Multi-variable Calculus

Archisman Bhattacharjee

Preface

This lecture note of Multi-variable calculus is a consequence of a series of forty lