\ f_{yx}(x, y) &= \lim_{h \to 0} \lim_{k \to 0} F(h, k) \end{aligned}$$

Now we proceed for the proof in earnest. Define $f_1(x) = f(x, k) - f(x, 0)$, which is continuous on $[0, h]$ and differentiable on $(0, h)$. Thus, by Lagrange's Mean Value Theorem, there exists $c_1 \in (0, h)$ (depending upon both h and k), such that, $h(f_1'(c_1)) = (f_1(h) - f_1(0))$, i.e.

$$\begin{aligned} \therefore\ f_x(c_1, k) - f_x(c_1, 0) &= \frac{1}{h}(f(h, k) - f(0, k) - f(h, 0) + f(0, 0)) = kF(h, k) \\ \implies F(h, k) &= \frac{1}{k}(f_x(c_1, k) - f_x(c_1, 0)) \end{aligned}$$

Again, define $f_2(y) = f_x(c_1, y)$, which again satisfies all conditions for the Mean Value Theorem. Thus, there exists $c_2 \in (0, k)$ such that $k(f_2'(c_2)) = (f_2(k) - f_1(0))$, which gives, $F(h, k) = f_{xy}(c_1, c_2)$. Repeating this entire construction, we can find $(c_1', c_2') \in [0, h] \times [0, k]$ such that $F(h, k) = f_{yx}(c_1', c_2')$. Thus, $f_{xy}(c_1, c_2) = f_{yx}(c_1', c_2')$. But, $0 < c_1, c_1' < h$ and $0 < c_2, c_2' < k$. Thus, as (h, k) can be made arbitrarily small, taking $(h, k) \to 0$, we have

$$\lim_{(c_1, c_2) \to 0} f_{xy}(c_1, c_2) = \lim_{(c_1', c_2') \to 0} f_{yx}(c_1', c_2')$$

By the continuity of f_{xy} and f_{yx}, we have $\boxed{f_{xy}(0, 0) = f_{yx}(0, 0)}$.

□

In particular, $f_{xy} = f_{yx}$ for C^2 functions over a given domain. In the next lecture, we sharpen this result slightly, and relate the partial derivatives to the total derivative.

Lecture 7

7.1 Schwarz Theorem

In the previous lecture, we discussed the notion of partial derivatives. In general, the partial derivatives depend on the order of differentiation. However, using Clairaut's Theorem, we found a necessary when $f_{xy} = f_{yx}$. Now, we conclude that discussion by the following result.

Theorem 7.1.1 (Schwarz)

Let $(a, b) \in \mathcal{O}_2$ and $f : \mathcal{O}_2 \to \mathbb{R}$. Suppose f_x, f_y, and f_{xy} exist on $\mathcal{O}_2$. If f_{xy} is continuous at (a, b), then f_{yx} exists in a neighbourbood of (a, b) and $f_{xy}(a, b) = f_{yx}(a, b)$.

Proof. Just as before, we take $(a, b) = (0, 0)$. From the proof of Clairaut's Theorem, we have $F(h, k) = f_{xy}(c_1, c_2)$ for some $0 < c_1 < h$ and $0 < c_2 < k$. By continuity of f_{xy} at $(0, 0)$, for any $\epsilon > 0$ there exists $h_\epsilon, k_\epsilon > 0$, such that

$$| f_{xy}(u, v) - f_{xy}(0, 0) | < \epsilon \;\; \forall\, (u, v) \in [0, h_\epsilon] \times [0, k_\epsilon]$$

But then $| F(h, k) - f_{xy}(0, 0) | < \epsilon \;\; \forall\, (u, v) \in [0, h_\epsilon] \times [0, k_\epsilon]$, that is, F is continuous at $(0, 0)$ with limit $f_{xy}(0, 0)$ at $(0, 0)$.
As $f_y(h, 0)$ exists for h sufficiently small, $\lim_{k\to 0} F(h, k)$ exists for h sufficiently small. Thus, by continuity of F at $(0, 0)$,

$$\lim_{h\to 0} \lim_{k\to 0} F(h, k) \text{ exists and is equal to } \lim_{(h,k)\to(0,0)} F(h, k)$$

Thus, $f_{yx}(0, 0)$ exists and is equal to $f_{xy}(0, 0)$ ☐

This is a slightly more useful version of Clairaut's Theorem. However, in many applications (say, partial differential equations), we work with C^2 or even C^∞ functions, in which case both of these hold automatically.

Exercise: Formulate and prove a similar result for higher order derivatives. In particular, provide a sufficient condition for $f : \mathcal{O}_n \to \mathbb{R}$ so that

$$\frac{\partial^n f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_m}} = \frac{\partial^n f}{\partial x_{\sigma(i_1)} \partial x_{\sigma(i_2)} \ldots \partial x_{\sigma(i_m)}}$$

over $\mathcal{O}_n$ for any permutation σ of the elements $\{i_1, i_2, \ldots, i_n\}$.

7.2 Partial and Total Derivatives

We will now see that the partial derivatives provide an effective way of proving the existence and computing the total derivatives of a function $f : \mathcal{O}_n \to \mathbb{R}^m$. In this lecture and the next, we will develop the relations between partial and total derivatives by a series of results.

Definition 7.2.1 ▶ Jacobian Matrix

For a function $f = (f_1, f_2, \ldots, f_m) : \mathcal{O}_n \to \mathbb{R}^m$, if all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ at $a \in \mathcal{O}_n$, we define the Jacobian of the function at a by the $m \times n$ matrix,

$$J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{m\times n}$$

Theorem 7.2.1

Consider a function $f = (f_1, f_2, \ldots, f_m) : \mathcal{O}_n \to \mathbb{R}^m$ differentiable at $a \in \mathcal{O}_n$. Then all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist at a. In particular, for f differentiable at a, we have,

$$(Df)(a) = J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{m\times n}$$

Proof. Without loss of generality, we take $m = 1$, and let $a = (a_1, a_2, \ldots, a_n)$. Fix an arbitrary index $i \in \{1, 2, \ldots, n\}$. We define $\eta_i : [a_i - \epsilon, a_i + \epsilon] \to \mathbb{R}^n$, defined by

$$\eta_i(t) = (a_1, \ldots, a_{i-1}, t, a_{t+1}, \ldots, a_n) = a + (t - a_i)e_i$$

As $\mathcal{O}_n$ is open and η_i is continuous, we can find ϵ small such that $f([[a_i - \epsilon, a_i + \epsilon]]) \subseteq \mathcal{O}_n \subseteq \mathbb{R}^n$. Evidently, η_i is differentiable and $(D\eta_i) = [0, \ldots, 1, \ldots, 0]^t = e_i^t$ over $[a_i - \epsilon, a_i + \epsilon]$. Now, by the definition of partial derivatives, $D(f \circ \eta_i)(a_i) = f_{x_i}(a)$.
Again, by chain rule, as f is differentiable at a, $D(f \circ \eta_i)(a_i) = f_{x_i}(a)$ exists, and

$$\begin{aligned} D(f \circ \eta_i)(a_i) &= Df(\eta_i(a_i)) \cdot D\eta_i(a_i) \\ \implies f_{x_i}(a) &= Df(a) \cdot e_i^t = [Df(a)]_i \end{aligned}$$

As the index i was arbitrary to begin with, this completes the proof. ☐

This theorem proves that differentiability of a function implies the existence of its partial derivatives, and gives the form of the derivative in the standard basis. But it is often quite elaborate and laborious to prove that a function is differentiable, whereas computation of the partial derivatives is much more straightforward. In the next lecture, we formulate a sufficient condition for differentiability of a function based on its partial derivatives.

Lecture 8

8.1 A kind of converse of Theorem 7.2.1

As we have seen in previous lecture, the differentiability of a function gives an explicit expression for derivative with the existence of partials. In this lecture, we will prove a sufficient condition for differentiability based on its partials, which will be our final reduction for derivatives.

Theorem 8.1.1 (Final Reduction)

Let $f : \mathcal{O}_n \to \mathbb{R}^m$ and $a \in \mathcal{O}_n$. Suppose, all partial derivatives $\frac{\partial f_i}{\partial x_j}$ exists on $\mathcal{O}_n$ and continuous at $a \in \mathcal{O}_n$. Then,

$$Df(a) = J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{m\times n}$$

Proof. Without loss of generality, we take $a = (0, \dots, 0) \in \mathcal{O}_n$ and $m = 1$.

Let's do some back calculation: We already "know",

$$L = J_f(a) = \begin{pmatrix} f_{x_1}(0) & \cdots & f_{x_n}(0) \end{pmatrix} \text{ and } \boxed{Lh = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(0)\ \forall\ h \in \mathbb{R}^n}$$

which we use in the following claim,

Claim

$$\frac{1}{\|h\|}\left|f(h) - f(0) - Lh\right| \to 0 \text{ as } h \to 0$$

Proof. Simply calculating,

$$\frac{1}{\|h\|}\left|f(h) - f(0) - Lh\right| = \frac{1}{\|h\|}\left|f(h) - f(0) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(0)\right|$$

For every i, we define, $\hat{h}_i = (h_1, \dots, h_i, \underbrace{0, \dots, 0}_{n-r})$ and $\hat{h}_0 = 0$. Then,

$$\begin{aligned} f(h) - f(0) &= \left(f(\hat{h}_1) - f(\hat{h}_0)\right) + \left(f(\hat{h}_2) - f(\hat{h}_1)\right) + \cdots + \left(f(\hat{h}_n) - f(\hat{h}_{n-1})\right) \\ &= \sum_{i=1}^{n} \left(f(\hat{h}_i) - f(\hat{h}_{i-1})\right) \end{aligned}$$

which implies,

$$\begin{aligned}\left|f(h)-f(0)-\sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(0)\right| &= \left|\sum_{i=1}^{n}\left(f(\hat{h}_i)-f(\hat{h}_{i-1})-h_i\frac{\partial f}{\partial x_i}(0)\right)\right| \\ &= \left|\sum_{i=1}^{n} \underbrace{h_i \frac{\partial f}{\partial x_i}\left(\hat{h}_{i-1}+c_i e_i\right)}_{\text{By MVT, as explained below}} - h_i\frac{\partial f}{\partial x_i}(0)\right|\end{aligned}$$

For a fixed h, we fix $i \in [n]$. And we consider the map, $\eta_i : (h_i-\epsilon, h_i+\epsilon) \to \mathbb{R}$

$$\begin{tikzcd} t \arrow[r, mapsto] \arrow[rr, bend left, "\eta_i"] & \hat{h}_{i-1}+te_i \arrow[r, mapsto] & f(\hat{h}_{i-1}+te_i) \end{tikzcd}$$

defined by, $\eta_i(t) = f(\hat{h}_{i-1}+te_i)$. Clearly, η_i is differentiable on $(0,h_i)$ and continuous on $[0,h_i]$. Then, by Mean Value Theorem,

$$\underbrace{\eta_i(h_i)}_{f(\hat{h}_i)} - \underbrace{\eta_i(0)}_{f(\hat{h}_{i-1})} = \eta_i'(c_i)h_i = f_{x_i}(\hat{h}_{i-1}+c_i e_i)h_i \qquad \text{(for some } c_i \in (0,h_i))$$

Now, observe that, as $h \to 0, \hat{h}_{i-1}+c_i e_i \to 0$ which in turn implies, $f_{x_i}(\hat{h}_{i-1}+c_i e_i) \to f_{x_i}(0)$. Therefore,

$$\begin{aligned}\frac{1}{\|h\|}\left|f(h)-f(0)-Lh\right| &= \frac{1}{\|h\|}\left|\sum_{i=1}^{n} h_i\frac{\partial f}{\partial x_i}\left(\hat{h}_{i-1}+c_i e_i\right)-h_i\frac{\partial f}{\partial x_i}(0)\right| \\ &\le \frac{1}{\|h\|}\sum_{i=1}^{n}|h_i|\left|\frac{\partial f}{\partial x_i}\left(\hat{h}_{i-1}+c_i e_i\right)-\frac{\partial f}{\partial x_i}(0)\right| && \text{(Triangle inequality)} \\ &\le \sum_{i=1}^{n}\left|\frac{\partial f}{\partial x_i}\left(\hat{h}_{i-1}+c_i e_i\right)-\frac{\partial f}{\partial x_i}(0)\right| && (\text{as } \|h\| \ge |h_i| \ \forall\ i) \\ &\longrightarrow 0 \text{ as } h \to 0\end{aligned}$$

□

And, this completes the proof. □

With Theorem 8.1.1, computation of derivative is much easier when we are in favorable situation. Note that,

(i) If f is differentiable at a then $\frac{\partial f_i}{\partial x_j}(a)$ exists for all i, j and $Df(a) = J_f(a)$.

(ii) If $\frac{\partial f_i}{\partial x_j}$ is continuous at a then f is differentiable and $Df(a) = J_f(a)$.

The gap between (i) and (ii) is the continuity of partials, which is removable.

8.2 Examples

We conclude the lecture with some instructive examples.

Example 8.2.1 (Differentiable but discontinuous)

Take,

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right), & (x,y) \neq 0 \\ 0, & (x,y) = 0 \end{cases}$$

Then,

$$\begin{aligned} |f(x,y) - f(0,0)| &= \left|x^2+y^2\right|\left|\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right)\right| \\ &\leq (x^2+y^2) = \|(x,y)\|^2 \end{aligned}$$

implies that f is continuous at $(0,0)$. For all $(x,y) \neq (0,0)$, the partials of f,

$$\begin{aligned} f_x(x,y) &= 2x\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) - \frac{x}{\sqrt{x^2+y^2}}\cos\left(\frac{1}{\sqrt{x^2+y^2}}\right) \\ f_y(x,y) &= 2y\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right) - \frac{y}{\sqrt{x^2+y^2}}\cos\left(\frac{1}{\sqrt{x^2+y^2}}\right) \end{aligned}$$

And at $(0,0)$,

$$\begin{aligned} f_x(0,0) &= \lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t} = 0 \\ f_y(0,0) &= \lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t} = 0 \end{aligned}$$

Also,

$$\frac{1}{\sqrt{h^2+k^2}}\left|f(h,k) - f(0,0) - \begin{pmatrix}0 & 0\end{pmatrix}\begin{pmatrix}h \\ k\end{pmatrix}\right| = \sqrt{h^2+k^2}\left|\sin\left(\frac{1}{\sqrt{h^2+h^2}}\right)\right| \leq \|(h,k)\|$$

which shows that, f is differentiable at $(0,0)$ and $Df(0,0) = \begin{pmatrix}0 & 0\end{pmatrix}$. But, f_x and f_y are not continuous at $(0,0)$!

So, even if a function is differentiable at some point, its partials may still not be continuous there!

Example 8.2.2 (Exercise)

Take,

$$f(x,y) = \begin{cases} x^{\frac{4}{3}}\sin\left(\frac{y}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

- Show that,
  1. f is differentiable on $\mathbb{R}^2$.
  2. f_x and f_y exist and continuous on $\mathcal{O}_2 = \left\{(x,y) \in \mathbb{R}^2 : x \neq 0\right\}$.
  3. f_x is not continuous at $(0,y)$ for all $y \neq 0$.
- Discuss the nature of continuity of f at the origin.

Example 8.2.3

Let, $f : \mathbb{R}^3 \to \mathbb{R}^4$ be defined by,

$$f(x, y) = (x + 2y + 3z, xyz, \cos x, \sin x)$$

Then, the at (x, y, z)

$$J_f(x, y, z) = \begin{pmatrix} 1 & 2 & 3 \\ yz & zx & xy \\ -\sin x & 0 & 0 \\ \cos x & 0 & 0 \end{pmatrix}$$

which has every entry continuous, thus

$$J_f(x, y, z) = Df(x, y, z)$$

Lecture 9

9.1 Directional Derivatives

We now introduce another extension of one-dimensional derivative, called the Directional Derivative.

Definition 9.1.1 ▶ Directional Derivative

Let $u \in \mathbb{R}^n$ (indicates a direction) be a unit vector and $f : \mathcal{O}_n \to \mathbb{R}$ be a scalar-valued function. Take, $a \in \mathcal{O}_n$, the directional derivative of f at a in the direction of u is defined as

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

the rate of change of f at a in the direction of u, provided the limit exists.

Note that $(D_{e_i} f)(a) = \frac{\partial f}{\partial x_i}(a)$, i.e., the partial derivatives are directional derivatives along the standard basis vectors.

Now, consider the following

$$\begin{tikzcd} t \arrow[r, mapsto, "\phi"] \arrow[rr, bend right, "\eta(t)=f(a+tu)"'] & a + tu \arrow[r, mapsto, "f"] & f(a + tu) \end{tikzcd}$$

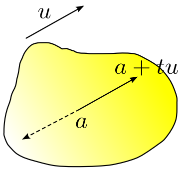


Figure 9.1: $t \mapsto a + tu$

Take, $\eta : (-\varepsilon, \varepsilon) \to \mathbb{R}$ as $\eta(t) := f(a + tu)$. If f is differentiable at a then η is differentiable at 0 (by chain rule) iff f has directional derivative at a along u and,

$$\begin{aligned} \eta'(0) &= (D_u f)(a) \\ &= Df(\phi(0)) \cdot \underbrace{(D\phi)(0)}_{u} \\ &= \begin{pmatrix} \frac{\partial f}{\partial x_i}(a) & \cdots & \frac{\partial f}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \end{aligned}$$

So,

$$\boxed{\eta'(0) = \sum_{i=1}^{n} u_i \frac{\partial f}{\partial x_i}(a) = (D_u f)(a)}$$

9.2 Gradient

Definition 9.2.1 ▶ Gradient

Given a scalar-valued function (a.k.a., scalar field) $f : \mathcal{O}_n \to \mathbb{R}$ and $x \in \mathcal{O}_n$, the Gradient of f at x is defined as

$$\nabla f(x) := \langle f_{x_1}(x), \ldots, f_{x_n}(x) \rangle$$

the dual of the total derivative, i.e., $Df(x)^t$, provided that all the partials f_{x_i} exists at x.

Observe that, for a differentiable function $f : \mathcal{O}_n \to \mathbb{R}$ at a

$$\begin{aligned} \left(D_u f\right)(a) &= \left(\nabla f\right)(a) \cdot u \\ &= \|\left(\nabla f\right)(a)\| \cos\theta_u && \text{(where, } \theta_u \text{ is angle between } \left(\nabla f\right)(a) \text{ and } u\text{)} \end{aligned}$$

which tells us, The steepest slope is achieved when $\theta_u \in \{0, \pi\}$, i.e., when u points along or opposite to the direction of $\left(\nabla f\right)(a)$ that means $\max\left(D_u f\right)(a)$ is attained at $\underbrace{u = \frac{\left(\nabla f\right)(a)}{\|\left(\nabla f\right)(a)\|}}_{\text{provided } \|(\nabla f)(a)\| \neq 0}$ (direction of the steepest slope). Hence, we get the following theorem,

Theorem 9.2.1

Let $f : \mathcal{O}_n \to \mathbb{R}$ be a differentiable function at $a \in \mathcal{O}_n$ and suppose $\left(\nabla f\right)(a) \neq 0$ then the vector $\left(\nabla f\right)(a)$ points in the direction of the greatest increment of f at a with the greatest rate $\|\left(\nabla f\right)(a)\|$

9.3 Examples

Example 9.3.1

Find the directional derivative of $f(x, y, z) = x^2yz$ along $\langle 1, 1, -1\rangle$ at $a = (1, 1, 0)$.
Solution. We have the unit vector $u = \frac{\langle 1,1,-1\rangle}{\|\langle 1,1,-1\rangle\|} = \left\langle \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right\rangle$. So,

$$\begin{aligned} \left(D_u f\right)(a) &= \left(\nabla f\right)(a) \cdot u \\ &= \begin{pmatrix} 2xyz & x^2z & x^2y \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix} \end{aligned}$$

Therefore, $\left(D_u f\right)(1, 1, 0) = -\frac{1}{\sqrt{3}}$ and hence maximum value of $\left(D_u f\right)(1, 1, 0)$ is $\|\left(\nabla f\right)(1, 1, 0)\|$ along the unit vector $\langle 0, 0, 1\rangle$.

Example 9.3.2

Take,

$$f(x, y) = \begin{cases} \frac{x^2y}{x^2+y^2} & (x, y) = (0, 0) \\ 0 & (x, y) \neq (0, 0) \end{cases}$$

then $|f(x, y) - f(0, 0)| = \left|\frac{x^2y}{x^2+y^2}\right| \leq |y| \leq \|(x, y)\|$ implies that f is continuous at $(0, 0)$. Now, fix

$u = \langle u_1, u_2 \rangle$ with $\|u\| = 1$. We get,

$$\begin{aligned}
\left(D_u f\right)(0,0) &= \lim_{t \to 0} \frac{f(tu) - 0}{t} \\
&= \lim_{t \to 0} \frac{1}{t} \cdot \frac{t u_1^2 u_2}{1} \\
&= u_1^2 u_2 \neq 0 && \text{(Because, } u \text{ is an unit vector)}
\end{aligned}$$

If we assume f to be differentiable then $\left(\boldsymbol{\nabla} f\right)(0,0) \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0 \neq u_1^2 u_2$ which is a contradiction!

Example 9.3.2 shows that, the existence of all partial and directional derivatives at a point fails to imply differentiability at that point.

Lecture 10

Example 10.0.3 (Exercise)

Take the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as,

$$f(x, y) = \begin{cases} \frac{x^3 y}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

(i) Prove that, f is continuous at $(0, 0)$.

(ii) Find $(D_u f)(0, 0)\ \forall\ u$.

(iii) Prove that, f is not differentiable.

10.1 Extension of MVT to Several Variables

In Analysis I, you learned about the Mean Value Theorem (MVT) for functions of a single variable. Now, we extend this concept to several variables in the context of multivariable calculus.

Theorem 10.1.1 (Multivariate MVT)

Let $\mathcal{O}_n \subseteq \mathbb{R}^n$ be an open and convex set, and let $f : \mathcal{O}_n \to \mathbb{R}$ be a differentiable function. For any two points $a, b \in \mathcal{O}_n$ define the line segment

$$L_{a,b} := \{tb + (1 - t)a : t \in [0, 1]\}$$

Then, there exists a point $c \in L_{a,b}$ such that,

$$f(b) - f(a) = (\boldsymbol{\nabla} f)(c) \cdot (b - a) = \langle f_{x_1}(c), \ldots, f_{x_n}(c)\rangle \cdot \langle (b_1 - a_1), \ldots, (b_n - a_n)\rangle$$

Proof. We consider the function $\eta : [0, 1] \to \mathcal{O}_n$

$$\begin{tikzcd}
{[0,1]} \arrow[r, "\eta"] \arrow[rr, bend right, "f \circ \eta"'] & \mathcal{O}_n \arrow[r, "f"] & \mathbb{R}
\end{tikzcd}$$

defined by $\eta(t) = (1 - t)a + tb$. This function is differentiable, and its derivative is

$$\eta'(t) = \begin{pmatrix} b_1 - a_1 \\ \vdots \\ b_n - a_n \end{pmatrix}$$

By applying the standard Mean Value Theorem to the composition $f \circ \eta$, there exists $t_0 \in (0,1)$ such that,

$$f(\eta(1)) - f(\eta(0)) = (f \circ \eta)'(t_0) = Df(\eta(t_0)) \cdot D\eta(t_0)$$

Expanding the dot product, we have,

$$f(b) - f(a) = \begin{pmatrix} f_{x_1}(\eta(t_0)) & f_{x_2}(\eta(t_0)) & \cdots & f_{x_n}(\eta(t_0)) \end{pmatrix} \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ \vdots \\ b_n - a_n \end{pmatrix}$$

Simplifying further, we obtain,

$$f(b) - f(a) = \langle f_{x_1}(\eta(t_0)), \cdots, f_{x_n}(\eta(t_0)) \rangle \cdot \langle (b_1 - a_1), \cdots, (b_n - a_n) \rangle$$

This expression can be rewritten as,

$$f(b) - f(a) = (\boldsymbol{\nabla} f)(\eta(t_0)) \cdot (b - a)$$

Hence, there exists $c = \eta(t_0) \in L_{a,b}$ such that $f(b) - f(a) = (\nabla f)(c) \cdot (b - a)$. ☐

10.2 More Partials and Chain Rules

In this section, we further explore the chain rule for differentiable functions of several variables. Consider two functions f and g as following,

$$\mathcal{O}_n \xrightarrow{\quad f \quad} \mathcal{O}_m \xrightarrow{\quad g \quad} \mathbb{R}^p$$

Assuming that f is differentiable at $a \in \mathcal{O}_n$ and g is differentiable at $b = f(a) \in \mathcal{O}_m$ the chain rule states that the derivative of the composite function $g \circ f$ is given by,

$$\begin{aligned} \underbrace{D(g \circ f)(a)}_{\mathbb{R}^n \to \mathbb{R}^p} &= (Dg)(f(a)) \cdot (Df)(a) \\ &= \underbrace{Dg(b)}_{\mathbb{R}^m \to \mathbb{R}^p} \cdot \underbrace{Df(a)}_{\mathbb{R}^n \to \mathbb{R}^m} \end{aligned}$$

This can be expressed in matrix form as,

$$J_{g \circ f}(a)_{p \times n} = J_g(f(a))_{p \times m} \cdot J_f(a)_{m \times n} \tag{10.1}$$

Moreover, if we consider the function components in each individual coordinates as

- $g \circ f = ((g \circ f)_1, (g \circ f)_2, \ldots, (g \circ f)_p)$
- $g = (g_1, g_2, \ldots, g_p)$
- $f = (f_1, f_2, \ldots, f_m)$.

Then, the $(i,j)^{\text{th}}$ entry of both sides of (10.1) would become,

$$\frac{\partial (g \circ f)_i}{\partial x_j}(a) = \begin{pmatrix} \frac{\partial g_i}{\partial y_1}(f(a)) & \frac{\partial g_i}{\partial y_2}(f(a)) & \cdots & \frac{\partial g_i}{\partial y_m}(f(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(a) \\ \frac{\partial f_2}{\partial x_j}(a) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(a) \end{pmatrix}$$

which is our familiar chain rule for partial derivatives,

$$\boxed{\frac{\partial \left(g \circ f\right)_i}{\partial x_j}\left(a\right) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}\left(b\right) \cdot \frac{\partial f_k}{\partial x_j}\left(a\right)}$$

where, $b = f(a)$.

Furthermore, if we define $y_k = f_k\left(x_1, x_2, \dots, x_n\right)$ and $z_i = g_i\left(y_1, y_2, \dots, y_m\right)$ we can express the chain rule for partial derivatives as,

$$\boxed{\frac{\partial z_i}{\partial x_j}\left(x\right) = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k}\left(y\right) \cdot \frac{\partial y_k}{\partial x_j}\left(x\right)}$$

Remark. In addition, the chain rule can be applied in the context of a composition map with respect to a parameter t. For functions $f : \mathcal{O}_n \to \mathbb{R}$, $\eta : \mathcal{O}_1 \to \mathcal{O}_n$ and $z = f \circ \eta$ as shown below,

$$\begin{tikzcd}
\mathcal{O}_1 \arrow[r, "\eta"] & \mathcal{O}_n \arrow[r, "f"] & \mathbb{R} \\
t \arrow[r, mapsto, "\eta"] \arrow[rr, bend right, "z=f\circ\eta"'] & (x_1(t), \dots, x_n(t)) \arrow[r, mapsto, "f"] & f\left(x_1(t), \dots, x_n(t)\right)
\end{tikzcd}$$

the chain rule states,

$$\boxed{\frac{\mathrm{d}z}{\mathrm{d}t} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \cdot \frac{\mathrm{d}x_k}{\mathrm{d}t}}$$

If we treat f as a function of t, the same can be written as,

$$\boxed{\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \cdot \frac{\mathrm{d}x_k}{\mathrm{d}t}}$$

Lecture 11

11.1 Chain Rule

We will begin by recalling some results from the previous lecture.

$$\begin{tikzcd}
\mathcal{O}_n \arrow[r, "f"'] \arrow[rr, bend left, "g\circ f"] & \mathcal{O}_m \arrow[r, "g"'] & \mathbb{R}^p
\end{tikzcd}$$

If we are given two differentiable function $f : \mathcal{O}_n \to \mathcal{O}_m$ and $g : \mathcal{O}_m \to \mathbb{R}^p$, then $g \circ f$ is also differentiable. We also derived how to compute $D_{g\circ f}$ by **chain rule** as following,

$$D_{g\circ f}(a) = D_g(f(a)).D_f(a)$$

Now, comparing the $(i, j)^{\text{th}}$ element, we get,

$$\frac{\partial(g \circ f)_i(a)}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial g_i(b)}{\partial y_k} \cdot \frac{\partial f_k(a)}{\partial x_j}$$

where $b = f(a)$. This can be rewritten in a slightly more suggestive form by introducing new variables,

$$\begin{aligned} y_k &= f_k(x_1, \ldots, x_n) \\ z_i &= g_i(y_1, \ldots, y_m) \end{aligned}$$

Then, since $(g \circ f)_i = g_i \circ f$, the equation above can be written as,

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_j}$$

This form of the chain rule is reminiscent of the one-variable chain rule.

Example 11.1.1

Let, $f(x, y, z) = xy^2z$ and $x = t, y = e^t, z = 1 + t$, we want to calculate $\frac{\mathrm{d}f}{\mathrm{d}t}$ in two ways.

First, we can write f as a function of t,

$$\begin{aligned} f(x, y, z) &= t(e^t)^2(1 + t) \\ &= (t + t^2)e^{2t} \end{aligned}$$

Hence, we have,

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}(t + t^2)e^{2t} \\ &= (1 + 2t)e^{2t} + 2(t + t^2)e^{2t} \end{aligned}$$

$$= (2t^2 + 4t + 1)e^{2t}$$

Alternatively, if we apply the chain rule, we obtain,

$$\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial f}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t} \\
&= y^2 z \cdot 1 + 2xyz \cdot e^t + xy^2 \cdot 1 \\
&= e^{2t}(1+t) + 2t(1+t)e^t.e^t + te^{2t} \\
&= e^{2t}(1 + t + 2t + 2t^2 + t) \\
&= (2t^2 + 4t + 1)e^{2t}
\end{aligned}$$

As we can see, both methods yield the same result!

11.2 Laplacian

The Laplacian operator plays a fundamental role in analyzing the behavior of functions and fields in multidimensional spaces. It quantifies the overall rate of change and spatial variations of a function, providing valuable insights into its properties and behavior.

Definition 11.2.1 ▶ Laplacian

$f : \mathcal{O}_n \to \mathbb{R}$ be a function. Then the Laplacian of f is defined as,

$$\Delta f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

Observe that,

$$\begin{aligned}
\Delta f &= \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} \\
&= \left\langle \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right\rangle \cdot \left\langle \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right\rangle \\
&= \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f
\end{aligned}$$

Hence, Laplacian can be written as, $\Delta f = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f = \nabla^2 f$.

Laplacian in Polar Coordinate

Let, f be a twice differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$. We can express $f(x, y)$ in polar coordinates as a function of (r, θ) by substituting, $x = r\cos\theta$ and $y = r\sin\theta$. Now, observe the following partial derivatives,

$$\begin{aligned}
\frac{\partial x}{\partial r} &= \cos\theta, \quad \frac{\partial x}{\partial \theta} = -r\sin\theta \\
\frac{\partial y}{\partial r} &= \sin\theta, \quad \frac{\partial y}{\partial \theta} = r\cos\theta
\end{aligned}$$

We want to express f_{xx} and f_{yy} in terms of partial derivatives of f in polar coordinates. Notice that,

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}$$

$$\text{i.e., } \frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta$$

Differentiating once more with respect to r, we have,

$$\begin{aligned}
\frac{\partial^2 f}{\partial r^2} &= \frac{\partial}{\partial r}\left[\frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\right] \\
&= \cos\theta\left[\frac{\partial}{\partial r}\frac{\partial f}{\partial x}\right] + \sin\theta\left[\frac{\partial}{\partial r}\frac{\partial f}{\partial y}\right] \\
&= \cos\theta\left[\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y \partial x}\frac{\partial y}{\partial r}\right] + \sin\theta\left[\frac{\partial^2 f}{\partial x \partial y}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial r}\right] \\
&= \cos\theta\left[\frac{\partial^2 f}{\partial x^2}\cos\theta + \frac{\partial^2 f}{\partial y \partial x}\sin\theta\right] + \sin\theta\left[\frac{\partial^2 f}{\partial x \partial y}\cos\theta + \frac{\partial^2 f}{\partial y^2}\sin\theta\right] \\
&= \cos\theta\left[\cos\theta f_{xx} + \sin\theta f_{xy}\right] + \sin\theta\left[\cos\theta f_{xy} + \sin\theta f_{y}\right]
\end{aligned}$$

Hence, we get,

$$\boxed{\frac{\partial^2 f}{\partial r^2} = \cos^2\theta f_{xx} + \sin^2\theta f_y + \sin 2\theta f_{xy}}$$

Similarly, we can find the expression for $\frac{\partial^2 f}{\partial \theta^2}$,

$$\boxed{\frac{\partial^2 f}{\partial \theta^2} = -r\left(\cos\theta f_x + \sin\theta f_y\right) + \left(r^2\sin^2\theta f_{xx} + r^2\cos^2\theta f_y - r^2\sin 2\theta f_{xy}\right)}$$

Combining the above two result we can write,

$$\boxed{\Delta f = f_{xx} + f_{yy} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\cdot\frac{\partial f}{\partial r} + \frac{1}{r^2}\cdot\frac{\partial^2 f}{\partial \theta^2}}$$

Example 11.2.1 (Writing Laplacian in New coordinate)

Let, $z = z(u, v)$ where,

$$u(x, y) = x^2 y \text{ and } v(x, y) = 3x + 2y$$

We want to express the Laplacian with respect to u and v. Starting with the given coordinates,

$$\begin{aligned}
\frac{\partial u}{\partial x} &= 2xy, \quad \frac{\partial u}{\partial y} = x^2 \\
\frac{\partial v}{\partial x} &= 3, \quad \frac{\partial v}{\partial y} = 2
\end{aligned}$$

We can find $\frac{\partial z}{\partial x}$ using the chain rule,

$$\begin{aligned}
\frac{\partial z}{\partial x} &= \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} \\
\implies \frac{\partial z}{\partial x} &= 2xy\frac{\partial z}{\partial u} + 3\frac{\partial z}{\partial v}
\end{aligned}$$

Differentiating once more with respect to x, we have,

$$\begin{aligned}
\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\left[2xy\frac{\partial z}{\partial u} + 3\frac{\partial z}{\partial v}\right] \\
&= 2y\frac{\partial z}{\partial u} + 2xy\frac{\partial}{\partial x}\left[\frac{\partial z}{\partial u}\right] + 3\frac{\partial}{\partial x}\left[\frac{\partial z}{\partial v}\right]
\end{aligned}$$

$$\begin{aligned}
&= 2y\frac{\partial z}{\partial u} + 2xy\left(\frac{\partial}{\partial u}\left[\frac{\partial z}{\partial u}\right]\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left[\frac{\partial z}{\partial u}\right]\frac{\partial v}{\partial x}\right) + 3\left(\frac{\partial}{\partial u}\left[\frac{\partial z}{\partial v}\right]\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left[\frac{\partial z}{\partial v}\right]\frac{\partial v}{\partial x}\right)\\
&= 2y\frac{\partial z}{\partial u} + 2xy\left(\frac{\partial^2 z}{\partial u^2}\frac{\partial u}{\partial x} + \frac{\partial^2 z}{\partial v \partial u}\frac{\partial v}{\partial x}\right) + 3\left(\frac{\partial^2 z}{\partial u \partial v}\frac{\partial u}{\partial x} + \frac{\partial^2 z}{\partial v^2}\frac{\partial v}{\partial x}\right)\\
&= 2y\frac{\partial z}{\partial u} + 2xy\left(2xy\frac{\partial^2 z}{\partial u^2} + 3\frac{\partial^2 z}{\partial v \partial u}\right) + 3\left(\frac{1}{2xy}\cdot\frac{\partial^2 z}{\partial u \partial v} + 3\frac{\partial^2 z}{\partial v^2}\right)
\end{aligned}$$

Hence, we get,

$$\boxed{\frac{\partial^2 z}{\partial x^2} = 2yz_u + 4x^2y^2z_{uu} + 6xyz_{uv} + 6xyz_{vu} + 9z_{vv}}$$

Exercise. Find z_{yy}, z_{yx}, z_{xy} and check if $z_{xy} = z_{yx}$.

11.3 Extrema of a function

Finding the extrema of a function is crucial in calculus, allowing us to identify maximum and minimum points and to relate the structure of functions. We will now extend this concept to functions of several variables.

Definition 11.3.1 ▶ Extrema

Let, a is an interior point of $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ be a function.

- f attains a **local maximum** at a if there exists an open neighborhood $\mathcal{O}_n$ of a such that, $f(a) \geq f(x)\ \forall\ x \in \mathcal{O}_n$.
- Similarly, f attains a **local minimum** at a if there exists an open neighborhood $\mathcal{O}_n$ of a such that, $f(a) \leq f(x)\ \forall\ x \in \mathcal{O}_n$.

Any point at which f attains a local(global) maxima (or minima) is called extremum point of that function. In plural, it is called **Extrema**.

Definition 11.3.2 ▶ Critical Point or Stationary Point

Let, $f : S(\subseteq \mathbb{R}^n) \to \mathbb{R}$ be a function and $a \in \mathcal{O}_n \subseteq S$. We say that a is a **critical point** or **stationary point**. If

$$(\boldsymbol{\nabla} f)(a) = 0$$

Or, equivalently all the partial derivatives $\frac{\partial f}{\partial x_i}$ are zero.

Theorem 11.3.1

Let, $f : \mathcal{O}_n \to \mathbb{R}$ is differentiable at $a \in \mathcal{O}_n$. If a is a local extremum, then

$$(\boldsymbol{\nabla} f)(a) = 0$$

Proof. Fix $i \in \{1, 2, \ldots, n\}$. We want to show $\frac{\partial f}{\partial x_i} = 0$. For this set, $\phi_i : (a_i - \epsilon, a_i + \epsilon) \to \mathbb{R}$ defined by

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

Notice that, $\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = f_{x_i}(a)$. Since a is local extremum of f, we can say that a_i is a local extremum of ϕ_i. So, $\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = 0$, which means, $\frac{\partial f(a)}{\partial x_i} = 0$. We can do this for all i and hence, $(\boldsymbol{\nabla} f)(a) = 0$. ☐

Question. When we did calculation for local extremum for the functions f with one variable, we used to evaluate the stationary points by calculating, $f'(x) = 0$. Then we used to check the second derivative in order to know whether the stationary point is local minima or maxima or saddle point. For multivariate case also, we need 2nd order derivative to know the behavior of the stationary point. Now what could be 2nd order total derivative?

Answer. For this purpose we will introduce **Hessian Matrix** in next class.

Lecture 12

12.1 Hessian Matrix

We start by defining Hessian matrix, which is a natural extension of the concept of the second derivative in higher dimensions, allowing us to analyze the rate of change and curvature of a function in multiple directions simultaneously.

Definition 12.1.1 ▶ Hessian

Suppose $f : \mathcal{O}_n \to \mathbb{R}$ is C^2 at $a \in \mathcal{O}_n$. The **Hessian of f at** a is defined as the matrix,

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{n\times n}$$

In explicit notation, it has the following form,

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

It is important to note that for any function f that is twice continuously differentiable ($f \in C^2$), its Hessian matrix H_f is symmetric, meaning that $H_f = {H_f}^t$.

Example 12.1.1

Let, $f : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by $f(x, y) = \sin^2 x + x^2 y + y^2$. Then,

$$Df = \begin{pmatrix}\sin 2x + 2xy & x^2 + 2y\end{pmatrix} : \mathbb{R}^2 \to \mathbb{R} \quad \text{is linear.}$$

The gradient is given by,

$$\boldsymbol{\nabla} f = \left\langle \sin 2x + 2xy, x^2 + 2y \right\rangle \in \mathbb{R}^2$$

And the Hessian matrix H_f is,

$$H_f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2(\cos 2x + y) & 2x \\ 2x & 2 \end{pmatrix} \quad [\because f \in C^2]$$

Now let's introduce some notation. Given $A = (a_{ij})_{n\times n} \in M_n(\mathbb{R})$ and $x \in \mathbb{R}^n$ we denote $Q_A(x)$ by,

$$Q_A(x) = x^t A x = \langle Ax, x \rangle_{\mathbb{R}^n}$$

$$= \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{31} & a_{32} & \cdots & a_{3n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

Definition 12.1.2 ▶ Quadratic Form

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called a **Quadratic Form** if it can be expressed as $f(x) = Q_A(x)$ for all x and for some symmetric $A \in M_n(\mathbb{R})$

It is important to note that a Quadratic Form represents a homogeneous polynomial of degree 2. For instance, in the case of a bivariate polynomial $p(x, y) = a_{11}x^2 + a_{22}y^2 + a_{12}xy$, the matrix

$$A = \begin{pmatrix} a_{11} & \frac{1}{2}a_{12} \\ \frac{1}{2}a_{12} & a_{22} \end{pmatrix}$$

corresponds to the quadratic form $p(x, y) = Q_A(x, y)$, capturing the essential quadratic behavior of p.

12.2 Positive Definite, Negative Definite, Semi Definite Matrices

Definition 12.2.1 ▶ Positive Definite, Negative Definite, Semi Definite

- A symmetric matrix $A \in M_n(\mathbb{R})$ is called **Positive Definite** if
$$\langle Ax, x\rangle > 0 \; \forall \; x \in \mathbb{R}^n \setminus \{0\}$$
- A symmetric matrix $A \in M_n(\mathbb{R})$ is called **Negative Definite** if
$$\langle Ax, x\rangle < 0 \; \forall \; x \in \mathbb{R}^n \setminus \{0\}$$
- A symmetric matrix $A \in M_n(\mathbb{R})$ is called **Semi Definite** if
$$\langle Ax, x\rangle \geq 0 \; \forall \; x \in \mathbb{R}^n \setminus \{0\}$$

Example 12.2.1

1. I_n is positive definite because for any vector $x \in \mathbb{R}^n \setminus \{0\}$ the inner product
$$\langle I_n x, x\rangle = \|x\|^2 > 0$$
is strictly positive.
2. For any matrix $A \in M_n(\mathbb{R})$, if there exists a matrix $B \in M_n(\mathbb{R})$ such that $A = B^tB$, we can examine the inner product $\langle Ax, x\rangle$ for any $x \in \mathbb{R}^n \setminus \{0\}$.

   Let's compute this inner product
$$\begin{aligned} \langle Ax, x\rangle &= \langle B^tBx, x\rangle \\ &= x^tB^tBx \\ &= (Bx)^t(Bx) \\ &= \|Bx\|^2 \quad \text{for all } x \in \mathbb{R}^n \setminus \{0\}. \end{aligned}$$

Therefore, we conclude that $\langle Ax, x\rangle \geq 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

Moreover, if $\langle Ax, x\rangle = 0$, then $Bx = 0$, which implies that x is in the kernel of B. Conversely, if A is positive definite, there is no non-zero vector x such that $\langle Ax, x\rangle = 0$. This implies that the columns of B are linearly independent.

In summary, when A can be written as $A = B^tB$ for some matrix B, we can conclude that A is positive semi-definite (and the converse also holds).

3. Consider the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

We can compute the quadratic form associated with A as $Q_A = x_1^2 - x_2^2$.

By examining this expression, we observe that it is the difference between the squares of two variables. This indicates that the sign of Q_A can change depending on the values of x_1 and x_2. Consequently, the matrix A is considered **indefinite**.

4. Consider the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

To determine the definiteness of A, we can compute the quadratic form associated with A as $Q_A(x) = x_1^2 \geq 0$ for all $x = \begin{pmatrix} x_1 & x_2 \end{pmatrix}$. This indicates that the matrix A is **positive semi-definite**.

Consider a positive definite matrix A. For any non-zero vector h, we have,

$$\langle Ah, h\rangle = \|Ah\|\|h\| \cos\theta > 0$$

where θ is the angle between vectors Ah and h. Since the cosine of any angle θ in the interval $[0, \pi/2)$ is positive, we can conclude that,

$$\cos\theta > 0 \implies \boxed{0 \leq \theta < \frac{\pi}{2}}$$

Thus, for any positive definite matrix A, the angle θ between Ah and h satisfies $0 \leq \theta < \frac{\pi}{2}$.

It is worth noting that classifying positive definite matrices becomes more challenging for higher dimensions $(n > 2)$. However, for 2×2 matrices, we can easily determine it from the next theorem.

Theorem 12.2.1

Let $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in M_2(\mathbb{R})$ be symmetric. Then,

(i) A is **Positive Definite** $\iff a > 0$ and $ac - b^2 > 0$

(ii) A is **Negative Definite** $\iff a < 0$ and $ac - b^2 > 0$

(iii) A is **Indefinite** $\iff ac - b^2 < 0$

Proof. We have $\langle Ah, h\rangle = h^tAh$ for any vector h. Now, consider a non-zero vector $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$ with $x_2 \neq 0$. Without loss of generality, we can scale $\mathbf{x}$ as $\mathbf{x} = (x, 1)$ for some $x \in \mathbb{R}$. Then, we have,

$$\langle A\mathbf{x}, \mathbf{x}\rangle = ax^2 + 2bx + c > 0 \ \forall\ x \in \mathbb{R}$$

If $x_2 = 0$, we can choose $\mathbf{x} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ (after scaling). Then, $\langle A\mathbf{x}, \mathbf{x}\rangle = a$. Therefore, we can summarize the conditions as follows,

$$A \text{ is Positive Definite}$$

$$\begin{aligned}
&\iff a > 0 \text{ and } ax^2 + bx + c > 0 \;\forall\; x \in \mathbb{R} \\
&\iff a > 0 \text{ and } (2b)^2 - 4ac < 0 \\
&\iff a > 0 \text{ and } ac - b^2 > 0
\end{aligned}$$

Similarly, we can derive the conditions for negative definite and indefinite matrices as,

$$\begin{aligned}
&A \text{ is Negative Definite} \\
&\iff a < 0 \text{ and } ax^2 + bx + c < 0 \;\forall\; x \in \mathbb{R} \\
&\iff a < 0 \text{ and } (2b)^2 - 4ac < 0 \\
&\iff a < 0 \text{ and } ac - b^2 > 0
\end{aligned}$$

And finally for indefinite ones,

$$\begin{aligned}
&A \text{ is Indefinite} \\
&\iff ax^2 + bx + c < 0 \text{ for some } x \in \mathbb{R} \\
&\qquad \text{and } ax^2 + bx + c > 0 \text{ for some } x \in \mathbb{R} \\
&\iff (2b)^2 - 4ac > 0 \\
&\iff ac - b^2 < 0
\end{aligned}$$

☐

> **Lemma 12.2.1**
>
> Let, $a \in \mathcal{O}_n,\ A(x) = \begin{pmatrix} a_1(x) & a_2(x) \\ a_2(x) & a_3(x) \end{pmatrix}$. Suppose, A is continuous at a (i.e., a_i's are continuous at a). Then, A is Positive Definite at a would imply that A is Positive Definite in a neighborhood of a.

Proof. $A(a)$ is Positive Definite, i.e., $a_1(a) > 0$ and $a_1(a)a_3(a) - {a_2}^2(a) > 0$. As $a_1(x)$ and $a_1(x)a_3(x) - a_2^2(x)$ are polynomial of continuous functions, we can find an $\epsilon > 0$ such that both are positive in $B_\epsilon(a)$. ☐

Lecture 13

13.1 Taylor's Theorem

Recall, Taylor's theorem for one variable.

Definition 13.1.1 ▶ Taylor's Polynomial

Let, $f : \mathcal{O}_1 \to \mathbb{R}$ be C^k $(k \in \mathbb{N})$. Then for all h such that $a + h \in \mathcal{O}_1$,

$$p_{a,k}(a+h) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!} h^n$$

is called the **Taylor's Polynomial** of f around a.

Question. "Is $f(x) \approx p_{a,k}(x)$, for x close to a"?
We have,

$$p_{a,k}(x) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!} (x-a)^n$$

Take, $f(x) - p_{a,k}(x) = r_{a,k}(x)$

Theorem 13.1.1 (Taylor's Theorem)

Let $f : \mathcal{O}_1 \to \mathbb{R}$ be C^{k+1}. Then, $f(x) = p_{a,k}(x) + r_{a,k}(x)$ where,

$$r_{a,k} = \frac{f^{k+1}(c)}{(k+1)!}(x-a)^{k+1}$$

for some c in between a and $x \in \mathcal{O}_1$.

We introduce the following notation for the sake of clarity in the multivariate Taylor expansion. Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$, and define

- $|\alpha| = \sum_{i=1}^{n} \alpha_i$
- $\alpha! = \alpha_1! \alpha_2! \cdots \alpha_n!$ (product of coordinate-wise factorials)
- $\partial^\alpha = \dfrac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1} \cdots \partial {x_n}^{\alpha_n}}$ (α^{th} derivative)
- $\boldsymbol{\nabla} \cdot h = \displaystyle\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i}$

The last definition when iterated gives,

$$(\boldsymbol{\nabla} \cdot h)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} h^\alpha \partial^\alpha$$

Theorem 13.1.2 (Taylor's Theorem in Multivariate Case)

Let, $f : \mathcal{O}_n \to \mathbb{R}$ be a C^{k+1} function and assume $\mathcal{O}_n$ is convex. If $h, a+h \in \mathcal{O}_n$, then

$$f(a+h) = \sum_{|\alpha| \le k} \frac{1}{\alpha!} (\partial^\alpha f)(a) h^\alpha + r_{a,k}(h)$$

where,

$$r_{a,k}(h) = \sum_{|\alpha| = k+1} \frac{1}{\alpha!} (\partial^\alpha f)(a + ch) h^\alpha \qquad \text{for some } c \in (0,1)$$

Proof. Define, $\eta : [0,1] \to \mathbb{R}$ as $\eta(t) = f(a+th)$

$$\begin{tikzcd} t \arrow[r, mapsto] \arrow[rr, bend right, "\eta"'] & a + th \arrow[r, mapsto] & f(a+th) \end{tikzcd}$$

which implies η is a C^{k+1} function around 0.

$$\therefore \eta'(t) = \boldsymbol{\nabla} f(a+th) \cdot h = (\boldsymbol{\nabla} \cdot h) f(a+th)$$

Claim

$$\eta^{(m)}(t) = (\boldsymbol{\nabla} \cdot h)^m f(a+th) \ \forall \ m \in \{0, 1, \dots, k+1\}$$

Proof. The first derivative of η,

$$\eta'(t) = \boldsymbol{\nabla} f(a+th) \cdot h = \sum_{i=1}^{n} f_{x_i}(a+th) h_i$$

which we use to compute the second derivative,

$$\begin{aligned} \eta''(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i=1}^{n} f_{x_i}(a+th) h_i \right) \\ &= \sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t} \left(f_{x_i}(a+th) h_i \right) \\ &= \sum_{i=1}^{n} h_i \sum_{j=1}^{n} f_{x_i x_j}(a+th) h_j && \text{(Chain rule of partials)} \\ &= \sum_{i,j=1}^{n} h_i h_j f_{x_i x_j}(a+th) \\ &= (\boldsymbol{\nabla} \cdot h)^2 f(a+th) \end{aligned}$$

Proceeding with induction on the order of the derivative, we get $\eta^{(m)}(t) = (\boldsymbol{\nabla} \cdot h)^m f(a+th)$ for all $0 \le m \le k+1$ which is our claim. ☐

By one-variable Taylor's Theorem,

$$\eta(1) = p_{0,k}(1) + r_{0,k}(c) \text{ for some } c \in (0,1) \tag{13.1}$$

with

$$p_{0,k}(1) = \eta(0) + \frac{\eta'(0)}{1!} + \cdots + \frac{\eta^{(k)}(0)}{k!} \tag{13.2}$$

and

$$r_{0,k}(c) = \frac{\eta^{(k+1)}(c)}{(k+1)!} \tag{13.3}$$

Substituting $\eta^{(m)}(t)$ in (13.1) we have,

$$f(a+h) = \sum_{|\alpha| \le k} \frac{1}{\alpha!}(\partial^\alpha f)(a)h^\alpha + r_{a,k}(h)$$

□

We note that, in particular, if $f : \mathcal{O}_2 \to \mathbb{R}$ is a C^2 function then we have,

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2}h^t H_f(a+ch)h \tag{13.4}$$

where,

$$H_f(a) = \begin{pmatrix} f_{xx}(a) & f_{xy}(a) \\ f_{xy}(a) & f_{yy}(a) \end{pmatrix}$$

and $c \in (0,1)$.

Theorem 13.1.3 (Extremum)

Let $f : \mathcal{O}_2 \to \mathbb{R}$ be a C^2 function such that $Df(a) = 0$. We write

$$H_f(a) = \begin{pmatrix} f_{xx}(a) & f_{xy}(a) \\ f_{xy}(a) & f_{yy}(a) \end{pmatrix}$$

Then,

(i) $f(a)$ is a local maximum if $f_{xx}(a) < 0$ and $\det(H_f(a)) > 0$

(ii) $f(a)$ is a local minimum if $f_{xx}(a) > 0$ and $\det(H_f(a)) > 0$

(iii) a is a saddle point if $\det(H_f(a)) < 0$

Proof. As a is an interior point of $\mathcal{O}_2$ we can get an $r > 0$ such that $a, a+h \in B_r(a) \subseteq \mathcal{O}_2$. By (13.4),

$$f(a+h) - f(a) = \nabla f(a) \cdot h + \frac{1}{2}h^t H_f(a+ch)h$$

We will prove (ii), other statements can be proved similarly. Our assumptions tell that $H_f(a)$ is positive definite. Hence, by the Lemma 12.2.1, there exist $\epsilon > 0$ such that $H_f(x)$ is positive definite $\forall\ x \in B_\epsilon(a)$. So for every $x \in B_\epsilon(a)$ the quadratic form $h^t H_f(x)h > 0$ with $h \neq 0$ which implies $f(x) - f(a) > 0$ in $B_\epsilon(a)$ that means a is a point of local minimum. □

Example 13.1.1 (Finding Critical points of a function and their nature)

Find the critical points and discuss the nature of the function

$$f(x,y) = x^3 - 6x^2 - 8y^2$$

Solution. Setting $\nabla f(x,y) = 0$, i.e., $(f_x, f_y)(x,y) = 0$, we get the system of equations

$$3x^2 - 12x = 0 \text{ and } -16y = 0$$

whose solution set is $\{(0,0),(4,0)\}$ implying that $(0,0),(4,0)$ are critical points.
The 2nd derivatives are,

$$f_{xx}(x,y) = 6x - 12,\ f_{yy}(x,y) = -16 \text{ and } f_{xy}(x,y) = 0$$

Now, we compute the determinant of the hessian at these points to tell their nature. For $(0,0)$,

$$\det(H_f(0,0)) = \begin{vmatrix} -12 & 0 \\ 0 & -16 \end{vmatrix} > 0 \text{ and } f_{xx}(0,0) = -12 < 0$$

So, f has a local maximum at $(0,0)$. And at $(4,0)$,

$$\det(H_f(4,0)) = \begin{vmatrix} 12 & 0 \\ 0 & -16 \end{vmatrix} < 0$$

which shows that $(4,0)$ is a saddle point. ■

Lecture 14

14.1 Compact subsets of $\mathbb{R}^n$

We start with the definition of Compactness which refers to a property of sets that captures the notion of being finite or having no “holes”.

Definition 14.1.1 ▶ Compact Subset

A subset $K \subseteq \mathbb{R}^n$ is said to be compact if every sequence $\{x_n\} \subseteq K$ has a subsequence $\{x_{n_k}\}$ that is convergent to some $x \in K$.

This is known as the Bolzano-Weierstrass Property.

Observe that a compact subset of $\mathbb{R}^n$ is always **closed**. To see this, note that every sequence $\{x_n\} \subseteq K$, where K is a compact subset of $\mathbb{R}^n$ that converges to some $x \in \mathbb{R}^n$ has a convergent subsequence $\{x_{n_k}\}$ that converges to the same x. Since K is compact, we can say that $x \in K$. So, the convergent sequence $\{x_n\}$ converges to a point in K. Hence, K is closed.

More is true. A compact subset of $\mathbb{R}^n$ is **bounded** too. Assume that a compact subset $K \subseteq \mathbb{R}^n$ is not bounded. Note that, a subset of $\mathbb{R}^n$ is bounded iff it is contained inside an open ball. Since K is unbounded, we can get a sequence $\{x_m\} \subseteq K$ with $\|x_m\| > m$, which doesn't have a convergent subsequence. This shows that K is not compact that contradicts our assumption.

Therefore, a compact subset of $\mathbb{R}^n$ is closed and bounded. What about the converse?

Theorem 14.1.1

A closed and bounded box in $\mathbb{R}^n$ is compact.

Proof. We take a closed and bounded box $K := \displaystyle\prod_{i=1}^{n}[a_i, b_i] \subseteq \mathbb{R}^n$. Fix $i \in [n]$. Consider a sequence $\{x_m\} \subseteq K$. We take its projection on the i^{th} coordinate, i.e., $\{\pi_i(x_m)\} \subseteq [a_i, b_i]$. Consider $i = 1$, by Bolzano-Weierstrass Theorem, it has a convergent subsequence $\{\pi_1(x_{m_t})\} \subseteq [a_1, b_1]$ converging to $\alpha_1 \in [a_1, b_1]$. Now we take $i = 2$ and repeat the process to get a convergent subsequence $\left\{\pi_2(x_{m_{t_l}})\right\} \subseteq [a_2, b_2]$ converging to $\alpha_2 \in [a_2, b_2]$. Continuing this way, we get a convergent subsequence of $\{x_m\}$ converging to $\alpha = (\alpha_1, \dots, \alpha_n) \in K$. Hence, K is compact. ☐

Theorem 14.1.2 (Heine-Borel Theorem)

A subset $K \subseteq \mathbb{R}^n$ is compact iff it is closed and bounded.

Proof. $\Longrightarrow$ Done!

$\Longleftarrow$ Since, K is bounded, it is contained in a closed box, i.e., there exists $r > 0$ such that $K \subseteq [-r, r]^n$. So, Theorem 14.1.1 implies that all sequences in K has a convergent subsequence, which must converge in K because K is closed. Hence, K is compact.

☐

Theorem 14.1.3

Let $f : \mathcal{O}_n \to \mathbb{R}^m$ be a continuous map. Then f sends compact sets to compact sets.

In other words, continuous image of a compact set is compact.

Proof. Let $K \in \mathcal{O}_n$ be compact. Take a sequence $\{x_k\} \subseteq K$ with a convergent subsequence $\{x_{k_t}\} \subseteq K$ converging to $x \in K$. Then, $\{f(x_k)\}$ is sequence in $f(K)$ with convergent subsequence $\{f(x_{m_t})\}$ converging to $f(x)$. The last statement about convergence follows from the continuity of f. This shows that $f(K)$ is compact. ☐

Theorem 14.1.4 (Extreme Value Theorem)

Let $K \subseteq \mathbb{R}^n$ be compact and $f : K \to \mathbb{R}$ a continuous map. Then $\exists\ a, b \in K$ such that $f(a) \leq f(x) \leq f(b)$ for all $x \in K$.

Proof. By Theorem 14.1.3, $f(K)$ is compact. So f is bounded which implies $\sup\limits_K f, \inf\limits_K f$ exist! Since $f(K)$ is closed, they must exist inside $f(K)$. ☐

14.2 Inverse Function Theorem

We are now going to study Inverse Function Theorem which relates the differentiability of a function to the differentiability of its inverse, enabling the study of local behavior and solving equations in higher dimensions. But before that we prove a lemma that is essential to prove the theorem.

> **Lemma 14.2.1**
>
> Let $\mathcal{O}_n \subseteq \mathbb{R}^n$ be open and convex. Suppose $f : \mathcal{O}_n \to \mathbb{R}^n$ be a C^1 function. If $\exists\ M > 0$ such that
> $$\sup_{x \in \mathcal{O}_n} \left| \frac{\partial f_i}{\partial x_j}(x) \right| \leq M \text{ for all } i, j$$
> Then $\|f(x) - f(y)\| \leq n^2 M \|x - y\|$ for every $x, y \in \mathcal{O}_n$.

Proof. Pick $x, y \in \mathcal{O}_n$ and $i \in [n]$. Then using Mean Value Theorem, we can get $c_i \in L_{x,y}$ such that

$$\begin{aligned}
f_i(x) - f_i(y) &= \boldsymbol{\nabla} f_i(c_i) \cdot (x - y) \\
\implies |f_i(x) - f_i(y)| &= \left| \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(c_i) \cdot (x_i - y_i) \right| \\
&\leq \sum_{j=1}^{n} \left| \frac{\partial f_i}{\partial x_j}(c_i) \right| |x_i - y_i| && \text{(Triangle inequality)} \\
&\leq M \sum_{j=1}^{n} |x_i - y_i| \leq nM \|x - y\|
\end{aligned}$$

The last inequality follows from the inequality $|x_i - y_i| \leq \|x - y\|$ which holds for all i.

Using the above,

$$\|f(x)-f(y)\| = \sqrt{\sum_{i=1}^{n} |f_i(x)-f_i(y)|^2} \leq \sqrt{\sum_{i=1}^{n} n^2M^2\|x-y\|^2} \leq n^2M\|x-y\|$$

we obtain the result. ☐

Theorem 14.2.1 (Inverse Function Theorem)

Let $f : \mathcal{O}_n \to \mathbb{R}^n$ be a C^1 function and $a \in \mathcal{O}_n$. Suppose $Df(a)$ is invertible. Then, there exist open sets V and W containing a and $f(a)$ respectively, such that $f : V \to W$ is invertible. Moreover, the local inverse $f^{-1} \equiv \left(f|_V\right)^{-1} : W \to V$ is differentiable and for all $y \in W$,

$$Df^{-1}(y) = \left((Df)\left(f^{-1}(y)\right)\right)^{-1}$$

i.e., locally, the derivative of the inverse is the matrix inverse of the derivative.

Proof.
We call $L = Df(a)$ which is given to be invertible and take $g(x) := L^{-1}f(x)$. Then,

$$\begin{aligned} Dg(a) &= L^{-1}(f(a)) \cdot (Df)(a) \\ &= \left[L^{-1}\right] \cdot Df(a) = I \end{aligned}$$

As this transformation can be made, without loss of generality, we may assume that $Df(a) = I_n$ which would imply that there exists a closed box U containing a such that for all $x \in U \setminus \{a\}, f(a) \neq f(x)$. To see this, let $f(a) = f(a+h)$ with arbitrarily small $\|h\|$. But then,

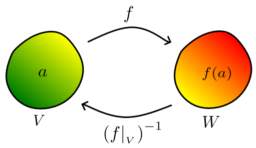


Figure 14.1

$$\frac{1}{\|h\|}\left(f(a+h)-f(a)-Ih\right) = \frac{h}{\|h\|} \neq 0$$

which contradicts the definition of derivative. Note that $\det J_f(a) \neq 0$. So, by continuity $\det J_f(x) \neq 0$ for all $x \in U$ (we may shrink U if necessary). Hence, $Df(x)$ is invertible for all $x \in U$. Again by continuity, for all $x \in U$ (we may shrink U if necessary),

$$\left| \frac{\partial f}{\partial x_j}(x) - \underbrace{\frac{\partial f}{\partial x_j}(a)}_{\delta_{ij}} \right| \leq \frac{1}{2n^2}$$

Now we claim the following,

Claim

For all $x, y \in U$,

$$\|f(x)-f(y)\| \geq \frac{1}{2}\|x-y\|$$

Proof. We take $g(x) = f(x) - x$ for all $x \in U$. Taking derivative, we get

$$\begin{aligned} Dg(x) &= Df(x) - I = Df(x) - Df(a) \\ \implies \frac{\partial g_i}{\partial x_j}(x) &= \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a) \\ \implies \left|\frac{\partial g_i}{\partial x_j}(x)\right| &\leq \frac{1}{2n^2} \ \forall\ x \in U \end{aligned}$$

Then by Lemma 14.2.1, for all $x, y \in U$,

$$\begin{aligned} \|g(x) - g(y)\| &\le n^2 \cdot \frac{1}{2n^2} \|x - y\| \\ \implies \|(f(x) - f(y)) - (x - y)\| &\le \frac{1}{2}\|x - y\| \\ \implies \boxed{\|f(x) - f(y)\| \ge \frac{1}{2}\|x - y\|} & \qquad (14.1) \end{aligned}$$

Where (14.1) follows from the Triangle inequality. So, we get the claim! It shows that f is injective. ☐

Next we look at the compact set $\partial U \subseteq U$. Since $a \notin \partial U$, we can say $f(x) \neq f(a)$ for all $x \in \partial U$. So, by continuity of f and compactness of ∂U, we can find a $d \in \mathbb{R}_{\ge 0}$ such that

$$\|f(x) - f(a)\| \ge d \quad \forall\, x \in \partial U$$

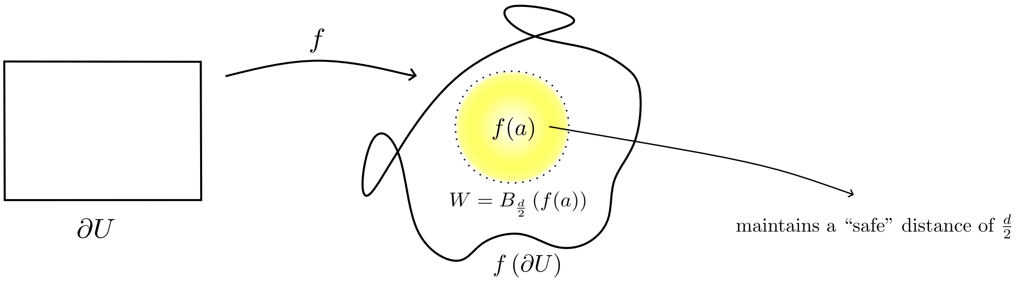


Figure 14.2

Now, we take $W = B_{\frac{d}{2}}(f(a))$. Then, for every $y \in W$ and $x \in \partial U$,

$$\underbrace{\|y - f(a)\|}_{\text{atmost } \frac{d}{2}} < \underbrace{\|y - f(x)\|}_{\text{atleast } d} \tag{14.2}$$

Claim

For a fixed $y \in W, \exists$ a unique $x_0 \in U^\circ$ such that $f(x_0) = y$

Proof. We define a continuous function $g : U \to \mathbb{R}$ with

$$g(x) = \|y - f(x)\|^2 = \sum_{i=1}^{n} (y_i - f_i(x))^2$$

Since, $\inf_U g$ cannot occur at the boundary ∂U, but must occur in U, there exists $x_0 \in U^\circ$ such that

$$\boldsymbol{\nabla} g(x_0) = 0 \text{ i.e., } \frac{\partial g}{\partial x_j}(x_0) = 0\ \forall j$$

Now, the partials of g,

$$\begin{aligned} \frac{\partial g}{\partial x_j}(x) &= \frac{\partial}{\partial x_j} \sum_{i=1}^{n} (y_i - f_i(x))^2 \\ &= -2 \sum_{i=1}^{n} (y_i - f_i(x)) \frac{\partial f_i}{\partial x_j}(x) \end{aligned}$$

At x_0, we get,

$$\begin{aligned} &\sum_{i=1}^{n} (y_i - f_i(x_0)) \frac{\partial f_i}{\partial x_j}(x_0) = 0 \; \forall j \\ \implies &\underbrace{\left(\tfrac{\partial f_i}{\partial x_j}(x_0) \right)^t}_{Df(x_0)^t} (y - f_i(x_0)) = 0 \end{aligned}$$

As Df is invertible in U and $x_0 \in U^\circ$ we obtain $y = f(x_0)$, which shows the existence of x_0. The uniqueness follows from (14.1). ☐

We now set $V := U \cap f^{-1}(W)$. Since U is closed, $V = U^\circ \cap f^{-1}(W)$. Hence, $f|_V : V \to W$ is invertible!

Claim

$f^{-1} \equiv (f|_V)^{-1} : W \to V$ is continuous.

Proof. (14.1) gives,

$$\|f(x_1) - f(x_2)\| \geq \frac{1}{2}\|x_1 - x_2\| \; \forall \; x_1, x_2 \in V \subseteq U$$

equivalently,

$$2\|y_1 - y_2\| \geq \left\|f^{-1}(y_1) - f^{-1}(y_2)\right\| \qquad \text{(where, } y_i = f(x_i)\text{)}$$

which shows that f^{-1} is Lipschitz, hence continuous. ☐

Claim

f^{-1} is differentiable.

Proof. We fix $y_0 = f(x_0) \in W$ for some $x_0 \in V$ and take $A = Df(x_0)$. As we know,

$$\begin{aligned} &\lim_{h \to 0} \frac{1}{\|h\|} \left(f^{-1}(y_0 + h) - f^{-1}(y_0) - A^{-1}h\right) = 0 \\ \iff &\lim_{y \to y_0} \frac{1}{\|y - y_0\|} \left(f^{-1}(y) - f^{-1}(y_0) - A^{-1}(y - y_0)\right) = 0 \end{aligned} \tag{14.3}$$

We set $\phi(h) = f(x_0 + h) - f(x_0) - Ah$ for h in a neighborhood of 0. Now,

$$\begin{aligned} A^{-1}\left(f(x_0 + h) - f(x_0)\right) &= h + A^{-1}\phi(h) \\ &= ((x_0 + h) - x_0) + A^{-1}\left(\phi((x_0 + h) - x_0)\right) \end{aligned}$$

Also set $y = f(x_0 + h)$. Then,

$$\begin{aligned} &A^{-1}(y - y_0) = f^{-1}(y) - f^{-1}(y_0) + A^{-1}\left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right) \\ \implies &-A^{-1}\left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right) = f^{-1}(y) - f^{-1}(y_0) - A^{-1}(y - y_0) \end{aligned} \tag{14.4}$$

So, it is now enough to prove that,

$$\begin{aligned} &\lim_{h \to 0} \frac{1}{\|h\|} A^{-1}\left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right) = 0 \\ \iff &\lim_{y \to y_0} \frac{1}{\|y - y_0\|} \left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right) = 0 \end{aligned}$$

But,

$$\frac{\left\|\phi\left(f^{-1}(y)-f^{-1}(y_0)\right)\right\|}{\|y-y_0\|}=\underbrace{\frac{\left\|\phi\left(f^{-1}(y)-f^{-1}(y_0)\right)\right\|}{\|f^{-1}(y)-f^{-1}(y_0)\|}}_{0}\cdot\underbrace{\frac{\left\|f^{-1}(y)-f^{-1}(y_0)\right\|}{\|y-y_0\|}}_{\leq 2}=0$$

Hence, the limit (14.3) is true. ☐

To show that f^{-1} is C^1, observe that all the partials of f^{-1} are rational polynomial functions (with non-zero denominators) of those of f. This completes the proof. ☐

Lecture 15

15.1 Inverse function theorem: Example

Recall that the inverse function theorem (14.2.1) states that if $f : \mathbb{R}^n \to \mathbb{R}^n$ is a C^1 function and there is $a \in \mathcal{O}_n$ such that $Df(a)$ is invertible, then there exists $\widetilde{\mathcal{O}_n}$ such that $f(a) \in \widetilde{\mathcal{O}_n}$ and $f^{-1} : \widetilde{\mathcal{O}_n} \to \mathcal{O}_n$ exists and is a C^1 function and further, $D(f^{-1}) = (Df)^{-1}$ at each point. Thus, given that f satisfies the conditions, the theorem guarantees that f is locally invertible with a differentiable inverse. We discuss an important example where the theorem is used.

Consider the polar coordinate transformation,

$$x = r\cos\theta \quad y = r\sin\theta$$

We can rephrase this with the function,

$$\begin{aligned} F : \mathbb{R}_{>0} \times \mathbb{R} &\to \mathbb{R}^2 \\ F(r,\theta) &= (x,y) \end{aligned}$$

F is defined on an open set, and it is C^1 because we have,

$$\begin{aligned} \frac{\partial x}{\partial r} &= \cos\theta & \frac{\partial y}{\partial r} &= \sin\theta \\ \frac{\partial x}{\partial \theta} &= -r\sin\theta & \frac{\partial y}{\partial \theta} &= r\cos\theta \end{aligned}$$

$$\implies J_F(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \implies \det J_F(r,\theta) = r$$

and hence, $\det J_F(r,\theta)$ is non-zero in the domain we chose. The inverse function theorem then guarantees that we can express (r,θ) as a C^1 function of (x,y), locally. Further, we also have

$$DF^{-1}(x,y) = \big(DF(F^{-1}(x,y))\big)^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

In other words,

$$\begin{aligned} \frac{\partial r}{\partial x} &= \cos\theta & \frac{\partial r}{\partial y} &= \sin\theta \\ \frac{\partial \theta}{\partial x} &= -\frac{1}{r}\sin\theta & \frac{\partial \theta}{\partial y} &= \frac{1}{r}\cos\theta \end{aligned}$$

which can also be verified directly.

15.2 Implicit Function Theorem

In one variable, we could differentiate functions given as $y = f(x)$. We have developed several variable calculus far enough that we can now do the same for functions given as $x_n = f(x_1, \ldots, x_{n-1})$. But what about functions that are given as $f(x, y) = 0$, that is, implicitly in both variables? We will now discuss an important theorem that will allow us to take derivatives of such functions as well without needing to solve the equation for the dependent variable.

Example 15.2.1

Let $F(x, y) = ax + by + c$. We have,

$$F(x,y) = 0 \iff ax + by + c = 0 \iff y = -\frac{a}{b}x - \frac{c}{b}$$

for all $b \neq 0$. Hence, given that $F(x, y) = 0$, we get

$$y = f(x) \text{ for a differentiable } f \iff b \neq 0 \iff \frac{\partial F}{\partial y} \neq 0$$

In other words, $\frac{\partial F}{\partial y} \neq 0 \iff F(x, f(x)) = 0$ for some differentiable f.

Example 15.2.2

Let $F(x, y) = x^2 + y^2 - 1$. We have,

$$F(x,y) = 0 \iff x^2 + y^2 = 1 \iff y^2 = 1 - x^2 \iff y = \pm\sqrt{1 - x^2}$$

but the last expression is not a function! More precisely,

$$\begin{aligned} y \geq 0 &\implies y = f_1(x) = \sqrt{1 - x^2} \\ y \leq 0 &\implies y = f_2(x) = -\sqrt{1 - x^2} \end{aligned}$$

Note that $\frac{\partial F}{\partial y} = 2y \neq 0$ for all $y \neq 0$. Hence, for $F(x_0, y_0) \neq 0$, $\frac{\partial F}{\partial y} \neq 0$, there exists a $C^1 f$ defined in a neighbourhood of x_0 such that $F(x, f(x)) = 0$ for all x in that neighbourhood.

Note

Consider $F(x, y) = 0$ and $y = f(x)$ for some $C^1 f$ such that $F(x, f(x)) = 0$. We have,

$$\begin{aligned} F(x, f(x)) &= 0 \\ \implies \frac{\mathrm{d}F}{\mathrm{d}x} &= 0 \\ \implies \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} &= 0 && \text{(Chain rule)} \\ \implies \frac{\mathrm{d}y}{\mathrm{d}x} &= -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \end{aligned}$$

Hence, the condition $\frac{\partial F}{\partial y} \neq 0$ is necessary for differentiating functions defined implicitly. Keeping the feeling of the above examples in mind, we now discuss the full theorem that shows that it is also sufficient.

15.2.1 Proof of the theorem

We first introduce some notation. Throughout this section, $X \in \mathbb{R}^n, Y \in \mathbb{R}^m$ so that $(X, Y) \in \mathbb{R}^{n+m}$. The point $(a, b) \in \mathbb{R}^{n+m}$ is defined with $a \in \mathbb{R}^n, b \in \mathbb{R}^m$. $\mathcal{O}$ denotes an open set in $\mathbb{R}^{n+m}$, and $F : \mathcal{O} \to \mathbb{R}^m$ has the coordinate functions $f_i(X, Y)$. Assuming $F \in C^1(\mathcal{O})$, its jacobian is

$$DF = \left(\frac{\partial f_i}{\partial x_j} \;\middle|\; \frac{\partial f_i}{\partial y_k} \right)$$

where $i, k \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$.

Theorem 15.2.1 (Implicit Function Theorem)

Let $F \in C^1(\mathcal{O})$ and $F(a, b) = 0$. If $\det \left(\frac{\partial f_i}{\partial y_k} \right)_{m \times m} \neq 0$, then there exists an open neighbourhood $U \subset \mathbb{R}^n$ of a and a C^1 function $f : U \to \mathbb{R}^m$ such that $f(a) = b$ and $F(x, f(x)) = 0$ for all $x \in U$.

Proof. We define the function $\tilde{F} : \mathcal{O} \to \mathbb{R}^n \times \mathbb{R}^m$ as

$$\widetilde{F}(X, Y) = (X, F(X, Y))$$

As F is C^1, so is $\widetilde{F}$, and we have

$$J_{\widetilde{F}} = \begin{pmatrix} I_n & 0 \\ \dfrac{\partial f_i}{\partial x_j} & \dfrac{\partial f_i}{\partial y_k} \end{pmatrix}$$

Hence, $\det J_{\widetilde{F}}(a, b) \neq 0$ as $\det \left(\frac{\partial f_i}{\partial y_k} \right) \neq 0$, and so we can use the inverse function theorem!

By Inverse Function Theorem (14.2.1), there exists a neighbourhood $U_0 \subseteq \mathcal{O}$ of (a, b) and a neighbourhood $V_0 \subseteq \mathbb{R}^{n+m}$ of $(a, 0)$ such that $\widetilde{F} : U_0 \to V_0$ has a C^1 inverse.

We shrink U_0 (also shrinking V_0 accordingly) so that $U_0 = A \times B$ where $a \in U \subseteq \mathbb{R}^n$ and $b \in U' \subseteq \mathbb{R}^m$, for open sets A, B. We also have,

$$\widetilde{F}^{-1}(X, Y) = (X, g(X, Y))$$

for some C^1 function g, from the definition of $\widetilde{F}$. Now, consider the map

$$\begin{aligned} \Pi_2 : \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R}^m \\ (X, Y) &\mapsto Y \end{aligned}$$

Then, $\Pi_2 \circ \widetilde{F} = F$. So we get,

$$\begin{aligned} & (X, g(X, Y)) = \widetilde{F}^{-1}(X, Y) \\ \implies \; & F(X, g(X, Y)) = \Pi_2(X, Y) = Y \end{aligned}$$

where $(X, Y) \in V_0, (X, g(X, Y)) \in U_0$. Hence, for $Y = 0$ and $X \in U$,

$$F(X, g(X, 0)) = 0$$

Therefore, $f(X) = g(X, 0), X \in U$, works. □

15.3 Solving systems of equations

Example 15.3.1

Consider the system of equations,

$$\begin{aligned} x^2 + 2y^2 + z^2 + w &= 6 \\ 2x^3 + 4y^2 + z + w^2 &= 9 \end{aligned}$$

We wish to know whether (z, w) can be expressed as a function of (x, y) locally. Construct the function $F : \mathbb{R}^4 \to R^2$ as $F = (f_1, f_2)$ where

$$\begin{aligned} f_1(x, y, z, w) &= x^2 + 2y^2 + z^2 + w - 6 \\ f_2(x, y, z, w) &= 2x^3 + 4y^2 + z + w^2 - 9 \end{aligned}$$

Consider $\alpha = (1, -1, -1, 2) \in \mathbb{R}^4$. We have,

$$\begin{aligned} J &= \begin{pmatrix} \frac{\partial f_1}{\partial z} & \frac{\partial f_1}{\partial w} \\ \frac{\partial f_2}{\partial z} & \frac{\partial f_2}{\partial w} \end{pmatrix} = \begin{pmatrix} 2z & 1 \\ 1 & 2w \end{pmatrix} \\ \implies \det J(\alpha) &= -9 \neq 0 \end{aligned}$$

Hence, by the implicit function theorem, there is a neighbourhood of $(1, -1)$ such that on that neighbourhood, $(z, w) = f(x, y)$ for some $C^1 f$.

Lecture 16

16.1 Riemann-Darboux Integration

Definition 16.1.1 ▶ Volume of a closed/open ball

We define the volume of a closed ball $B^n = \prod_{i=1}^{n} [a_i, b_i]$ as $\text{Vol}(B^n) = \prod_{i=1}^{n} (b_i - a_i)$. We also define the volume of the open set $O^n = \prod_{i=1}^{n} (a_i, b_i)$ to be equal to that of B^n, i.e., $\text{Vol}(O^n) = \text{Vol}(B^n)$.

We now introduce some notation. Fix $i \in \{1, 2, \ldots, n\}$. We define a partition of the i^{th} interval $[a_i, b_i]$ as

$$P_i : a_i = a_{i,0} < a_{i,1} < \cdots < a_{i,n_i} = b_i$$

and the intervals of its partition as

$$I_{i,t} = [x_{i,t-1}, x_{i,t}], \forall\ 1 \le t \le n_i$$

and set

$$B^n_{(t_1,t_2,\ldots,t_n)} = B_\alpha = [x_{1,t_1-1}, x_{1,t_1}] \times \cdots \times [x_{n,t_n-1}, x_{n,t_n}] = I_{1,t_1} \times \cdots \times I_{n,t_n}$$

where α is chosen from the indexing set $\Lambda(P) = \{\alpha = (t_1, \ldots, t_n) \mid 1 \le t_i \le n_i, i = 1, \ldots, n\}$.

Note

1. $B^n = \bigcup_{\alpha \in \Lambda(P)} B^n_\alpha$
2. $\text{Vol}(B^n) = \sum_{\alpha \in \Lambda(P)} \text{Vol}(B^n_\alpha)$

We call $\mathscr{P}(B) = \{P_1 \times \cdots \times P_n \mid P_i \in \mathscr{P}[a_i, b_i]\}$ the set of all partitions of B^n.

Definition 16.1.2 ▶ Refinement of Partitions

Given $P = \prod_{i=1}^{n} P_i$ and $\widetilde{P} = \prod_{i=1}^{n} \widetilde{P}_i$ with $P, \widetilde{P} \in \mathscr{P}[a, b]$, then $\widetilde{P}$ is called a refinement of P if $\widetilde{P}_i \supset P_i\ \forall i = 1, 2, \ldots, n$.

Theorem 16.1.1

Let f be a bounded function over B^n. Let $P, \widetilde{P} \in \mathscr{P}(B^n)$ and $\widetilde{P} \supset P$. Then

$$L(f, P) \le L(f, \widetilde{P}) \le U(f, \widetilde{P}) \le U(f, P)$$

Proof. Note that $L(f, \widetilde{P}) \leq U(f, \widetilde{P})$ follows directly from the fact that $m_\alpha(\widetilde{P}) \leq M_\alpha(\widetilde{P})\ \forall \widetilde{P} \in \mathscr{P}[a, b]$, where $m_\alpha = \inf_{B_\alpha^n} f$ and $M_\alpha = \sup_{B_\alpha^n} f$. ☐

Corollary (Inequality of upper and lower sums)

For all $P, \widetilde{P} \in \mathscr{P}(B^n)$, the following inequality holds.

$$m \times \mathrm{Vol}(B^n) \leq L(f, P) \leq U(f, \widetilde{P}) \leq M \times \mathrm{Vol}(B^n)$$

We denote $\mathscr{B}(A) = \{f : A \to \mathbb{R} \mid \sup_A |f| < \infty\}$ as the set of all bounded functions over A for any $A \subseteq \mathbb{R}^n$.

Definition 16.1.3 ▶ Upper and Lower Darboux Integrals

For $f \in \mathscr{B}(B^n)$, we define

$$\overline{\int}_{B^n} f = \inf_{P \in \mathscr{P}(B^n)} U(f, P) \text{ and } \underline{\int}_{B^n} f = \sup_{P \in \mathscr{P}(B^n)} L(f, P)$$

as the Upper and Lower Darboux Integrals, respectively.

We have $L(f, P) \leq U(f, P')$ for all $P, P' \in \mathscr{P}(B^n)$ by taking the common refinement $\widehat{P} = P \cup P'$. Hence,

$$\underline{\int}_{B^n} f \leq \overline{\int}_{B^n} f$$

Definition 16.1.4 ▶ Darboux Integral

Let $f \in \mathscr{B}(B^n)$. f is said to be Riemann-Darboux Integrable or Riemann Integrable or just Integrable if

$$\underline{\int}_{B^n} f = \overline{\int}_{B^n} f$$

In this case, we introduce the notation,

$$\int_{B^n} f \, \mathrm{d}V = \int_{B^n} f(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n = \underline{\int}_{B^n} f = \overline{\int}_{B^n} f$$

At this point, the notation $\int_{B^n} f(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n$ **does not** indicate repeated integration, but we will see that it represents repeated integration for “nice” functions.

Lecture 17

17.1 Properties of Riemann-Darboux Integration

In the previous lecture, we introduced the Riemann-Darboux Integral. In this lecture, we will explore some important properties of this integral, starting with a characterization.

Theorem 17.1.1 (Classification of Riemann Integrable Functions)

Let $f \in \mathscr{B}(B^n)$. Then $f \in \mathscr{R}(B^n)$ if and only if for every $\epsilon > 0$, there exists a partition $P \in \mathscr{P}(B^n)$ of B^n such that

$$(0 \leq)\ U(f,P) - L(f,P) < \epsilon$$

Proof. $\Longrightarrow$ Suppose $f \in \mathscr{R}(B^n)$. Then we have

$$\overline{\int} f - \underline{\int} f = 0$$

Thus,

$$\begin{aligned} 0 = \overline{\int} f - \underline{\int} f &= \inf_{P \in \mathscr{P}(B^n)} U(f,P) - \sup_{P \in \mathscr{P}(B^n)} L(f,P) \\ &= \inf_{P \in \mathscr{P}(B^n)} (U(f,P) - L(f,P)) \end{aligned}$$

Hence for all $\epsilon > 0$, there exists a partition $P \in \mathscr{P}(B^n)$ such that $U(f,P) - L(f,P) < \epsilon$.

$\Longleftarrow$ Conversely, assume that for every $\epsilon > 0$, there exists a partition $P \in \mathscr{P}(B^n)$ such that $U(f,P) - L(f,P) < \epsilon$. We want to show that $\overline{\int} f = \underline{\int} f$. Since $U(f,P) \geq \overline{\int} f$ and $L(f,P) \leq \underline{\int} f$, it follows that for all $\epsilon > 0$,

$$0 \leq \overline{\int} f - \underline{\int} f < \epsilon$$

This implies $\overline{\int} f = \underline{\int} f$, showing that $f \in \mathscr{R}(B^n)$.

□

Exercise. Let $f, g \in \mathscr{R}(B^n)$. Then show that,

- $|f| \in \mathscr{R}(B^n)$ and $\left| \int_{B^n} f \right| \leq \int_{B^n} |f|$

- $\mathscr{R}(B^n)$ is a $\mathbb{R}-$algebra by showing the following,
  - (i) For any $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g \in \mathscr{R}(B^n)$
  - (ii) $fg \in \mathscr{R}(B^n)$

Next, we demonstrate the Riemann integrability of a class of "nice" functions (continuous). However, before proceeding, let's introduce the concepts of the diameter of a set and the mesh of a partition.

Definition 17.1.1 ▶ Diameter of a set

For a set $A \subseteq \mathbb{R}^n$, the diameter $d(A)$ is defined as $d(A) = \sup\{|x - y| \mid x, y \in A\}$.

Exercise. Show that $d(B^n) = \max\{\|v_i - v_j\| \mid v_i, v_j \text{ are vertices of } B^n\}$

Definition 17.1.2 ▶ Mesh of a Partition

For a partition $P \in \mathscr{P}(B^n)$, the mesh $\|P\|$ is defined as $\|P\| = \max\{d(B^n_\alpha) \mid \alpha \in \Lambda(P)\}$.

Theorem 17.1.2 (Continuous Functions are Riemann Integrable)

The set of all continuous functions over B^n is contained in $\mathscr{R}(B^n)$, i.e.,

$$C(B^n) \subset \mathscr{R}(B^n)$$

Proof. Let $f \in C(B^n)$. Since f is uniformly continuous, for any $\epsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in B^n$ with $\|x - y\| < \delta$, we have,

$$|f(x) - f(y)| < \underbrace{\frac{\epsilon}{2\,\mathrm{Vol}(B^n)}}_{\text{Call it } \tilde{\epsilon}} \tag{17.1}$$

Let $P \in \mathcal{P}(B^n)$ be a partition such that $\|P\| < \delta$. For each $\alpha \in \Lambda(P)$, let $a_\alpha \in B^n_\alpha$. Then $\|x - a_\alpha\| < \delta$ for all $x \in B^n_\alpha$. It follows from the uniform continuity condition (17.1) that,

$$\begin{gathered} |f(x) - f(a_\alpha)| < \tilde{\epsilon} \\ \text{i.e., } f(a_\alpha) - \tilde{\epsilon} < f(x) < f(a_\alpha) + \tilde{\epsilon} \end{gathered} \tag{17.2}$$

Since, (17.2) holds for all $\alpha \in \Lambda(P)$, $a_\alpha \in B^n_\alpha$ and for all $x \in B^n_\alpha$ we have,

$$f(a_\alpha) - \tilde{\epsilon} \le m_\alpha \le M_\alpha \le f(a_\alpha) + \tilde{\epsilon}$$

Multiplying the volumes of B^n_α and summing over $\Lambda(P)$, we obtain,

$$\sum_{\alpha \in \Lambda(P)} f(a_\alpha)\,\mathrm{Vol}(B^n_\alpha) - \frac{\epsilon}{2} \le L(f,P) \le U(f,P) \le \sum_{\alpha \in \Lambda(P)} f(a_\alpha)\,\mathrm{Vol}(B^n_\alpha) + \frac{\epsilon}{2}$$

Thus, $U(f,P) - L(f,P) < \epsilon$, and since ϵ is arbitrary, we conclude that $f \in \mathscr{R}(B^n)$. ☐

Now, let's consider an important question: Does an analogue of the Fundamental Theorem of Calculus exist in higher dimensions?

In one dimension ($n = 1$), we have the useful relationship $\displaystyle\int_{[a,b]} \mathrm{d}f = f\Big|_{\partial[a,b]}$, which aids in computation.

However, this relationship becomes less practical in higher dimensions. For instance, in $n = 1$, the continuous counterpart to a sum $\sum a_n$ is the one-dimensional integral $\displaystyle\int_{B^1} f$. Similarly, in $n = 2$, the continuous analogue to a double sum $\sum a_{mn}$ is the two-dimensional integral $\displaystyle\int_{B^2} f$.

17.2 Iterated Integrals

Before delving deeper into the concept of integrability, let's take a brief detour to discuss the idea of a double sum.

Definition 17.2.1 ▶ Convergence of Double Sequence

A double sequence $\{a_{mn}\}$ converges to a if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|a_{mn} - a| < \epsilon$ for all $m, n \geq N$

Consider the following examples,

Example 17.2.1

- Let's take the sequence $\{a_{mn}\}$ defined by $a_{mn} = \frac{1}{m+n}$ for all $m, n \in \mathbb{N}$. This sequence is bounded, and for $N > \frac{1}{2\epsilon}$, we have $|a_{mn} - 0| = a_{mn} < \epsilon$ for all $m, n \geq N$.
- Now consider the sequence $\{a_{mn}\}$ defined as follows,

$$a_{mn} = \begin{cases} n & \text{if } m = 1 \\ \frac{1}{m+n} & \text{otherwise} \end{cases}$$

  This sequence is also convergent but not bounded.

Recall the relation between total limit and iterated limits in double sequence,

Theorem

For a double sequence $\{a_{mn}\}$ if $\lim_{m,n\to\infty} a_{mn}$ exists and $\lim_{m\to\infty} a_{mn}$ exists for all n, then

$$\lim_{n\to\infty}\left(\lim_{m\to\infty} a_{mn}\right) = \lim_{m,n\to\infty} a_{mn}$$

An important analogue of the above result is Fubini's Theorem. Computation of the Darboux integral is typically a challenging task. However, Fubini's Theorem offers a valuable approach that simplifies the computation by utilizing iterated integrals.

Visualization

We look at slice functions along each axis, which enables us to simplify computations and apply Fubini's Theorem for efficient evaluation of multivariable integrals.

Consider a function $f : B^2 \to \mathbb{R}$. For each $x \in [a_1, b_1]$ we define a slice function $f_x : [a_2, b_2] \to \mathbb{R}$ given by $f_x(y) = f(x, y)$ for all $y \in [a_2, b_2]$. This function is obtained by fixing x and slicing along the y-axis at that x-coordinate. Then an iterated integral becomes

$$\int_{[a_1,b_1]}\left(\int_{[a_2,b_2]} f_x(y)\,\mathrm{d}y\right)\mathrm{d}x$$

The question arise whether this quantity is invariant under the interchange of x and y, i.e., we may slice f along $x-$axis at y to obtain $f_y : [a_1, b_1] \to \mathbb{R}$ given by $f_y(x) = f(x, y)$ for every $x \in [a_1, b_1]$ and want to investigate the equality of

$$\int_{[a_1,b_1]}\left(\int_{[a_2,b_2]} f_x(y)\,\mathrm{d}y\right)\mathrm{d}x \overset{?}{=} \int_{[a_2,b_2]}\left(\int_{[a_1,b_1]} f_y(x)\,\mathrm{d}x\right)\mathrm{d}y \overset{?}{=} \int_{B^2} f \tag{17.3}$$

In this context, we observe that a partition $P \in \mathscr{P}(B^2)$ can be decomposed into the partitions of the individual coordinates. Specifically, we have $P = P_1 \times P_2$ for the two coordinate intervals $[a_1, b_1]$ and $[a_2, b_2]$, and the corresponding indexing sets satisfy $\Lambda(P) = \Lambda(P_1) \times \Lambda(P_2)$.

Now consider the following example,

Example 17.2.2 (A discrepancy: Function integrable, slices not)

Let $I = [0, 1]$ and $B^2 = I \times I$. Consider the function $f : B^2 \to \mathbb{R}$ given by,

$$f(x,y) = \begin{cases} 1 & \text{if } x = \frac{1}{2}, y \in \mathbb{Q} \cap [0,1] \\ 0 & \text{otherwise} \end{cases}$$

So the $x-$slice becomes,

$$f_x \equiv 0 \text{ for all } x \neq \frac{1}{2} \qquad \text{and} \qquad f_{\frac{1}{2}}(y) = \underbrace{\begin{cases} 1 & \text{if } y \in \mathbb{Q}^c \cap [0,1] \\ 0 & \text{if } y \in \mathbb{Q} \cap [0,1] \end{cases}}_{\text{Dirichlet Function}}$$

and $y-$slice,

$$\text{For } y \in \mathbb{Q} \cap [0,1],\ f_y(x) = \begin{cases} 1 & \text{if } x = \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$
$$\text{For } y \in \mathbb{Q}^c \cap [0,1],\ f_y \equiv 0$$

Clearly, $f_x \in \mathscr{R}(I)$ for all $x \in \frac{1}{2}$ but $f_{\frac{1}{2}} \notin \mathscr{R}(I)$ and $f_y \in \mathscr{R}(I)$ for every y. So,

$$\int_I f_y = 0 \implies \int_I \left(\int_I f_y(x)\, \mathrm{d}x \right) \mathrm{d}y = 0$$

But, $\int_I f_x$ doesn't exist for $x = \frac{1}{2}$ which means $x \longmapsto \int_I f_x$ is not a well-defined function on $[0, 1]$. Hence, $\int_I \left(\int_I f_x(y)\, \mathrm{d}y \right) \mathrm{d}x$ doesn't exist. Yet $f \in \mathscr{R}(B^2)$. To see this, we fix $\epsilon > 0$ and consider the partition $P = P_1 \times P_2$ where,

$$\begin{cases} P_1 : 0 < \frac{1}{2} - \epsilon < \frac{1}{2} + \epsilon < 1 \\ P_2 : 0 < 1 \end{cases}$$

$$\text{So, } P = \Bigg\{ \underbrace{\left[0, \frac{1}{2} - \epsilon\right] \times I}_{B_{\alpha_1}}, \underbrace{\left[\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon\right] \times I}_{B_{\alpha_2}}, \underbrace{\left[\frac{1}{2} - \epsilon, 1\right] \times I}_{B_{\alpha_3}} \Bigg\}.$$

Then $m_{\alpha_1} = m_{\alpha_2} = m_{\alpha_3} = 0$, $M_{\alpha_1} = M_{\alpha_3} = 0$ and $M_{\alpha_2} = 1$, which implies $U(f, P) - L(f, P) = 2\epsilon < 3\epsilon$. This shows that $f \in \mathscr{R}(B^2)$. Again $L(f, P) = 0$ for all $P \in \mathscr{P}(B^2)$ and hence,

$$\int_{B^2} f = 0$$

Question. Under which conditions does (17.3) hold?
Answer. Fubini's Theorem. The conditions for (17.3) to hold will be discussed in the next lecture.

Lecture 18

18.1 Fubini's Theorem

In this lecture, we explore Fubini's Theorem. Let's begin by setting up the necessary framework.

Consider the $(m+n)$-dimensional space, where m and n are positive integers. We can decompose a box $B^{m+n} \subseteq \mathbb{R}^{m+n}$ as the Cartesian product of two boxes, $B^{m+n} = B^m \times B^n$. Here, B^m represents a box in $\mathbb{R}^m$, and B^n represents a box in $\mathbb{R}^n$.

Now, suppose we have a partition $P \in \mathscr{P}(\mathbb{R}^{m+n})$. We can express this partition as the Cartesian product of two partitions, $P = P^m \times P^n$, where $P^m \in \mathscr{P}(\mathbb{R}^m)$ and $P^n \in \mathscr{P}(\mathbb{R}^n)$. The corresponding indexing set for this partition becomes $\Lambda(P) = \Lambda(P^m) \times \Lambda(P^n)$. Consequently, the elements of $\Lambda(P)$ can be written as $\alpha(P) = (\alpha(P^m), \alpha(P^n))$.

By extending this decomposition, we can also break down the elements of the boxes, $B_{\alpha(P)} = B_{\alpha(P^m)} \times B_{\alpha(P^n)}$.

Throughout this section, we take $x \in B^m$ and $y \in B^n$ to represent the point $(x, y) \in B^{m+n}$. For a bounded function $f \in \mathscr{B}(B^{m+n})$, we define the slice functions

- $f_x : B^n \to \mathbb{R}$ as $y \mapsto f(x, y)$ for all $x \in B^m$.
- $f_y : B^m \to \mathbb{R}$ as $x \mapsto f(x, y)$ for all $y \in B^n$.

It is worth noting that $f_x \in \mathscr{B}(B^n)$ and $f_y \in \mathscr{B}(B^m)$. For a fixed $x \in B^m$, we can compute the lower and upper integrals of f_x over B^n, denoted as $\underline{f}(x)$ and $\overline{f}(x)$ respectively. Similarly, we can compute the lower and upper integrals of f_y over B^m for fixed $y \in B^m$. These are given by,

$$\underline{f}(x) = \underline{\int}_{B^n} f_x(y)\ \mathrm{d}V(y) \quad \text{and} \quad \overline{f}(x) = \overline{\int}_{B^n} f_x(y)\ \mathrm{d}V(y)$$

with similar expressions for y. Now, let's state Fubini's Theorem.

Theorem 18.1.1 (Fubini's Theorem)

Let $f \in \mathscr{R}(B^{m+n})$. Then $\underline{f}, \overline{f} \in \mathscr{R}(B^n)$ and,

$$\int_{B^m} \underline{f} = \int_{B^m} \overline{f} = \int_{B^{m+n}} f$$

Consequently, we have the following corollaries,

Corollary

For any $f \in \mathscr{R}(B^{m+n})$, the following equalities hold,

$$\begin{aligned}\int_{B^m}\left(\underline{\int}_{B^n} f(x,y)\ \mathrm{d}V(y)\right)\mathrm{d}V(x) &= \int_{B^m}\left(\overline{\int}_{B^n} f(x,y)\ \mathrm{d}V(x)\right)\mathrm{d}V(y)\\ &= \int_{B^{m+n}} f(x,y)\ \mathrm{d}V(x,y)\end{aligned}$$

Furthermore, if $f_x \in \mathscr{R}(B^n)$ for all x, then $\underline{f} = \overline{f}$ and

$$\int_{B^m}\left(\int_{B^n} f(x,y)\ \mathrm{d}V(y)\right)\mathrm{d}V(x) = \int_{B^{m+n}} f(x,y)\ \mathrm{d}V(x,y)$$

Corollary

If $f \in C(B^n)$, then all possible slice functions are continuous and hence Riemann Integrable. Thus, multidimensional integral becomes the iterated one-dimensional integrals,

$$\int_{B^n} f = \int\left(\int\cdots\int\left(\int f\ \mathrm{d}x_1\right)\mathrm{d}x_2\cdots\mathrm{d}x_{n-1}\right)\mathrm{d}x_n$$

where, x_i's can be in any order.

($n = 2$) Hence, if $f \in C(B^2)$, then (17.3) holds,

$$\int_{B^2} f = \int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} f(x,y)\ \mathrm{d}y\right)\mathrm{d}x = \int_{a_2}^{b_2}\left(\int_{a_1}^{b_1} f(x,y)\ \mathrm{d}x\right)\mathrm{d}y$$

Proof of the Fubini's Theorem. Let $P = P^m \times P^n$ be a partition of B^{m+n}. Then, we can express the lower sum $L(f,P)$ as follows,

$$\begin{aligned}L(f,P) &= \sum_{\alpha(P)\in\Lambda(P)} m_{\alpha(P)}\,\mathrm{Vol}\left(B^{m+n}_{\alpha(P)}\right)\\ &= \sum_{\alpha(P^m)\in\Lambda(P^m)} \underbrace{\sum_{\alpha(P^n)\in\Lambda(P^n)} m_{(\alpha(P^m),\alpha(P^n))}\,\mathrm{Vol}\left(B^n_{\alpha(P^n)}\right)}_{l_{P^m}}\mathrm{Vol}\left(B^m_{\alpha(P)^m}\right)\end{aligned}$$

For each $x \in B^m$ and $\alpha(P^n) \in \Lambda(P^n)$, let $m_{\alpha(P^n)}(x) = \inf_{y\in B^n_{\alpha(P^n)}} f_x(y)$. It follows that $m_{\alpha(P^n)}(x) \geq m_{(\alpha(P^m),\alpha(P^n))}$ for every $x \in B^m_{\alpha(P^m)}$. Consequently, we have,

$$\begin{aligned}l_{P^m} &\leq \sum_{\alpha(P^n)\in\Lambda(P^n)} m_{\alpha(P^n)}\,\mathrm{Vol}\left(B^n_{\alpha(P^n)}\right)\\ &= L(f_x, P^n) \leq \underline{\int}_{B^n} f_x\end{aligned}$$

Taking infimum over all $x \in B^m_{\alpha(P^m)}$, we obtain,

$$l_{P^m} \leq \inf_{x\in B^m_{\alpha(P^m)}} \underline{\int}_{B^n} f_x$$

$$= \inf_{x \in B^m_{\alpha(P^m)}} \underline{f}(x) = \underline{m}_{\alpha(P^m)}$$

Thus, the lower sum becomes,

$$L(f, P) \leq \sum_{\alpha(P^m) \in \Lambda(P^m)} \underline{m}_{\alpha(P^m)} \operatorname{Vol}\left(B^m_{\alpha(P^m)}\right) = L(\underline{f}, P^m)$$

Similarly, we can show that $U(f, P) \geq U(\underline{f}, P^m)$. Consequently, $\underline{f} \in \mathscr{R}(B^m)$, and we have,

$$\int_{B^m} \underline{f} \; \mathrm{d}V(x) = \int_{B^{m+n}} f$$

By following analogous arguments, we can show that $\overline{f} \in \mathscr{R}(B^m)$ and,

$$\int_{B^m} \overline{f} \; \mathrm{d}V(x) = \int_{B^{m+n}} f = \int_{B^m} \underline{f} \; \mathrm{d}V(x)$$

□

Question: Will the function be Riemann integrable if all the slices are Riemann integrable? We will address this question later. In the meantime, let's conclude this lecture with an example.

Example 18.1.1

Consider the integral

$$\int_{[0,1]^2} xy \; \underbrace{\mathrm{d}x \, \mathrm{d}y}_{\mathrm{d}v}$$

We can evaluate this integral by iterated integration as follows,

$$\begin{aligned} \int_0^1 \left(\int_0^1 xy \; \mathrm{d}x \right) \mathrm{d}y &= \int_0^1 y \left(\int_0^1 x \; \mathrm{d}x \right) \mathrm{d}y \\ &= \int_0^1 \frac{y}{2} \; \mathrm{d}y = \frac{1}{4} \end{aligned}$$

Alternatively, we can also express it as, $\int_0^1 y \left(\int_0^1 x \; \mathrm{d}x \right) \mathrm{d}y = \left(\int_0^1 x \; \mathrm{d}x \right) \left(\int_0^1 y \; \mathrm{d}y \right)$, which yields the same result.

Lecture 19

19.1 Integration over Bounded Domain

Now that we know how to do integration over boxes, in this lecture we will discuss how to integrate a bounded function over an arbitrary bounded set.

Let $\Omega \subseteq \mathbb{R}^n$, and $f \in \mathscr{B}(\Omega)$ and Ω bounded, then there exists $B^n \supseteq \Omega$ where $B^n = \prod_{i=1}^n [a_i, b_i]$.

Definition 19.1.1

Given Ω bounded, let $B^n \supseteq \Omega$. For $f \in \mathscr{B}(\Omega)$, define

$$\tilde{f}_{B^n}(x) = \begin{cases} f(x) & \text{if } x \in \Omega \\ 0 & \text{if } x \in B^n \setminus \Omega \end{cases}$$

An immediate question that arises now is: If $B^n \supseteq \Omega$ and $\hat{B^n} \supseteq \Omega$ then will it be true that

(I) $\tilde{f}_{B^n} \in \mathscr{R}(B^n)$

(II) And if (I) holds, is it necessarily true that $\displaystyle\int_{B^n} \tilde{f}_{B^n} = \int_{\hat{B^n}} \tilde{f}_{\hat{B^n}}$.

It is in fact true that (I) $\implies$ (II), but we won't cover the proof here.

Definition 19.1.2

Let $f \in \mathscr{B}(\Omega)$. We say that $f \in \mathscr{R}(\Omega)$ if $\displaystyle\int_{B^n} \tilde{f}_{B^n}$ exists for some $B^n \supseteq \Omega$, and in this case we define

$$\int_\Omega f := \int_{B^n} \tilde{f}_{B^n}$$

Definition 19.1.3 ▶ Content Zero Sets

Let $S \subseteq \mathbb{R}^n$, we say that S is of content zero if for all $\varepsilon > 0$, there exists boxes $\{B_j^n\}_{j=1}^p$ (for some $p \in \mathbb{N}$) such that

$$S \subseteq \bigcup_{j=1}^p B_j^n \quad \text{and} \quad \sum_{j=1}^n \text{Vol}(B_j^n) < \varepsilon$$

For example a line segment in $\mathbb{R}^n$ is of content zero, provided $n > 1$. We then have the following theorems:

Theorem 19.1.1

(i) Let $f \in \mathscr{B}(B^n)$ and let $\mathcal{D} = \{x \in B^n \mid f \text{ is not continuous at } x\}$ be the set of discontinuities of f, if $\mathcal{D}$ is of content zero, then $f \in \mathscr{R}(B^n)$.

(ii) If S is a content zero set then $\operatorname{int}(S) = \emptyset$.

(iii) Let $\Omega \subseteq \mathbb{R}^n$ and $\mathcal{O}_n \subseteq \Omega$ is bounded. Let $f \in \mathscr{B}(\Omega)$ and $f|_{\mathcal{O}_n} \in C(\mathcal{O}_n)$, if $\bar{\Omega} \setminus \mathcal{O}_n$ is content zero then $f \in \mathscr{R}(\Omega)$ and

$$\int_{\Omega} f = \int_{\mathcal{O}_n} f$$

Particularly, (i) and (iii) of Theorem 19.1.1 are very important.

Now that we know how to integrate on arbitrary domains, the next question that comes to our mind is, does there exist a Fubini's theorem for integration over arbitrary sets? Before that we define elementary regions.

19.2 Two Elementary Regions

Definition 19.2.1 ▶ Elementary Regions

A set $\Omega \subseteq \mathbb{R}^2$ is y-simple/type I if there exists functions $\varphi_1, \varphi_2 \in \mathscr{B}([a,b])$ such that

$$\Omega = \{(x,y) \mid x \in [a,b],\ y \in [\varphi_1(x), \varphi_2(x)]\}$$

Similarly a set $\Omega \subseteq \mathbb{R}^2$ is x-simple/type II if there exists functions $\psi_1, \psi_2 \in \mathscr{B}([c,d])$ such that

$$\Omega = \{(x,y) \mid y \in [c,d],\ x \in [\psi_1(x), \psi_2(x)]\}$$

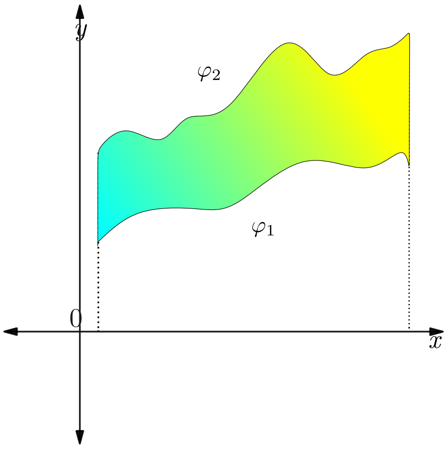


Figure 19.1: Example of a y-simple region.

Example 19.2.1 (Examples of elementary regions)

The region H given by

$$H = \{(x,y) \mid 0 \le x \le 1, \text{ and } x^2 \le y \le x\}$$

is a y-simple region. (see Figure 19.1)

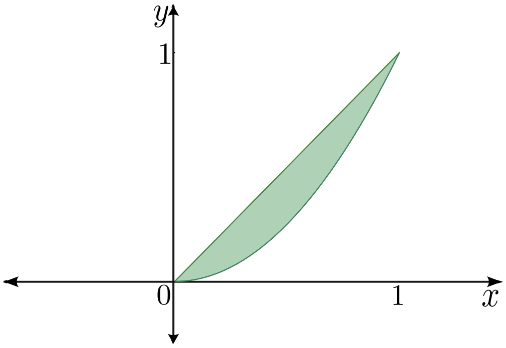


Figure 19.2: Plot of the region H.

Exercise. Show that the region bounded by $x^2 + y^2 \leq 1$ and $y \geq 0$ in $\mathbb{R}^2$ is an x-simple as well as a y-simple region.

Lecture 20

In the previous lecture we extended integration over boxes to over what we called elementary regions. This lecture explores the extention of Fubini's theorem to integration over such sets, and talk about some applications thereof. Towards the end, we discuss the celebrated change of variables formula of multivariable calculus.

20.1 Fubini's Theorem on Elementary Regions

For integration over elementary regions, Fubini's theorem takes the following form.

Theorem 20.1.1

Let $f \in \mathscr{R}(\Omega)$ where $\Omega \subseteq \mathbb{R}^2$ is a bounded elementary domain.

(1) If $\Omega = \{(x,y) \mid a \le x \le b, \text{ and } \varphi_1(x) \le y \le \varphi_2(x)\}$ and if $\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y$ exists for all $x \in [a,b]$ then

$$\int_\Omega f\,\mathrm{d}A = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

(2) Similarly we have

$$\int_\Omega f\,\mathrm{d}A = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y$$

when Ω is x-simple.

Proof. There exists $c, d \in \mathbb{R}$ such that $\Omega \subseteq [a,b] \times [c,d] = B^2$. We know $\tilde{f} \in \mathscr{R}(B^2)$ where

$$\tilde{f}(x,y) = \begin{cases} f(x,y) & \text{if } (x,y) \in \Omega \\ 0 & \text{if } (x,y) \in B^2 \setminus \Omega \end{cases}$$

Since $\tilde{f} \in \mathscr{R}(B^2)$, and since $\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y$ exists for fixed x, hence

$$\tilde{f}(x,\cdot)\big|_{[\varphi_1(x),\varphi_2(x)]} \quad \text{and} \quad \tilde{f}(x,\cdot)\big|_{[c,d]\setminus[\varphi_1(x),\varphi_2(x)]} \equiv 0$$

are both Riemann integrable. Thus, we get that $\tilde{f}(x,\cdot)|_{[c,d]} \in \mathscr{R}([c,d])$ and hence $\int_c^d \tilde{f}(x,y)\,\mathrm{d}y$ exists for all $x \in [a,b]$ and further we have

$$\int_c^d \tilde{f}(x,y)\,\mathrm{d}y = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y \quad \forall x \in [a,b]$$

Then we get

$$\begin{aligned}
&\Longrightarrow \int_a^b \left(\int_c^d \tilde{f}(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \\
&\overset{(*)}{\Longrightarrow} \int_{B^2} \tilde{f}\,\mathrm{d}A = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \\
&\Longrightarrow \int_\Omega f\,\mathrm{d}A = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x
\end{aligned}$$

where $(*)$ follows from Theorem 18.1.1. An analogous argument works for the x-simple regions. ☐

Example 20.1.1

Let $f \in C(\Omega)$ where $\Omega = \left\{(x,y) \mid 0 \le x \le 1 - \frac{y}{2}, \text{ and } 0 \le y \le 2\right\}$. Then we can write Ω as a y-simple region as follows:

$$\Omega = \{(x,y) \mid 0 \le x \le 1, \text{ and } 0 \le y \le 2(1-x)\}\,.$$

Now, using Theorem 20.1.1 we get the required result.

$$\int_\Omega f\,\mathrm{d}A = \int_0^2 \left(\int_0^{1-\frac{y}{2}} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int_0^1 \left(\int_0^{2(1-x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

Example 20.1.2

Let $B^2 = [0,\pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$, and we want to evaluate the integral $\displaystyle\int_{B^2} \sin(x+y)\,\mathrm{d}A$.

$$\begin{aligned}
\int_{B^2} \sin(x+y)\,\mathrm{d}A &= \int_{B^2} \sin x \cos y\,\mathrm{d}A + \int_{B^2} \sin y \cos x\,\mathrm{d}A \\
&= \left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos y\,\mathrm{d}y \right) \left(\int_0^\pi \sin x\,\mathrm{d}x \right) + \left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin y\,\mathrm{d}y \right) \left(\int_0^\pi \cos x\,\mathrm{d}x \right) \\
&= 4
\end{aligned}$$

Example 20.1.3

Let Ω be the region bounded by $y = 1$ and $y = x^2$, and we want to find $\displaystyle\int_\Omega x^2 y\,\mathrm{d}V$. We can write Ω as a y-simple region as follows:

$$\Omega = \{(x,y) \mid -1 \le x \le 1, \text{ and } x^2 \le y \le 1\}$$

Then using Theorem 20.1.1 we get that

$$\begin{aligned}
\int_\Omega x^2 y\,\mathrm{d}A &= \int_{-1}^1 \left(\int_{x^2}^1 x^2 y\,\mathrm{d}y \right) \mathrm{d}x \\
&= \int_{-1}^1 x^2 \left(\frac{y^2}{2} \right) \Bigg|_{x^2}^1 \mathrm{d}x = \frac{2}{15}
\end{aligned}$$

Example 20.1.4

Compute $\int_{[0,1]^2} f \, \mathrm{d}A$ where $f : [0,1]^2 \to \mathbb{R}$ is given by

$$f(x,y) = \begin{cases} x & \text{if } y \leq x^2 \\ y & \text{if } y > x^2 \end{cases} \quad \forall\, (x,y) \in [0,1]^2$$

We define the regions

$$\Omega_1 = \{(x,y) \mid 0 \leq x \leq 1,\, 0 \leq y \leq x^2\} \quad \text{and} \quad \Omega_2 = \{(x,y) \mid 0 \leq x \leq 1,\, x^2 \leq y \leq 1\}.$$

Then $f|_{\Omega_1}$ and $f|_{\Omega_2}$ are both Riemann integrable, and while $f|_{y=x^2}$ is not continuous, the set $\{(x,x^2) \mid 0 \leq x \leq 1\}$ is of content zero. Hence f is integrable, and we can make sense of writing the given integral as a sum

$$\int_{[0,1]^2} f \, \mathrm{d}A = \int_{\Omega_1} f \, \mathrm{d}A + \int_{\Omega_2} f \, \mathrm{d}A$$

Using Theorem 20.1.1, we can simplify each of these parts:

$$\int_{\Omega_1} f \, \mathrm{d}A = \int_0^1 \left(\int_0^{x^2} x \, \mathrm{d}y \right) \mathrm{d}x = \frac{1}{4}$$

$$\int_{\Omega_2} f \, \mathrm{d}A = \int_0^1 \left(\int_{x^2}^1 y \, \mathrm{d}y \right) \mathrm{d}x = \frac{2}{5}$$

Thus, we have the result

$$\int_{[0,1]^2} f \, \mathrm{d}A = \frac{13}{20}$$

Example 20.1.5

Compute $\int_0^{\frac{\pi}{2}} \int_x^{\frac{\pi}{2}} \frac{\sin y}{y} \, \mathrm{d}y \, \mathrm{d}x$ using Fubini's theorem.

We consider the region $\Omega = \left\{(x,y) \mid 0 \leq x \leq \frac{\pi}{2},\, x \leq y \leq \frac{\pi}{2}\right\} \subseteq \mathbb{R}^2$ can be written as a x-simple region as follows:

$$\Omega = \left\{(x,y) \mid 0 \leq y \leq \frac{\pi}{2},\, 0 \leq x \leq y\right\}$$

This shows that

$$\begin{aligned} \int_\Omega \frac{\sin y}{y} \, \mathrm{d}A &= \int_0^{\frac{\pi}{2}} \int_x^{\frac{\pi}{2}} \frac{\sin y}{y} \, \mathrm{d}y \, \mathrm{d}x \\ &= \int_0^{\frac{\pi}{2}} \int_0^{y} \frac{\sin y}{y} \, \mathrm{d}x \, \mathrm{d}y \\ &= \int_0^{\frac{\pi}{2}} \sin y \, \mathrm{d}y \\ &= 1 \end{aligned}$$

20.2 Change of Variables

Before discussing the theorem in the multivariable case, we recall the change of variable rule for real valued functions on the real line.

Theorem 20.2.1 (Change of Variable on $\mathbb{R}$)

Let $\varphi : \mathcal{O}_1 \to \mathbb{R}$ be a C^1 function where $\varphi'(x) \neq 0$ for all $x \in \mathcal{O}_1$. Then, for $[a, b] \subseteq \mathcal{O}_1$ and $f \in C(\varphi[a, b])$, we have

$$\int_{\varphi(a)}^{\varphi(b)} f = \int_a^b (f \circ \varphi)\varphi'.$$

Here we effectively compensate for the change of variable by introducing the scale change factor of φ'. As we have seen, the scale change factor at a point for a transformation on $\mathbb{R}^n$ is given by the detereminant of the Jacobian matrix at that point. Thus, this theorem has the following natural extension to $\mathbb{R}^n$:

Theorem 20.2.2 (Change of Variable on $\mathbb{R}^n$)

Let $\varphi : \mathcal{O}_n \to \mathbb{R}^n$ be an injective and C^1 function, where $\det(J_\varphi(x)) \neq 0$ for all $x \in \mathcal{O}_n$. Let $\Omega \subseteq \mathcal{O}_n$, then for $f \in \mathscr{R}(\varphi(\Omega))$

$$\int_{\varphi(\Omega)} f \, \mathrm{d}V = \int_\Omega (f \circ \varphi) |\det J_\varphi|$$

Although it is not too hard to get a feel for the theorem from its applications, the proof is quite long and technical, and thus omitted. We recommend the interested and daring readers to have a look at page 67 of *Calculus on Manifolds* by Michael Spivak. In the next lecture, we will discuss some applications of this result.

Lecture 21

In the previous lecture, we extended our notion of Riemann integration over boxes to elementary regions. Additionally, we had discussed the change of variable formula for multivariable calculus (Theorem 20.2.2). We now try to motivate the use of the same with some rather important applications.

21.1 Change of Variables (Continued)

We start off with a particularly useful example. When dealing with functions on $\mathbb{R}^2$, particularly if the situation is radially symmetric, it is often useful to work in polar coordinates. Here we analyse how that change of coordinates transforms the integrals over a given region.

Example 21.1.1 (Polar coordinates)

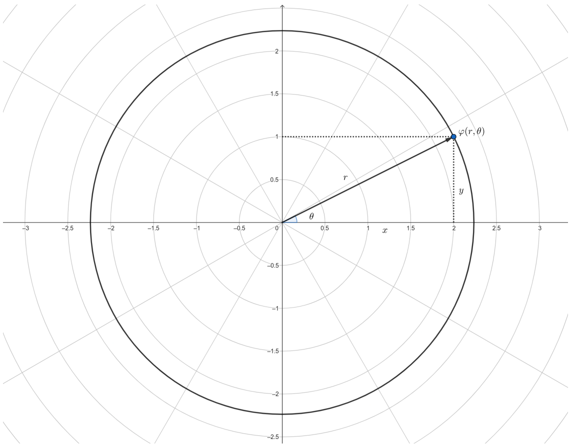


Figure 21.1: Transforming into polar coordinates

This example illustrates how we can compute integrals when converting to polar coordinates from Cartesian coordinates.

Consider $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ given by $\varphi(r,\theta) = (r\cos\theta, r\sin\theta)$. Then, the Jacobian matrix of φ is given by

$$J_\varphi(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Thus, $\det(J_\varphi(r,\theta)) = r \neq 0$ for all $r > 0$. Taking the domain to be $\mathcal{O}_2 := (0,\infty) \times (0, 2\pi)$, the function $\varphi|_{\mathcal{O}_2} : \mathcal{O}_2 \to \mathbb{R}^2$ is C^1 and injective, with $\det(J_\varphi(r,\theta)) \neq 0$. For some fixed $0 < r_1 < r_2$ and $0 \leq \theta_1 < \theta_2 < 2\pi$, consider the set $\Omega = (r_1, r_2) \times (\theta_1, \theta_2)$. Clearly the boundary $\partial\Omega$ is of content zero (since it is union of finitely many line segments) and thus using Theorem 20.2.2 we see that for $f \in \mathscr{R}(\varphi(\Omega))$,

$$\begin{aligned} \int_{\varphi(\Omega)} f &= \int_\Omega (f \circ \varphi)|\det(J_\varphi(r,\theta))| \\ &= \int_\Omega r f(r\cos\theta, r\sin\theta) \\ &= \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} r f(r\cos\theta, r\sin\theta)\, \mathrm{d}\theta\, \mathrm{d}r \end{aligned}$$

As a simple example, consider the following integral:

$$\int_{x^2+y^2<1} e^{-(x^2+y^2)} \mathrm{d}A$$

The domain of integration can be written as

$$\{(x,y) \mid x^2+y^2<1\} = \varphi((0,1)\times[0,2\pi))$$

Clearly $f(x,y)=e^{-(x^2+y^2)} \in C^1(\varphi((0,1)\times[0,2\pi)))$ which shows that:

$$\begin{aligned}\int_{x^2+y^2<1} e^{-(x^2+y^2)}\,\mathrm{d}A &= \int_0^1 \left(\int_0^{2\pi} re^{-r^2}\,\mathrm{d}\theta\right)\mathrm{d}r\\ &= 2\pi\int_0^1 re^{-r^2}\,\mathrm{d}r\\ &= \pi\left(1-\frac{1}{e}\right)\end{aligned}$$

Definition 21.1.1 ▶ Area or volume of a region

For $\Omega \subseteq \mathbb{R}^n$, the volume of the region Ω is defined by the integral

$$\int_{\mathbb{R}^n} \chi_\Omega$$

if it exists, where χ_Ω is the indicator function of the region Ω, given by

$$\chi_\Omega = \begin{cases} 1 & \text{if } x\in\Omega,\\ 0 & \text{otherwise.}\end{cases}$$

Example 21.1.2

Compute the area of $\Omega = \left\{(x,y) \mid x^{\frac{2}{3}}+y^{\frac{2}{3}}<1\right\}$. Consider the function

$$\begin{aligned}\varphi : [0,1]\times[0,2\pi] &\to \mathbb{R}^2\\ (r,\theta) &\mapsto (r\cos^3\theta, r\sin^3\theta)\end{aligned}$$

then clearly $\varphi([0,1]\times[0,2\pi])=\Omega$. Also φ is injective and C^1, but we have

$$\det(J_\varphi(r,\theta)) = 3r\sin^2\theta\cos^2\theta$$

and thus $\det(J_\varphi(r,\theta))=0$ if $r=0$ or $\theta\in\left\{0,\frac{\pi}{2},\pi,\frac{3\pi}{2},2\pi\right\}$, but set of all such points are of content zero, hence we can safely ignore them while doing our integration. We get

$$\begin{aligned}\text{Area of }\Omega = &\ \text{Area of }\varphi\left([0,1]\times[0,2\pi]\right)\\ = &\int_{\varphi([0,1]\times[0,2\pi])} 1\,\mathrm{d}A\\ = &\int_0^1\int_0^{2\pi}\frac{3}{4}r\sin^2 2\theta\,\mathrm{d}\theta\,\mathrm{d}r = \frac{3\pi}{8}\end{aligned}$$

Next up, consider a change of coordinates in $\mathbb{R}^3$, from Cartesian system to the spherical co-ordinate system. Just like the previous case, this comes in handy when dealing with functions and regions which are spherically symmetric. One canonical example may be its use in the theory of central forces in physics.

Example 21.1.3 (Spherical coordinates)

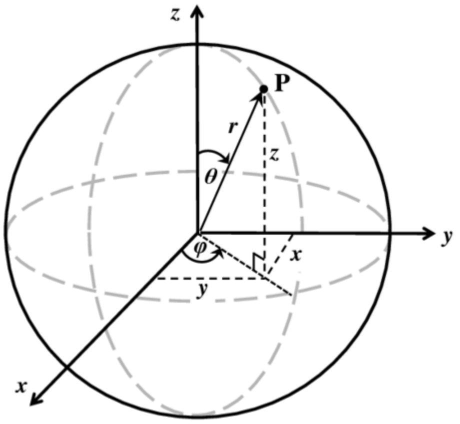


Figure 21.2: Transforming into spherical coordinates

The spherical co-ordinate system gives a unique representation to all points in $\mathbb{R}^3$ not lying on the z-axis. For all $(x, y, z) \in \mathbb{R}^3 \setminus \{(0, 0, \alpha) \mid \alpha \in \mathbb{R}\}$, set

$$(x, y, z) = (r \sin\phi \cos\theta, r \sin\phi \sin\theta, r \cos\phi)$$

where $r > 0$, $0 < \theta < 2\pi$ and $0 < \phi < \pi$. Define the set

$$\mathcal{O}_3 := \{(r, \phi, \theta) \mid r > 0,\ 0 < \phi < \pi \text{ and } 0 < \theta < 2\pi\}$$

and the map

$$\varphi : \mathcal{O}_3 \to \mathbb{R}^3,\ \ \varphi(r, \phi, \theta) = (x, y, z)$$

Then, the Jacobian matrix of the map φ is:

$$J_\varphi(r, \phi, \theta) = \begin{pmatrix} \sin\phi\cos\theta & r\cos\phi\cos\theta & -r\sin\phi\sin\theta \\ \sin\phi\sin\theta & r\cos\phi\sin\theta & r\sin\phi\cos\theta \\ \cos\phi & -r\sin\phi & 0 \end{pmatrix}$$

This gives $\det(J_\varphi(r, \phi, \theta)) = r^2 \sin\phi$, which is non-vanishing in the given domain. As φ is injective and C^1, we can use Theorem 20.2.2 to transform from Cartesian to spherical coordinates. We provide a simple example for some clarity.

Consider the solid sphere of radius a, given by $\Omega = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq a^2\}$. Then, the volume is:

$$\begin{aligned} \mathrm{Vol}(\Omega) &= \int_0^{2\pi} \int_0^{\pi} \int_0^{a} r^2 \sin\phi \, \mathrm{d}r \, \mathrm{d}\phi \, \mathrm{d}\theta \\ &= \frac{4}{3}\pi a^3 \end{aligned}$$

We leave it to the reader to do a similar analysis for the cylindrical co-ordinate system. This formula also finds extensive use in probability theory, where it is commonly referred to as the change of density formula (see, for instance *A First Course in Probability* by Sheldon Ross, or any introductory probability book for that matter). Hopefully, we have demostrated to the reader the central role Theorem 20.2.2 plays in analysis of several variables, enough to convince him to actually read the proof! We will now depart from the general study of functions and integrals in $\mathbb{R}^n$, and delve into the theory of curves and surfaces, dealing primarily with $\mathbb{R}^3$.

Lecture 22

22.1 Curves and Surfaces

We now study the basic concepts of curves and surfaces as subsets of $\mathbb{R}^2$ or $\mathbb{R}^3$ (mainly) with a given parametrization, but also as subsets defined by equations. The connection from equations to parametrization is drawn by means of the Implicit function theorem.

Definition 22.1.1 ▶ Parametrized Curve

A **parametrized curve** is a continuous function $\gamma : [a, b] \to \mathbb{R}^n$. We say that parametrized curve is C^1 if $t \mapsto \gamma_i(t)$ is C^1 for all $i = 1, \ldots, n$. A parametrized curve $\gamma : I \to \mathbb{R}^n$ is **smooth** if $\gamma'(t) \neq \mathbf{0}$ for all $t \in I$. The **path** of a parametrized curve γ is the set

$$\{\gamma(t) \mid t \in [a, b]\}$$

Let us consider some examples.

Example 22.1.1

1. Let $\gamma : [0, 1] \to \mathbb{R}^2$ be $\gamma(t) = (1 - 2t, 2 + t)$. Clearly γ is C^1 and $\gamma'(t) = (-2, 1) \neq \mathbf{0}$ and thus γ is smooth.
2. $\gamma : [0, 2\pi] \to \mathbb{R}^2$ given by $\gamma : t \mapsto (r \cos t, r \sin t)$ where $r > 0$ is constant. Then $\gamma'(t) = (-r \sin t, r \cos t) \neq \mathbf{0} \,\forall\, t \in [0, 2\pi]$, thus γ is smooth.
3. Fix $r > 0$ and $c \neq 0$, and define

   $$\begin{aligned} \gamma : [0, n\pi] &\to \mathbb{R}^3 \\ t &\mapsto (r \cos t, r \sin t, ct) \end{aligned}$$

   Then $\gamma'(t) = (-r \sin t, r \cos t, c) \neq \mathbf{0}$ and hence γ is smooth. The path of γ is called a **helix**.
4. $\gamma : [-1, 1] \to \mathbb{R}^2$ given by $\gamma(t) = (|t|, t)$, then γ is not C^1, hence it is not smooth.
5. $\gamma : [0, 1] \to \mathbb{R}^2$ given by $\gamma(t) = (0, t^2)$, then even though γ is C^1, but it is not smooth as $\gamma'(0) = \mathbf{0}$.
6. $\gamma : [0, 2\pi] \to \mathbb{R}^2$ given by $\gamma : t \mapsto (r \cos t, r \sin t)$, then path of γ is given by

   $$\begin{aligned} \{(r \cos t, r \sin t) \mid t \in [0, 2\pi]\} &= \{(x, y) \mid x^2 + y^2 = r^2\} \\ &= \text{path of } \tilde{\gamma} \end{aligned}$$

   where $\tilde{\gamma}(t) = (r \cos 2t, r \sin 2t)$.

Definition 22.1.2 ▶ Piecewise Smooth Curve

A parametrized curve $\gamma : [a,b] \to \mathbb{R}^n$ is called piecewise smooth if there exists a partition $a = t_0 < t_1 < \cdots < t_m = b$ such that

$$\gamma|_{[t_{i-1},t_i]} \text{ is smooth } \forall\, i \in [m]$$

Definition 22.1.3 ▶ Equivalent Curves

Two parametrized curves $\gamma : [a,b] \to \mathbb{R}^n$ and $\tilde{\gamma} : [\tilde{a},\tilde{b}] \to \mathbb{R}^n$ are equivalent, denoted by $\gamma \sim \tilde{\gamma}$ if there exists a strictly increasing surjective function which is differentiable (even C^1), $\varphi : [\tilde{a},\tilde{b}] \to [a,b]$ such that $\tilde{\gamma} = \gamma \circ \varphi$.

$$\begin{tikzcd}
{[\tilde{a},\tilde{b}]} \arrow[rr, "\tilde{\gamma}"] \arrow[dr, "\varphi"] & & \mathbb{R}^n \\
& {[a,b]} \arrow[ur, "\gamma"] &
\end{tikzcd}$$

Definition 22.1.4

Let $\gamma : [a,b] \to \mathbb{R}^n$ be a C^1 curve, then

(i) $\|\gamma'(t)\| :=$ speed of γ at time t.

(ii) $\displaystyle\int_a^b \|\gamma'(t)\| \mathrm{d}t :=$ arc length of γ.

Let's try to look at more natural how equation (ii) in the previous definition gives us the arc length of a curve γ.

Let $\gamma : [a,b] \to \mathbb{R}^n$ and let $\mathcal{P} := a = t_0 < t_1 < \cdots < t_m = b$ be a partition of the interval $[a,b]$. Now define

$$\ell(\gamma, \mathcal{P}) = \sum_{i=1}^{m} \|\gamma(t_{i-1}) - \gamma(t_i)\|$$

Definition 22.1.5

A curve $\gamma : [a,b] \to \mathbb{R}^n$ is **rectifiable** or said to have arc length if

$$\lim_{\substack{\|\mathcal{P}\| \to 0 \\ \mathcal{P} \in \mathscr{P}[a,b]}} \ell(\gamma, \mathcal{P}) = \ell(\gamma) \text{ exists}$$

which is equivalent to saying that for all $\varepsilon > 0$, there exists $\delta > 0$, such that

$$\|\ell(\gamma, \mathcal{P}) - \ell(\gamma)\| < \varepsilon \quad \forall\, \mathcal{P} \in \mathscr{P}([a,b]) \text{ such that } \|\mathcal{P}\| < \delta$$

Theorem 22.1.1

For a piecewise smooth curve $\gamma : [a,b] \to \mathbb{R}^n$ it is rectifiable and $\ell(\gamma) = \displaystyle\int_a^b \|\gamma'(t)\| \mathrm{d}t$.

Remark. Rectifiable curve $\not\Rightarrow$ piecewise smooth, counter example: Cantor's function (popularly called the Devil's staircase).

Theorem 22.1.2

Let $\gamma : [a, b] \to \mathbb{R}^n$ be a rectifiable parametrized curve and let $\tilde{\gamma} = \gamma \circ \varphi$, where φ is a strictly increasing surjective and continuous function, then $\tilde{\gamma}$ is rectifiable and $\ell(\gamma) = \ell(\tilde{\gamma})$.

Theorem 22.1.3

Let $\gamma : [a, b] \to \mathbb{R}^n$ be a smooth curve, then there exists a parametrization φ such that $\|\tilde{\gamma}'(s)\| = 1$ for all $s \in [c, d]$.

Lecture 23

We will begin this lecture with few examples.

Example 23.0.2 (Path of a Projectile)

$$\begin{aligned} \gamma(t) &= (\alpha t, \beta t - 16t^2) \\ \implies \gamma'(t) &= (\alpha, \beta - 32t) \\ \text{path length} &= \int \|\gamma'(t)\| \, \mathrm{d}t \\ &= \int \sqrt{\alpha^2 + (\beta - 32t)^2} \, \mathrm{d}t \end{aligned}$$

Example 23.0.3 (Perimeter of a Circle)

Parametrization of a circle of radius r is given by $\gamma(t) = (r\cos(t), r\sin(t)), t \in [0, 2\pi)$.

$$\begin{aligned} \gamma'(t) &= (-r\sin(t), r\cos(t)) \\ \ell(\gamma) &= \int_0^{2\pi} \|\gamma'(t)\| \mathrm{d}t \\ &= \int_0^{2\pi} r \, \mathrm{d}t \\ &= 2\pi r \end{aligned}$$

Example 23.0.4 (Arc Length of graph of functions)

Let, $f : [a, b] \to \mathbb{R}$ be a C^1 function. Consider $\gamma(t) = (t, f(t))$. It is a smooth curve.

$$\begin{aligned} \gamma'(t) &= (1, f'(t)) \\ \ell(\gamma) &= \int_a^b \|\gamma'(t)\| \mathrm{d}t \\ &= \int_a^b \sqrt{1 + f'(t)^2} \, \mathrm{d}t \end{aligned}$$

23.1 Line Integrals

To integrate a function over a curve we use **Line integral**. The function we should integrate maybe a **Scalar Field** or a **Vector Field**.(A quick example of a Vector Field: $f : \mathcal{O}_n \to \mathbb{R}$ be a differentiable function, then $\boldsymbol{\nabla} f$ is a vector field.)

Question. Given a scalar field $f : \mathcal{O}_n \to \mathbb{R}$ and $\gamma \equiv \mathcal{C}$ be a curve, we want to define $\int_{\mathcal{C}} f$. But exactly how we can do this?
Answer. $\mathcal{C}$ is a curve, so it is bounded subset of $\mathbb{R}^n$. How about thinking of **Riemann Integration**? For $n \geq 2$, $\mathcal{C}$ is **content zero** in $\mathbb{R}^n$. This does not make any sense! The right way is as following.

Let, $\gamma : [a, b] \to \mathbb{R}^n$ be a **smooth curve** (or piecewise smooth) and $\mathcal{C} := \operatorname{ran}(\gamma)$ (on other words path of γ). Let, $f \in \mathscr{B}(\mathcal{C})$. Given $\mathcal{P} \in \mathscr{P}[a, b]$, $\mathcal{P} : a = t_0 < t_1 < \cdots < t_m = b$.

Let, $I_i = [t_{i-1}, t_i]$ be the sub-intervals and $\mathcal{C}_i = \gamma(I_i)$. Since, γ is smooth there is nice correspondence between I_i and $\mathcal{C}_i$. Also denote s_i by $\|\gamma(t_i) - \gamma(t_i)\|$. As previous, define $m_i = \inf_{\mathcal{C}_i} f$ and $M_i = \sup_{\mathcal{C}_i} f$.

$$U(f, \mathcal{P}) = \sum_{i=1}^{m} M_i \cdot s_i$$
$$L(f, \mathcal{P}) = \sum_{i=1}^{m} m_i \cdot s_i$$

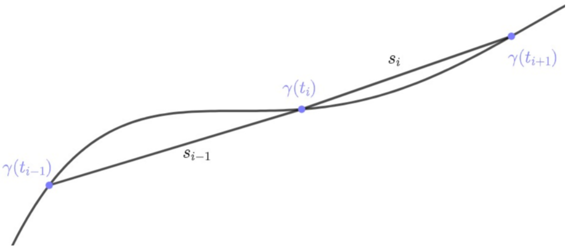


Figure 23.1: Curve $\mathcal{C}$

The above expressions are same as upper and lower Riemann sum respectively. This opens up "The Pandora's box!".

We can now use all the theory we used for the standard Riemann Integrals. We say f is line integrable over γ if,

$$\inf_{\mathcal{P} \in \mathscr{P}[a,b]} U(f, \mathcal{P}) = \sup_{\mathcal{P} \in \mathscr{P}[a,b]} L(f, \mathcal{P})$$

More over we will write the common value of the above equality as $\int_{\mathcal{C}} f$ and call this "The Line Integral over curve $\mathcal{C}$".
Now we can invoke all theory we derived for 1 variable integration! We call $\mathscr{R}(\mathcal{C})$ the set of all Riemann integrable functions over $\mathcal{C}$.

Theorem 23.1.1

Let, γ be a "Rectifiable" smooth(or piecewise smooth) and $\mathcal{C} = \operatorname{ran}(\gamma)$ and $f \in \mathscr{B}(\mathcal{C})$. Then,

1. $f \in C^0(\mathcal{C}) \implies f \in \mathscr{R}(\mathcal{C})$
2. $$f \in \mathscr{R}(\mathcal{C}) \iff \lim_{||\mathcal{P}|| \to 0} \sum_{i=1}^{m} f(\zeta_i) s_i \quad \text{exist and equal to} \int_{\mathcal{C}} f.$$
   Here, ζ_i is tag of the interval I_i.
3. (This requires smoothness) If γ is C^1 and smooth, $f \in \mathscr{R}(\mathcal{C})$, then
   $$\int_{\mathcal{C}} f = \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, \mathrm{d}t$$

Proof. **Exercise.**

In the above theorem equation in 3 is also independent of choice of Parametrization of γ. Because any other smooth parametrized curve $\tilde{\gamma} = \gamma \circ \varphi$ where, φ is an onto continuous function.

Facts: $\mathcal{C}$ be a piecewise smooth, parametrized curve γ. $f, g \in \mathscr{R}(\mathcal{C})$ and $r \in \mathbb{R}$,then

- $\displaystyle\int f + rg = \int f + r \int g$
- $f \geq g$ over $\mathcal{C}$ then $\displaystyle\int f \geq \int g$
- $\displaystyle\int |f| \geq \left|\int f\right|$
- If $a < d < b$, if $\gamma_1 := \gamma|_{[a,d]}$ and $\gamma_2 := \gamma|_{[d,b]}$ then
$$\int_{\mathcal{C}} f = \int_{\gamma_1} f + \int_{\gamma_2} f$$

We have resolved the problems for Scalar field. **What about vector fields?**

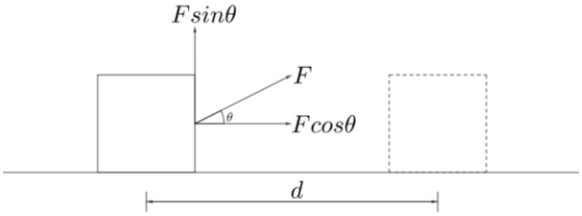


Figure 23.2: Work done by a constant Force

Suppose a particle moves a distance d under a constant force F, then work done by the force is $Fd\cos\theta = \vec{F} \cdot \vec{d}$.

If the force was not constant throughout the path then how can we calculate work done by that force? Consider the case where F is the vector field (Force in this case) defined over a curve (path)γ. Here, $\gamma : [a, b] \to \mathbb{R}^n$ and $\mathcal{C} = \operatorname{ran}(\gamma)$. So, work done throughout the whole path will be,
$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r}$$
Which is equal to, $\displaystyle\int_a^b \vec{F}(\gamma(t)) \cdot \boldsymbol{\nabla}\gamma(t)\,\mathrm{d}t$

Now we will look into some examples.

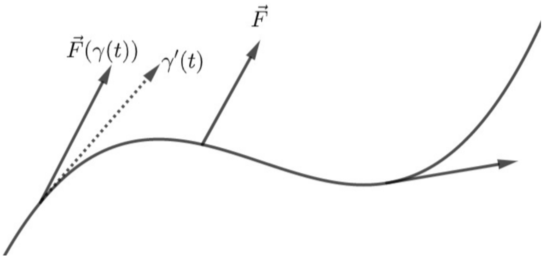


Figure 23.3: $\vec{F}$ is the vector field over the curve γ

Example 23.1.1

Find work done by the field $F(x, y, z) = (xy, xz, yz)$ along the curve $\gamma(t) = (t^2, -t^3, t^4), t \in [0, 1]$.

Answer.

$$\gamma'(t) = (2t, -3t^2, 4t^3)$$

$$\begin{aligned}\implies \text{Work done} &= \int_0^1 (-t^5, t^6, -t^7) \cdot (2t, -3t^2, 4t^3)\,\mathrm{d}t \\ &= -\frac{31}{88}\end{aligned}$$

23.1.1 Line Integration of a Vector Field

$F : \mathcal{O}_n \to \mathbb{R}^n$ be a vector field and $\gamma : [a, b] \to \mathcal{O}_n$ be a curve. We consider a partition $\mathcal{P} : a = t_0 < t_1 < \cdots < t_m = b$, Let $\mathcal{C}_i = \gamma|_{[t_{i-1}, t_i]}$ and $\gamma_i = \gamma(t_i)$, $\Delta r_i = \gamma_i - \gamma_{i-1}$.

$$R(F; \mathcal{P}) = \sum_{i=1}^{m} F(\gamma_i) \cdot \Delta r_i$$

$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r} = \lim_{\|\mathcal{P}\| \to 0} R(F, \mathcal{P}) \qquad \text{(if the limit exists)}$$

Just like the scalar field, if γ is C^1 and smooth, then

$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r} = \int_a^b F(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t \tag{23.1}$$

Lecture 24

Theorem 24.0.2

Let, $f : \mathcal{O}_n \to \mathbb{R}$ be a C^1 function and γ be a piecewise smooth C^1 curve on $\mathcal{O}_n$, joining two points A and B. Then

$$\int_c \boldsymbol{\nabla} f \cdot \mathrm{d}\vec{r} = f(B) - f(A)$$

Which means the above line integral is "independent of parametrization".

Proof.
Let, r be a parametrization of the curve γ and $r : [a, b] \to \mathcal{O}_n$ such that $r(a) = A$ and $r(b) = B$.
Evaluating the LHS gives,

$$\begin{aligned}
\int_c \boldsymbol{\nabla} f \cdot \mathrm{d}\vec{r} &\overset{(23.1)}{=} \int_a^b \boldsymbol{\nabla} f(r(t)) \cdot r'(t)\, \mathrm{d}t \\
&= \int_a^b \left(\sum_{i=1}^{n} \frac{\partial f}{\partial r_i} \times r_i'(t) \right) \mathrm{d}t \\
&= \int_a^b \frac{\mathrm{d}}{\mathrm{d}t}(f(r(t)))\, \mathrm{d}t \\
&= f(r(b)) - f(r(a)) \\
&= f(B) - f(A)
\end{aligned}$$

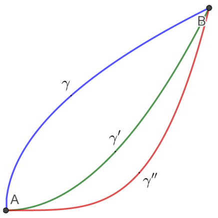


Figure 24.1: Paths from A to B

□

This result has an important consequence that we often use in Physics.

> "Work done by a conservative force always depends on the starting point and the end point, not on the path followed by the particle."

We know if the force is conservative then we can define potential energy U as $\vec{F} = -\boldsymbol{\nabla} U$. Work done by the force is simply the change of potential.

Now we will recall the basics of the Planes and Normals.

24.1 Planes and Normals

Let, $\vec{P_0} = \langle x_0, y_0, z_0 \rangle$ be a fixed vector in $\mathbb{R}^3$. $\vec{N} = \langle a, b, c \rangle \neq \vec{0}$. The plane through $\vec{P_0}$ with $\vec{N}$ as normal to this plane is,

$$\left\{ \vec{P_0} + \vec{P} \mid \vec{P} \cdot \vec{N} = 0 \right\} = \left\{ \vec{r} \mid (\vec{r} - \vec{P_0}) \cdot \vec{N} = 0 \right\}$$

Equation of the plane: Consider an arbitrary point $\vec{P} = \langle x, y, z\rangle$ on the plane then,

$$\langle x - x_0, y - y_0, z - z_0\rangle \cdot \langle a, b, c\rangle = 0$$

So, equation of the plane is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \tag{24.1}$$

Let $\vec{Q_1}, \vec{Q_2}$ be independent in $\mathbb{R}^3$ and also satisfying $\vec{Q_i} \cdot \vec{N} = 0$. Clearly, $\vec{Q_1} \times \vec{Q_2} \neq \vec{0}$. We can see $(\vec{Q_1} \times \vec{Q_2}) \cdot \vec{Q_i} = 0$, i.e., $\{\vec{Q_1}, \vec{Q_2}, \vec{Q_1} \times \vec{Q_2}\}$ form a basis of $\mathbb{R}^3$.

$$\therefore \quad \vec{Q_1} \times \vec{Q_2} = c\vec{N}$$

So, $\left\{\vec{P_0} + r_1\vec{Q_1} + r_2\vec{Q_2} \mid r_1, r_2 \in \mathbb{R}\right\}$ describes the same plane as (24.1).

24.2 Surface and Surface Integrals

Definition 24.2.1 ▶ Region

A subset $\mathcal{R} \subseteq \mathbb{R}^2$ is called a "Region" if $\mathcal{R}$ is Open ans $\mathcal{R}$ has an area (i.e. $\partial\mathcal{R}$ is **content zero**)

Definition 24.2.2 ▶ Parametrized Surface

Let $\mathcal{R} \subseteq \mathbb{R}^2$ be a region. A C^1 function $r : \mathcal{R} \to \mathbb{R}^3$ said to be a "Parametrized Surface" if :

- The component functions r_i have bounded partials
- r is $1 - 1$ function
- $\frac{\partial\vec{r}}{\partial u} \times \frac{\partial\vec{r}}{\partial v} \neq \vec{0}$ for all $(u, v) \in \mathbb{R}^2$. This means total derivative of r has rank 2.

We will call range of r as a **Surface**, $\mathcal{S} = \text{ran}(r)$.

Let, η be a map defined on $(-\varepsilon, \varepsilon)$ that maps $t \xmapsto{\eta} r(u_o = 0 + t, v_0)$. Clearly, η defines smooth curve on $\mathcal{S}$. Similarly, we can define $\tilde{\eta}$ on $(-\varepsilon, \varepsilon)$ that maps $t \xmapsto{\tilde{\eta}} r(u_0, v_0 + t)$.
Thus, $\tilde{\eta}$ also defines a smooth curve on $\mathcal{S}$.
From the Figure 24.2 we can see η and $\tilde{\eta}$ are the curves. $r_u(u_0, v_0) = \frac{\mathrm{d}\eta}{\mathrm{d}t}\Big|_{t=0}$, which gives the tangent of the curve η at point (u_0, v_0) along x axis.
Similarly, for $\tilde{\eta}$, $r_v(u_0, v_0)$ gives the tangent of $\tilde{\eta}$ along y axis. The vectors $r_u(u_0, v_0), r_v(u_0, v_0)$ spanned together to form a plane. This plane is known as **Tangent Plane**.
Since r is C^1, both $\vec{r_u}$ and $\vec{r_v}$ are continuous and hence $\vec{r_u} \times \vec{r_v}$ is continuous. Also, $\vec{r_u} \times \vec{r_v}$ is along the normal vector of $\mathcal{S}$ at $r(u_0, v_0)$ which follows from the previous section.

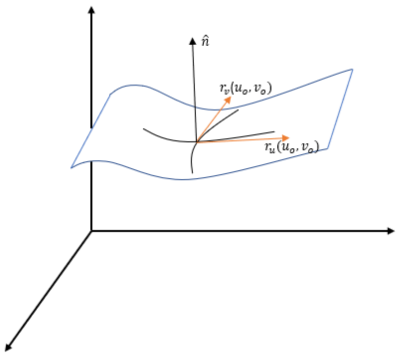


Figure 24.2

Now we will move towards a very important definition.

Definition 24.2.3 ▶ Tangent Plane

For a Parametrized Surface r, let $\operatorname{ran}(r) = \mathcal{S}$ and $r(u_0, v_0) = P$, then the plane generated by $r_u(u_0, v_0)$ and $r_v(u_0, v_0)$ through $r(u_0, v_0)$ is called the **Tangent Plane** of $\mathcal{S}$ through P.
We denote it by $T_P\mathcal{S}$.
Every element of $T_P\mathcal{S}$ is called **Tangent Vectors** at P on $\mathcal{S}$.

It can be shown that $T_P\mathcal{S}$ is independent of parametrization of $\mathcal{S}$, i.e., $T_P\mathcal{S}$ is independent of r (**Exercise**). Actually, for different $\tilde{r}$ of $\mathcal{S}$ the basis for $T_P\mathcal{S}$ will be changed. But they still generate the same plane.

24.3 Examples

We will go through some examples.

Example 24.3.1 (Graph of Function)

Let $f : \mathcal{O}_2 \to \mathbb{R}$ be a C^1 function then the graph of f is $\mathcal{G}(f) = \{(x, y, f(x, y)) : (x, y) \in \mathcal{O}_2\}$. Under the conditions $\mathcal{O}_2$ is bounded and partial derivatives of f is bounded, we want to find a parametrization of this Surface $\mathcal{G}(f)$.
Answer. Here, we use the trivial parametrization $r : \mathcal{O}_2 \to \mathbb{R}^3$ that is $r(x, y) = (x, y, f(x, y))$. Clearly, r is one - one. Now, $r_u(u, v) = (1, 0, f_u)$ and $r_v(u, v) = (0, 1, f_v)$.
So, $r_u \times r_v = (-f_u, -f_v, 1) \neq \vec{0}$. So, it is a parametrization of the surface.

Example 24.3.2 (Torus)

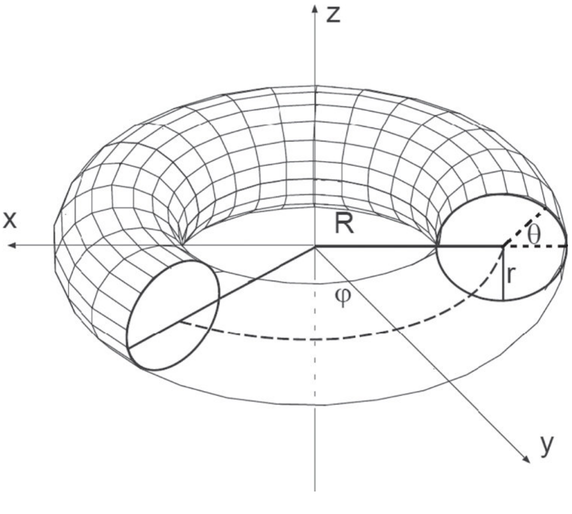


Figure 24.3: Torus.

Parametrization is given by, $(0 < r < R)$

$$r(\theta, \varphi) = ((R + r\cos\theta)\cos\varphi, (R + r\cos\theta)\sin\varphi, r\sin\theta); 0 \leq \theta, \varphi \leq 2\pi \qquad (24.2)$$

Example 24.3.3 (Surface of Revolution)

Let f, g be C^1 functions on $[0, b]$. Consider the curve $t \longmapsto (0, f(t), g(t))$ is a C^1 curve. If we rotate the curve with respect to z axis, we must get a surface. Parametrization of the surface is given by,

$$r(u, v) = (f(u)\cos v, f(u)\sin v, g(u)) \quad ; u \in [0, b], v \in [0, 2\pi]$$

Exercise. Find the Parametrization of a sphere of radius R.

Lecture 25

25.1 Tangent Plane Of $\mathcal{G}(f)$

Let, $f : \mathcal{O}_2 \to \mathbb{R}$ be a C^1 function and $r(u,v) = (u, v, f(u,v))$. Here $\operatorname{ran}(r)$ defines a surface $\mathcal{G}(f)$ as we have showed in Example 24.3.1. We also have calculated $r_v \times r_u = (-f_u, -f_v, 1)$. Now using (24.1) we can write down the equation of **Tangent space** at a point $P = (a, b, f(a,b))$ on $\mathcal{G}(f)$. Equation of the tangent space $T_P\mathcal{S}$,

$$\begin{aligned} & f_u(a,b)(x-a) + f_v(a,b)(y-b) - (z - f(a,b)) = 0 \\ \implies & z = f(a,b) + f_u(a,b)(x-a) + f_v(a,b)(y-b) \end{aligned} \tag{25.1}$$

Equation of **Normal** at the point P on the surface $\mathcal{G}(f)$ is,

$$\frac{x-a}{-f_u(a,b)} = \frac{y-b}{-f_v(a,b)} = \frac{z-f(a,b)}{1} \tag{25.2}$$

Example 25.1.1 (Equation of Tangent and Normal to $z = f(x,y) = \frac{2x}{y} - y^2$ at $(1,1,1)$)

Solution. This is a graph function so obviously a surface $f_x(x,y) = \frac{2}{y}$ and $f_y = -\frac{2x}{y^2} - 2y$. So,$\langle -f_x, -f_y, 1\rangle = \langle -2, 4, 1\rangle$. So equation of Normal is $\frac{x-1}{-2} = \frac{y-1}{4} = \frac{z-1}{1}$ and the equation of Tangent Plane is, $2(x-1) - 4(y-1) - (z-1) = 0$. ■

Example 25.1.2 (Use Tangent Plane to approximate $(1.99)^2 - \frac{1.99}{1.01}$)

Solution. Consider $z = x^2 - \frac{x}{y} = f(x,y)$. This describes a surface. Now consider $P = (2,1,2)$ be the point on the surface. Here,$\langle -f_x, -f_y, 1\rangle = \langle -3, -2, 1\rangle$. So, equation of Tangent plane at P is,

$$z = 2 + 3(x-2) + 2(y-1)$$

The given expression can be approximated as, (by putting value of x, y in the above equation of tangent plane) $z(1.99, 1.01) \approx 1.99$. ■

Our next goal is to calculate area of different surfaces. We will start with very basic example, that is, area of a plane.

25.2 Surface Area

Suppose P_0, P_1, P_2 be the points on $\mathbb{R}^3$ and coordinate vector of the points is given by,

$$\begin{aligned}\overrightarrow{OP_0} &= \langle a_0, b_0, c_0\rangle\\ \overrightarrow{OP_1} &= \langle a_1, b_1, c_1\rangle\\ \overrightarrow{OP_2} &= \langle a_2, b_2, c_2\rangle\end{aligned}$$

We will actually look at the parallelogram generated by,

$$\begin{aligned}\vec{v_1} &= \overrightarrow{P_0P_1}\\ \vec{v_2} &= \overrightarrow{P_0P_2}\end{aligned}$$

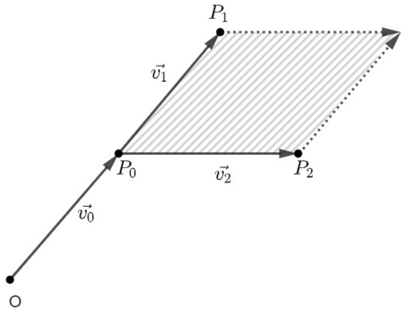


Figure 25.1: Plane $\mathcal{S}$

Any point inside the parallelogram must look like $\vec{v_0} + t_1\vec{v_1} + t_2\vec{v_2}$ for some $t_1, t_2 \in [0,1]$. So the parallelogram can be explicitly written as,

$$\mathcal{S} = \{\vec{v_0} + t_1\vec{v_1} + t_2\vec{v_2} \mid 0 \leq t_1, t_2 \leq 1\}$$

We know area of $\mathcal{S}$ is $\|\vec{v_1} \times \vec{v_2}\|$. We can describe this plane differently. If the equation of the plane was $z = ax+by+c$, then the surface of the plane can be described by $\mathcal{S} = \{(x, y, ax+by+c) \mid (x, y) \in B^2\}$. Area of $\mathcal{S} = \sqrt{1+a^2+b^2} \times \text{Area}(B^2)$. Now we should move forward to the general case.

Let, $r : B^2 \to \mathbb{R}^3$ be a function defined as $r(x, y) = (x, y, f(x, y))$ (Here f is C^1 function). Let, $\text{ran}(r)$ be the surface $\mathcal{S}$.

As we have done in the case of Riemann Integration. We should make partition of B^2 into tiny boxes. Let, $\mathcal{P} \in \mathscr{P}(B^2)$. Then,

$$B = \bigcup_{\alpha \in \Lambda(\mathcal{P})} B^2_\alpha$$

For any $\alpha \in \Lambda(\mathcal{P})$ fix $(x_\alpha, y_\alpha) \in B^2_\alpha$. Consider the tangent plane of $\mathcal{S}$ at $r(x_\alpha, y_\alpha)$ over B^2_α. Now the Tangent Plane at $r(x_\alpha, y_\alpha)$ is given by,

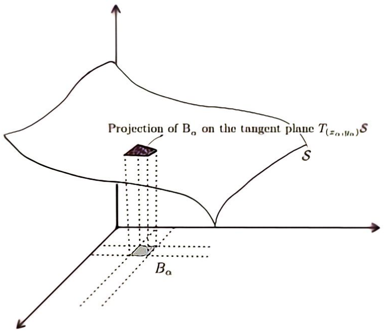


Figure 25.2: $T_{r(x_\alpha, y_\alpha)}\mathcal{S}$

$$\begin{aligned}z - f((x_\alpha, y_\alpha)) &= (Df)(x_\alpha, y_\alpha) \cdot (x_\alpha, y_\alpha)\\ \implies z &= (Df)(x_\alpha, y_\alpha) + f(x_\alpha, y_\alpha)\\ \implies z &= f_x(x_\alpha, y_\alpha) \cdot (x - x_\alpha) + f_y(x_\alpha, y_\alpha) \cdot (y - y_\alpha) + f(x_\alpha, y_\alpha)\end{aligned}$$

So, Area of $T_{r(x_\alpha, y_\alpha)}\mathcal{S}$ over B^2_α is

$$\sqrt{1 + f_x^2(x_\alpha, y_\alpha) + f_y^2(x_\alpha, y_\alpha)} \times \text{Area}(B_\alpha)^2$$

$$= \|r_x \times r_y\| \times \text{Area}(B_\alpha)^2$$

Since f is C^1 function so is r. So the total area of the surface $\mathcal{S}$ is given by,

$$\begin{aligned}\text{Area}(\mathcal{S}) &= \lim_{\|\mathcal{P}\|\to 0} \|r_x \times r_y\| \times \text{Area}(B_\alpha)^2 \\ &= \int_{B^2} \|r_x \times r_y\| \, \mathrm{d}A \\ &= \int_{B^2} \sqrt{1 + f_x^2 + f_y^2} \, \mathrm{d}A \quad \text{(In this case)}\end{aligned}$$

Will will do the above integration over any bounded set Ω as we have done in Riemann integration chapter. Over a bounded set Ω area of $\mathcal{S}$ is given by,

$$\text{Area}(\mathcal{S}) = \int_\Omega \sqrt{1 + f_x^2 + f_y^2} \, \mathrm{d}A \tag{25.3}$$

The General method for finding surface area is described by the following theorem.

Theorem 25.2.1

Let $\mathcal{R} \subseteq \mathbb{R}^2$ be a region. $r : \mathcal{R} \to \mathbb{R}^3$ be the parametrization of the surface $\mathcal{S}$. Then,

$$\text{Area}(\mathcal{S}) = \int_{\mathcal{R}} \|r_u \times r_v\| \, \mathrm{d}A$$

Lecture 26

26.1 Examples

Recall in previous lecture, we discussed how to find area of a surface over a bounded region. Now we will look towards some examples.

Example 26.1.1 (Truncated Cylinder)

Let $\mathcal{R} \subseteq \mathbb{R}^2$ be a region. $r : \mathcal{R} \to \mathbb{R}^3$ be a map defined by $r(x, y) = (\cos(x), \sin(x), y)$. Here the region is given by,

$$\mathcal{R} = \{(x, y) \mid 0 \le x \le \frac{\pi}{2}, 0 \le y \le 1\} = \left[0, \frac{\pi}{2}\right] \times [0, 1]$$

It can be seen easily that r is parametrization of a surface $\operatorname{ran}(r) = \mathcal{S}$. We want to calculate the surface area of it. For this case, $r_x = (-\sin x, \cos x, 0)$ and $r_y = (0, 0, 1)$. So, $r_x \times r_y = (\cos x, \sin y, 0)$ and hence, $\|r_x \times r_y\| = 1$. Therefore,

$$\operatorname{Area}(\mathcal{S}) = \int_{\mathcal{R}} 1 \mathrm{d}A = \int_0^{\frac{\pi}{2}} \int_0^1 1 \mathrm{d}A = \frac{\pi}{2}$$

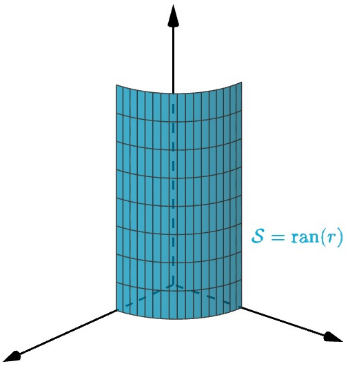


Figure 26.1: Truncated Cylinder

Example 26.1.2 (Hemisphere)

$\mathcal{S} = \left\{(x, y, z) \mid x^2 + y^2 + z^2 = 1, z > 0\right\}$ is a surface. This surface actually surface of unit Hemisphere. Parametrization of the surface $\mathcal{S}$ is given by, $r(\theta, \varphi) = (\cos\theta\cos\varphi, \sin\theta\cos\varphi, \sin\varphi)$ where, $0 \le \varphi \le \frac{\pi}{2}, 0 \le \theta \le 2\pi$. So,

$$\begin{aligned}
r_\theta &= (-\sin\theta\cos\varphi, \cos\theta\cos\varphi, 0) \\
r_\varphi &= (-\cos\theta\sin\varphi, -\sin\theta\sin\varphi, \cos\varphi) \\
\therefore r_\theta \times r_\varphi &= (\cos\theta\cos^2\varphi, \sin\theta\cos^2\varphi, \cos\varphi\sin\varphi) \\
\implies \|r_\theta \times r_\varphi\| &= \cos\varphi
\end{aligned}$$

Hence, $\operatorname{Area}(\mathcal{S}) = \displaystyle\int_0^{\frac{\pi}{2}} \int_0^{2\pi} \cos\varphi \,\mathrm{d}\theta\mathrm{d}\varphi = 2\pi$

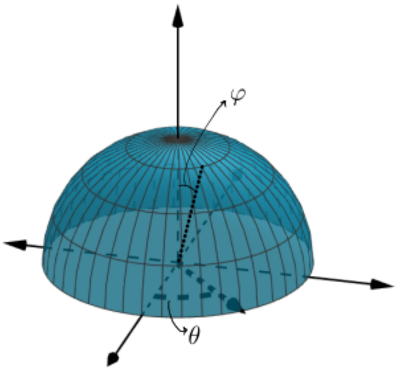


Figure 26.2: Hemisphere

26.2 Surface Integral over Scalar fields

Let, $\mathcal{R} \in \mathbb{R}^2$ be a region. $r : \mathcal{R} \to \mathbb{R}^3$ be parametrization of a surface $\mathcal{S}$. f be a scalar function defined on the surface $\mathcal{S}$. Let $f \in C(\mathcal{S})$ then we can define integration of f over the surface as following,

$$\begin{aligned}\int_{\mathcal{S}} f \, \mathrm{d}\mathcal{S} &:= \lim_{\|\mathcal{P}\| \to 0} \sum_{\alpha \in \Lambda(\mathcal{P})} f(r(x_\alpha)) \|r_u(x_\alpha) \times r_v(x_\alpha)\| \operatorname{Area}(B_\alpha^2) \\ &= \int_{\mathcal{R}} (f \circ r)(u,v) \|r_u \times r_v\| \, \mathrm{d}A \qquad (26.1)\end{aligned}$$

As we have noticed in the case of **Line Integral of over a scalar field** the integration do not depend on the parametrization of the path. In this case also if we have different two parametrization $r, \tilde{r}$ for the surface $\mathcal{S}$ then there exist a one-one continuous function φ so that the following diagram commutes, (in other words $r = \tilde{r} \circ \tilde{\varphi}$)

$$\begin{tikzcd} & \mathcal{S} & \\ \mathcal{R} \arrow[ur, "r"] \arrow[rr, "\varphi"'] & & \tilde{\mathcal{R}} \arrow[ul, "\tilde{r}"'] \end{tikzcd}$$

We can show that the integration in (26.1) is also same if we replace r by $\tilde{r}$. In other words the surface Integral over a surface is independent of its parametrization (See Figure 26.3).

Example 26.2.1 (Surface Integral over a Cone)

We want to evaluate

$$\int_{\mathcal{S}} (x^2 + y^2 + z^2) \mathrm{d}\mathcal{S}$$

where, $\mathcal{S} = \left\{(x,y,z) \mid z = \sqrt{x^2+y^2}, 0 \le z \le 1\right\}$

Solution. Let,$\mathcal{R} = \left\{(x,y) | x^2 + y^2 \le 1\right\}$ be the region and $r(x,y) = \left(x, y, \sqrt{x^2+y^2}\right)$ is the parametrization of the surface $\mathcal{S}$. Here, $r_x = \left(1, 0, \frac{x}{\sqrt{x^2+y^2}}\right)$ and $r_y = \left(0, 1, \frac{y}{\sqrt{x^2+y^2}}\right)$.

$$\begin{aligned}\therefore \int_{\mathcal{S}} (x^2+y^2+z^2) \, \mathrm{d}\mathcal{S} &= \int_{\mathcal{R}} (f \circ r) \cdot \|r_x \times r_y\| \, \mathrm{d}A \\ &= \int_{\mathcal{R}} (f \circ r) \sqrt{1 + \frac{x^2}{x^2+y^2} + \frac{y^2}{x^2+y^2}} \, \mathrm{d}A \\ &= 2\sqrt{2} \int_{\mathcal{R}} (x^2+y^2) \, \mathrm{d}A \\ &= 2\sqrt{2} \int_0^1 \int_0^{2\pi} r^3 \, \mathrm{d}\theta \, \mathrm{d}r \quad \text{(It's the polar substitution)} \\ &= \sqrt{2}\pi\end{aligned}$$

26.3 Surface Integral over a Vector field

Let $\vec{F} : \mathcal{S} \to \mathbb{R}^3$ be a vector field defined on a surface $\mathcal{S}$. We want to calculate the flux of $\vec{F}$ over the surface $\mathcal{S}$.

Let $\mathcal{S}$ be a surface and $\vec{F}$ (This function is C^1 over $\mathcal{S}$) be the Vector field defined over it. We can assume $\vec{F}$ to be a force. We want to compute the extent to which $\vec{F}$ is pushing the surface along normal to $\mathcal{S}$. In other words, we want to calculate **Flux / Flow** of $\vec{F}$ through $\mathcal{S}$.

Let $\mathrm{d}S$ be a tiny part of this surface. Flux through $\mathrm{d}S$ is given by $\vec{F} \cdot \vec{n}\,\mathrm{d}S$ where, $\vec{n}$ is the unit normal vector through the surface at a point $x \in \mathrm{d}S$. Since $\mathrm{d}S$ is small enough we can assume $\vec{n}$ is almost constant over $\mathrm{d}S$. So, the total flux must be

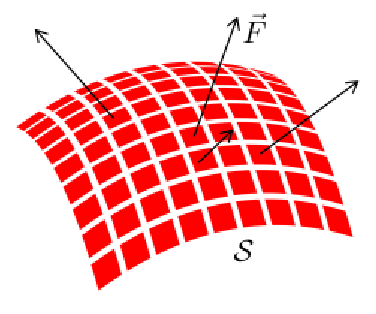


Figure 26.3: A surface with Vector field

$$\textbf{Flux} = \int_{\mathcal{S}} \vec{F} \cdot \vec{n}\,\mathrm{d}S$$

But, there is a problem! If the surface $\mathcal{S}$ has a point where two normal vectors with different direction are the at same point then the concept of flux does not make any sense. That's why we have to come up with some conditions. Namely, *Orientation of Surface.*

So we formalize the notion of "Orientation" with the following definition. Thought it is only for the surfaces in $\mathbb{R}^3$. Later in Differential Geometry, you will see this notion in a more general way (for manifolds).

Definition 26.3.1 ▶ Oriented Surface

A surface is **oriented** if there exists a continuous function $\vec{n} : \mathcal{S} \to \mathbb{R}^3$ such that $\vec{n}(x)$ is normal to $\mathcal{S}$ at the point x and $\|\vec{n}(x)\| = 1$.

Surfaces like **Möbius strip** or **Klein Bottle** are not oriented.

Definition 26.3.2 ▶ Surface Integral over Vector field

Let, $\mathcal{S}$ be an oriented surface with normal vector $\vec{n}$ and $\vec{F} : \mathcal{S} \to \mathbb{R}^3$ be a C^1 vector field defined over $\mathcal{S}$. Then the "Surface Integral" of $\vec{F}$ over $\mathcal{S}$ is defined as following,

$$\int_{\mathcal{S}} \vec{F} \cdot \mathrm{d}\vec{S} = \int_{\mathcal{S}} \vec{F} \cdot \vec{n}\,\mathrm{d}S$$

Lecture 27

27.1 Conservative Vector Fields

In the previous lecture we introduced the notion of an oriented surface. For an oriented surface $S \subseteq \mathbb{R}^3$, we call the orientation vector field $\vec{n} : S \to \mathbb{R}^3$ the **normal vector field**. Now we give an example of such a vector field.

Example 27.1.1

Take $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$.
Then

$$\vec{n}_1(x) = x \ \forall x \in \mathbb{S}^{n-1}$$

and

$$\vec{n}_2(x) = -x \ \forall x \in \mathbb{S}^{n-1}$$

are the two normal vector fields on the sphere.

Generally we consider the outward normal vector, i.e., the normal vector field given by $\vec{n}_1$ as the standard normal vector field on the sphere.

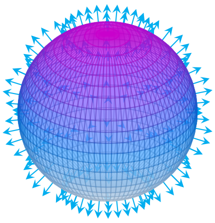


Figure 27.1: Standard normal vector field on a sphere

Formula for Normal Vector Field.
Let $\mathcal{G}(f) = \{(x, y, f(x, y)) \mid (x, y) \in \mathcal{O}_2\}$ where $f : \mathcal{O}_2 \to \mathbb{R}$ is a C^1 function. Then a parametrization of the surface $\mathcal{G}(f)$ is given by the function

$$\begin{aligned} \vec{r} &: \mathcal{O}_2 \to \mathbb{R}^3 \\ (x, y) &\mapsto (x, y, f(x, y)) \end{aligned}$$

Then we have

$$\vec{r}_x \times \vec{r}_y = (-f_x, -f_y, 1)$$

Then a normal vector field is given by

$$\vec{n}(x, y) = \frac{\vec{r}_x \times \vec{r}_y}{\|\vec{r}_x \times \vec{r}_y\|}$$

Unless otherwise mentioned this will be our standard orientation of the normal vector field.

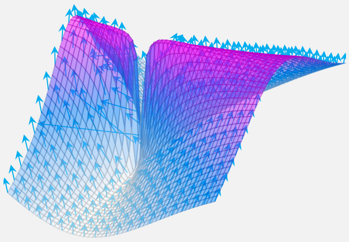


Figure 27.2: Normal vector fields on a graph surface

Usually computation of $\int_S \vec{F} \cdot \mathrm{d}\vec{S}$ is complicated, let us look at some examples to gain more familiarity.

Example 27.1.2

Consider the vector field $\vec{F}(x,y,z) = (x,y,z)$ on $S = \operatorname{ran}(r)$, where

$$\vec{r}(x,y) = (\cos x, \sin x, y) \quad 0 \le x \le \frac{\pi}{2}, 0 \le y \le 1$$

Then $\vec{r}_x \times \vec{r}_y = (\cos x, \sin x, 0)$, so $\vec{n}(x,y) = (\cos x, \sin x, 0)$ is a normal vector field.

$$\begin{aligned}\int_S \vec{F} \cdot \mathrm{d}\vec{S} &= \int_S \vec{F} \cdot \vec{n} \mathrm{d}s \\ &= \int_0^1 \int_0^{\frac{\pi}{2}} \vec{F}(\vec{r}(x,y)) \cdot (\vec{r}_x \times \vec{r}_y) \, \mathrm{d}A \\ &= \int_0^1 \int_0^{\frac{\pi}{2}} (\cos x, \sin x, y) \cdot (\cos x, \sin x, 0) \mathrm{d}A \\ &= \int_0^1 \int_0^{\frac{\pi}{2}} \mathrm{d}A \\ &= \frac{\pi}{2}\end{aligned}$$

We already know that $\int_{\mathcal{C}} \nabla f \cdot \mathrm{d}r = f(B) - f(A)$, now a natural question that arises is

Question: Given $\vec{F}$, does there exist f a scalar field such that $\nabla f = \vec{F}$?

Definition 27.1.1 ▶ Conservative Vector Field

A vector field $\vec{F}$ on $\mathcal{O}_n$ is called **conservative** if there exists a scalar field $f \in C^1(\mathcal{O}_n)$ such that $\nabla f = \vec{F}$, then f is called the **potential function**.

Theorem 27.1.1

Let $\vec{F}$ be a vector field over $\mathcal{O}_n$, the following are equivalent:

1. $\vec{F}$ is conservative.
2. $\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}r = 0$, for all closed and piecewise smooth curve $\mathcal{C}$.
3. $\int_{\mathcal{C}_1} \vec{F} \cdot \mathrm{d}r = \int_{\mathcal{C}_2} \vec{F} \cdot \mathrm{d}r$, for all curves $\mathcal{C}_1$ and $\mathcal{C}_2$ with same initial and end points.

Question: Given a vector field $\vec{F}$, can we conclude $\vec{F}$ is conservative? (NO!)

We will give a general picture for the most common case, when $n = 3$. Let $\vec{F} = (P, Q, R)$ where P, Q, R are scalar fields. Now if $\vec{F} = \nabla f$ for some scalar field f, then we would have

$$\begin{aligned}f_x &\equiv \frac{\partial f}{\partial x} = P \\ f_y &\equiv \frac{\partial f}{\partial y} = Q \\ f_z &\equiv \frac{\partial f}{\partial z} = R\end{aligned} \tag{27.1}$$

Then we can define **curl** of a vector field

$$\nabla \times \vec{F} := \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}$$

Then expanding this out and using the relations (27.1) and others we get that $\nabla \times \vec{F} = 0$. So, we have proved that if $\vec{F}$ is conservative then $\nabla \times \vec{F} = 0$.

Remark. Thus, a necessary condition for a vector field to be conservative is that, its curl should be the zero vector field.

Example 27.1.3

Let $\vec{F}(x, y) = (y - 3, x + 2) = (P, Q)$ (say), then $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} = 1$. Let f be a possible potential function, then

$$\frac{\partial f}{\partial x} = y - 3 \quad \text{and} \quad \frac{\partial f}{\partial y} = x + 2$$

Then by **Fundamental Theorem of Calculus** (assuming domain is convex) we get

$$f(x, y) = xy - 3x + g(y)$$

But then using $\frac{\partial f}{\partial y} = x + 2$ we get

$$x + g'(y) = \frac{\partial f}{\partial y} = x + 2 \Rightarrow g'(y) = 2$$

Therefore taking $f(x, y) = xy - 3x + 2y$ gives us a potential function for the vector field $\vec{F}$.

Remark. This approach works for all $\vec{F}$ such that $\nabla \times \vec{F} = 0$ and the domain is convex.

Example 27.1.4

Let $\vec{F}(x, y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right) = (P, Q)$ (say) on $\mathbb{R}^2 \setminus \{0\}$. Then we have $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, but we will show that $\vec{F}$ is not conservative. Consider the curve

$$\mathcal{C} : \gamma(t) = (\cos t, \sin t), \quad 0 \le t \le 2\pi$$

then

$$\begin{aligned} \int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}r &= \int_0^{2\pi} \vec{F}(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t \\ &= \int_0^{2\pi} (-\sin t, \cos t) \cdot (-\sin t, \cos t)\,\mathrm{d}t \\ &= \int_0^{2\pi} \mathrm{d}t \\ &= 2\pi \end{aligned}$$

But $\mathcal{C}$ is clearly a closed curve, hence by Theorem 27.1.1 we must have $\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}r = 0$. (Contradiction!)

27.2 Green's Theorem

Definition 27.2.1 ▶ Simply Connected Domain

Let $\mathcal{D}$ be an open and connected set. Let $\mathcal{C}$ be a simple and closed curve if $\mathcal{C}$ can be shrunk continuously to a point inside $\mathcal{D}$, then we say $\mathcal{D}$ is **simply connected**.

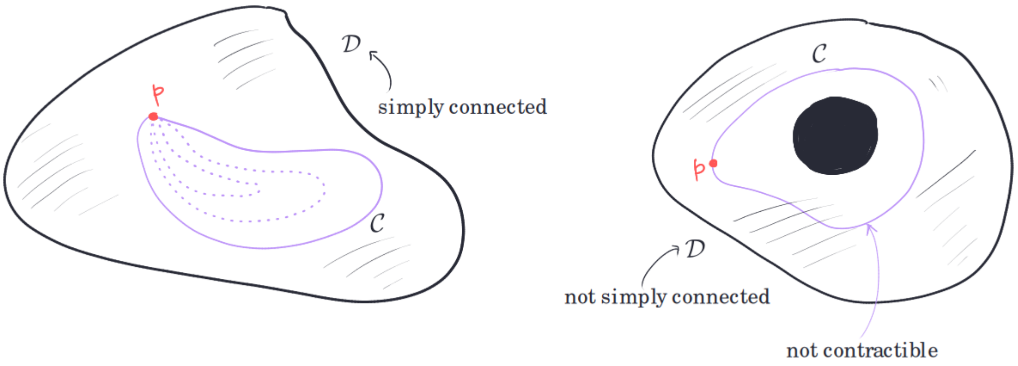


Figure 27.3: Examples of simply connected and not simply connected region

Theorem 27.2.1 (Green's Theorem)

Let $\mathcal{R} \subseteq \mathbb{R}^2$ be a simply connected domain with boundary curve $\mathcal{C}$ where parametrization is taken in anti-clockwise direction. Let $\vec{F} = (P, Q)$ be a C^1 vector field on $\mathcal{R}$, then

$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}r := \int_{\mathcal{C}} P \, \mathrm{d}x + Q \, \mathrm{d}y = \int_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A$$

What happens when $\mathcal{R}$ is not simply connected?

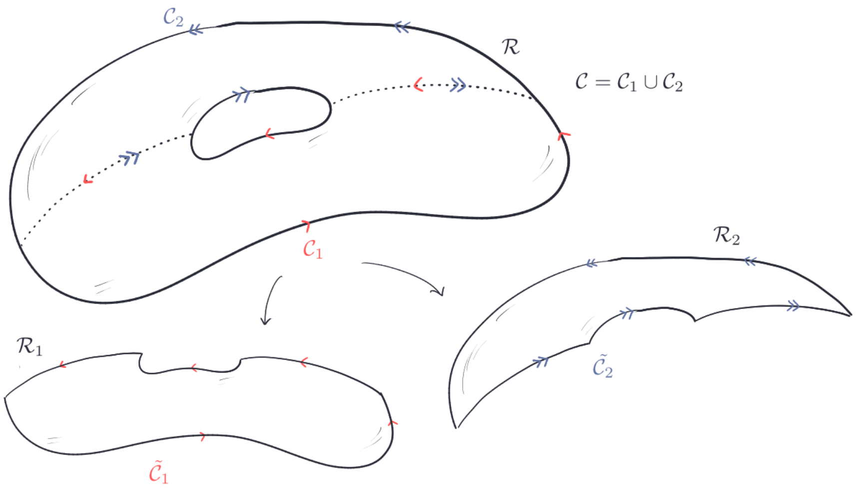


Figure 27.4: You break up the region $\mathcal{C}$ with the hole into two regions without holes $\mathcal{C}_1$ and $\mathcal{C}_2$.

$$\begin{aligned}\int_{\mathcal{C}} P\,\mathrm{d}x + Q\,\mathrm{d}y &= \int_{\tilde{\mathcal{C}}_1} P\,\mathrm{d}x + Q\,\mathrm{d}y + \int_{\tilde{\mathcal{C}}_2} P\,\mathrm{d}x + Q\,\mathrm{d}y \\ &= \int_{\mathcal{R}_1} (Q_x - P_y)\,\mathrm{d}A + \int_{\mathcal{R}_2} (Q_x - P_y)\,\mathrm{d}A \\ &= \int_{\mathcal{R}} (Q_x - P_y)\,\mathrm{d}A\end{aligned}$$

Lecture 28

28.1 Green's Theorem

Theorem 28.1.1 ($\mathbb{R}^2$ version of Green's Theorem)

Let $\mathcal{R} \subseteq \mathbb{R}^2$ be a simply connected domain with boundary curve $\mathcal{C}$ where parametrization is taken in anti-clockwise direction. Let $\vec{F} = (P, Q)$ be a C^1 vector field on $\mathcal{R}$, then

$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}r := \int_{\mathcal{C}} P \,\mathrm{d}x + Q \,\mathrm{d}y = \int_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}A$$

Proof.
(for Simple region)
Let $\mathcal{R} = \{(x, y) \mid a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}$ be a simple region. Here $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{V}_2 \cup \mathcal{C}_2 \cup \mathcal{V}_1$ is the curve bounding the region along anti-clockwise direction (as shown in Figure 28.1).

Now,

$$\begin{aligned} -\int_{\mathcal{R}} \frac{\partial P}{\partial y} \,\mathrm{d}A &= -\int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y} \,\mathrm{d}y \,\mathrm{d}x \\ &= -\int_a^b (P(x, \varphi_2(x)) - P(x, \varphi_1(x))) \,\mathrm{d}x \end{aligned}$$

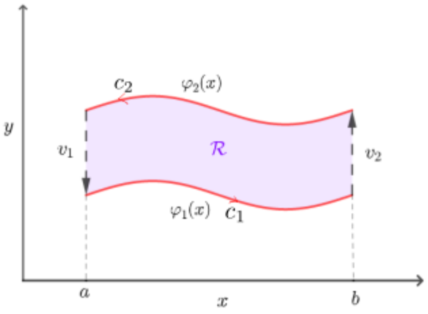


Figure 28.1: A simple region

The curves $\mathcal{C}_1, \mathcal{C}_2, \mathcal{V}_1, \mathcal{V}_2$ can be explicitly written as,

$$\begin{aligned} \mathcal{V}_1 &= \{(a, t) \mid \varphi_1(a) \leq t \leq \varphi_2(a)\} \\ \mathcal{C}_1 &= \{(x, \varphi_1(x)) \mid a \leq x \leq b\} \\ \mathcal{V}_2 &= \{(a, t) \mid \varphi_1(b) \leq t \leq \varphi_2(b)\} \\ \mathcal{C}_2 &= \{(x, \varphi_2(x)) \mid a \leq x \leq b\} \end{aligned}$$

We compute the integrals for P over these curves and obtain,

$$\begin{aligned} \int_{\mathcal{V}_1} P \,\mathrm{d}x &= \int_{\mathcal{V}_1} P(x(t), y(t)) \frac{\mathrm{d}x(t)}{\mathrm{d}t} \mathrm{d}t = 0 \\ \int_{\mathcal{C}_1} P \,\mathrm{d}x &= \int_a^b P(t, \varphi_1(t)) \,\mathrm{d}t \\ \int_{\mathcal{C}_2} P \,\mathrm{d}x &= \int_a^b P(t, \varphi_2(t)) \,\mathrm{d}t \end{aligned}$$

$$\implies \int_{\mathcal{C}} P\mathrm{d}x = -\int_{\mathcal{R}} \frac{\partial P}{\partial y}\,\mathrm{d}A$$

By similar mechanism we can show $\int_{\mathcal{C}} Q\,\mathrm{d}y = \int_{\mathcal{R}} \frac{\partial Q}{\partial y}\,\mathrm{d}A$. The rest follows from here. ☐

Example 28.1.1

Let $\mathcal{C}$ be the boundary of $[0,1]^2$, i.e., $\partial[0,1]\times[0,1]=\mathcal{C}$. Evaluate

$$\int_{\mathcal{C}} \langle x^2-y^2, 2xy\rangle$$

Solution. We can decompose $\mathcal{C}=\mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3\cup\mathcal{C}_4$ (as in the following picture)
Let $P(x,y)=x^2-y^2, Q(x,y)=2xy$. Then the integral,

$$\begin{aligned}\int_{\mathcal{C}} P\,\mathrm{d}x + Q\,\mathrm{d}y &= \iint_{[0,1]^2} (2y+2y)\,\mathrm{d}A \qquad \text{(Green's Theorem)}\\ &= \int_0^1\int_0^1 4y\,\mathrm{d}y\,\mathrm{d}x\\ &= 2\end{aligned}$$

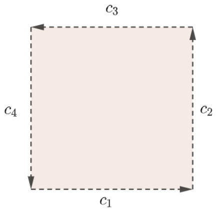


Figure 28.2: $\partial([0,1]^2)$

If we try to calculate the integral **directly**, we will end up getting same result.

Area of a closed Region. Let $\mathcal{R}$ (simply connected) be a closed region and $\mathcal{C}=\partial\mathcal{R}$ be the curve enclosing the region. Using Green's Theorem we get,

$$\text{Area}(\mathcal{R}) = \int_{\mathcal{R}} \mathrm{d}A = \int_{\mathcal{C}} x\mathrm{d}y = \int_{\mathcal{C}} -y\mathrm{d}x = \int_{\mathcal{C}} \frac{x\mathrm{d}y - y\mathrm{d}x}{2}$$

Example 28.1.2 (Area inside the ellipse: $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$)

Solution. Parametrization of ellipse $x=a\cos t, y=b\sin t$ where $t\in[0,2\pi)$. Using the above application of Green's Theorem we can write,

$$\text{Area} = \int_{\mathcal{C}} x\mathrm{d}y = ab\int_0^{2\pi} \cos^2 t\mathrm{d}t = \pi ab$$

Theorem 28.1.2 (Independence of path)

Let $\vec{F}$ be a C^1 vector field on $\mathbb{R}^2$ such that $\int_{\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r}$ is independent of path. Then $\vec{F}$ is conservative over an open and simply connected domain.

Proof. Let $\mathcal{D}$ be an open and connected domain. $\vec{F}=\langle P,Q\rangle$ is defined over $\mathcal{D}$. Also let $P_0=\langle x_0,y_0\rangle$ be a fixed point in the domain $\mathcal{D}$ and $P_1=\langle x,y\rangle\in\mathcal{D}$ be a variable point. $\mathcal{C}$ be a smooth curve joining P_0 and P_1. Define

$$\varphi(x,y) = \int_{\mathcal{C}} \vec{F}\cdot\mathrm{d}\vec{r}$$

Since, $\mathcal{D}$ is open set, so we must get an open ball centered at P_1 contained in $\mathcal{D}$. Take a point $P_1' = \langle x_1, y\rangle$ inside that open ball such that $x_1 < x$. Let $\mathcal{C}_1$ be a smooth curve from P_0 to P_1 and $\mathcal{C}_2$ be a line segment from P_1' to P_1. So, $\mathcal{C}_1 \cup \mathcal{C}_2$ defines a smooth curve from P_0 to P_1.

$$\begin{tikzcd}
 & P_1'(x_1, y) \arrow[r, "\mathcal{C}_2"] & P_1(x, y) \\
P_0(x_0, y_0) \arrow[ur, bend left, "\mathcal{C}_1"] \arrow[urr, bend right, "\mathcal{C}"'] & &
\end{tikzcd}$$

As $\displaystyle\int_{\mathcal{C}} \vec{F}\cdot \mathrm{d}\vec{r}$ is path independent We can write,

$$\varphi(x, y) = \int_{\mathcal{C}_1} \vec{F}\cdot \mathrm{d}\vec{r} + \int_{\mathcal{C}_2} \vec{F}\cdot \mathrm{d}\vec{r}$$

Now we take the partial derivative of both sides of this equation with respect to x. The first integral does not depend on the variable x since $\mathcal{C}_1$ is the path from $P_0(x_0, y_0, z_0)$ to $P_1'(x_1, y, z)$ and so partial differentiating this line integral with respect to x is zero.

$$\begin{aligned}
\frac{\partial \varphi}{\partial x} &= \frac{\partial}{\partial x}\left(\int_{\mathcal{C}_1} \vec{F}\cdot \mathrm{d}\vec{r} + \int_{\mathcal{C}_2} \vec{F}\cdot \mathrm{d}\vec{r}\right) \\
&= \underbrace{\frac{\partial}{\partial x}\left(\int_{\mathcal{C}_1} \vec{F}\cdot \mathrm{d}\vec{r}\right)}_{=0} + \frac{\partial}{\partial x}\left(\int_{\mathcal{C}_2} \vec{F}\cdot \mathrm{d}\vec{r}\right)
\end{aligned}$$

Also, $\mathcal{C}_2$ can be parametrized as $r(t) = \langle t, y\rangle$ where $t \in [x_1, x]$. So,

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\int_{\mathcal{C}_2} \vec{F}\cdot \mathrm{d}\vec{r}\right) &= \frac{\partial}{\partial x}\left(\int_{x_1}^{x} \langle P(t, y), Q(t, y)\rangle \cdot \langle 1, 0\rangle \,\mathrm{d}t\right) \\
&= \frac{\partial}{\partial x}\left(\int_{x_1}^{x} P(t, y)\,\mathrm{d}t\right) \\
&= P(x, y) \qquad \text{[Fundamental Theorem of calculus]}
\end{aligned}$$

Similarly, we can show that, $\frac{\partial \varphi}{\partial y} = Q(x, y)$. And hence, $\boldsymbol{\nabla}\varphi = \vec{F}(x, y)$. We can define φ as the potential of $\vec{F}$. ☐

Theorem 28.1.3

Let $\mathcal{D}$ be a simply connected domain in $\mathbb{R}^2$ and $\vec{F}$ is a C^1 vector field on $\mathcal{D}$. Then $\vec{F}$ is conservative iff $\boldsymbol{\nabla}\times \vec{F} = 0$ on $\mathcal{D}$.

Proof. ($\Rightarrow$) This direction is trivial.

($\Leftarrow$) From Green's Theorem we can say that $\displaystyle\int_{\mathcal{C}} \vec{F}\cdot \mathrm{d}\vec{r} = 0$ over all closed curve $\mathcal{C}$. For any two point $p_0, p_1 \in \mathcal{D}$ if $\gamma_1, \gamma_2 : [0, 1] \to \mathcal{D}$ are two smooth curves joining p_0 and p_1. (i.e., $\gamma_1(0) = \gamma_2(0) = p_0$ and $\gamma_1(1) = \gamma_2(1) = p_1$) then $\gamma_1 \cup \gamma_2(1-t)$ is a closed curve. So, $\displaystyle\int_{\gamma_1} \vec{F}\cdot \mathrm{d}\vec{r} = \int_{\gamma_2} \vec{F}\cdot \mathrm{d}\vec{r}$. Which means the integral is path independent. Using the previous theorem we can say, $\vec{F}$ is conservative on $\mathcal{D}$. ☐

28.2 Gauss Divergence Theorem

Definition 28.2.1 ▶ Divergence of a vector field

Given a vector field $\vec{F} = (f_1, \cdots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$, the "Divergence" of $\vec{F}$ is,

$$\operatorname{div}(F) = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} \equiv \boldsymbol{\nabla} \cdot \vec{F}$$

Theorem 28.2.1 (Gauss Divergence Theorem)

Let $\mathcal{D} \subseteq \mathbb{R}^3$ a solid domain, $\partial\mathcal{D}$ be an oriented surface. Let $\vec{F} = \langle P, Q, R \rangle$ be a C^1 vector field on an open surface containing $\mathcal{D} \cup \partial\mathcal{D}$. Then,

$$\underbrace{\int_{\partial\mathcal{D}=\mathcal{S}} \vec{F} \cdot \mathrm{d}\vec{S}}_{\text{surface integral}} = \underbrace{\int_{\mathcal{D}} \boldsymbol{\nabla} \cdot \vec{F} \,\mathrm{d}V}_{\text{volume integral}}$$

Just like FTC, the behavior over a volume is fully determined by the behavior at the boundary. Proof of this theorem is beyond our reach. But we can see the proof for simple cases.

Proof. (For a simple case) Consider $\mathcal{D} = \{(x, y, z) \mid \varphi_1(x, y) \le z \le \varphi_2(x, y), (x, y) \in [a, b] \times [c, d]\}$. (**Exercise.**) Complete the proof! ☐

Example 28.2.1

$F(x, y, z) = \left\langle x + y, z^2, x^2 \right\rangle$ and S be the hemisphere $x^2 + y^2 + z^2 = 1, z > 0$. Compute,

$$\int_S \vec{F} \,\mathrm{d}\vec{S}$$

Solution. Notice that S is open surface. We want to use Gauss Theorem 28.2.1. So we need a close surface. Let S_1 be the surface $x^2 + y^2 \le 1$. Then $S \sqcup S_1$ is a closed surface.

$$\begin{aligned}
\int_{S \sqcup S_1} \vec{F} \cdot \mathrm{d}\vec{S} &= \int_{x^2+y^2,z^2 \le 1, z \ge 0} \boldsymbol{\nabla} \cdot \vec{F} \,\mathrm{d}V \\
&= \int_{x^2+y^2,z^2 \le 1, z \ge 0} \mathrm{d}V \\
&= \frac{2\pi}{3}
\end{aligned}$$

Parametrization of the surface $S_1 = \left\{(x, y, 0) \mid x^2 + y^2 = 1\right\}$. So, $r_x \times r_y = \langle 1, 0, 0 \rangle \times \langle 0, 1, 0 \rangle$.

$$\begin{aligned}
\int_{S_1} \vec{F} \cdot \mathrm{d}\vec{S} &= \int_{x^2+y^2 \le 1} \left\langle x + y, z^2, x^2 \right\rangle \cdot \langle 0, 0, 1 \rangle \,\mathrm{d}A = \int_{x^2+y^2 \le 1} x^2 \,\mathrm{d}A \\
&= \int_0^{2\pi} \int_0^1 r^3 \cos^2\theta \,\mathrm{d}r \,\mathrm{d}\theta = \frac{\pi}{4} \\
\Rightarrow \int_S \vec{F} \cdot \mathrm{d}\vec{S} &= \frac{11\pi}{12}
\end{aligned}$$

28.3 Stokes' Theorem

Theorem 28.3.1 (Stokes' Theorem)

Let $\mathcal{C}$ be a C^1 curve enclosing an oriented surface $\mathcal{S}$ in $\mathbb{R}^3$. Let, $\vec{F} = \langle P, Q, R\rangle$ be a C^1 vector field on an open set containing $\mathcal{S}$. Then,

$$\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r} = \int_{\mathcal{S}} \left(\boldsymbol{\nabla} \times \vec{F}\right) \cdot \mathrm{d}\vec{S}$$

Here orientation of $\mathcal{S}$ and direction of $\mathcal{C}$ is same.

Example 28.3.1

Compute $\displaystyle\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r}$, where $\mathcal{C} : x^2 + y^2 = 9, z = 4$ and $\vec{F} = \langle -y, x, xyz\rangle$.

Solution. $\boldsymbol{\nabla} \times \vec{F} = \langle xz, -yz,, 2\rangle$. By convention, we should assume direction of $\mathcal{C}$ is along counter-clockwise direction. So, The normal vector of $\mathcal{S}$ is along negative z axis. So, required integral,

$$\begin{aligned}\int_{\mathcal{C}} \vec{F} \cdot \mathrm{d}\vec{r} &= \int_{\mathcal{S}} (\boldsymbol{\nabla} \times \vec{F}) \cdot \mathrm{d}\vec{S} \\ &= \int_{x^2+y^2\leq 1, z=4} (\boldsymbol{\nabla} \times \vec{F}) \cdot \langle 0, 0, -1\rangle \, \mathrm{d}A \\ &= -2 \int_{x^2+y^2\leq 1, z=4} \mathrm{d}A \\ &= -18\pi\end{aligned}$$

Stoke's Theorem is the $\mathbb{R}^3-$analogue of Green's Theorem 28.1.1. If we take the third component of $\vec{F}$ to be zero, i.e., $R = 0$, then Stoke's Theorem 28.3.1 gives us back Green's Theorem 28.1.1.

There is a generalized version of Stokes' theorem. Just for information the theorem is stated below.

- If Ω is an oriented n-manifold (with boundary) and ω is a differential form ($(n-1)$ form). Then integral of ω over the boundary $\partial\Omega$ of the manifold Ω is given by,

$$\int_{\partial\Omega} \omega = \int_{\Omega} \mathrm{d}\omega$$

Index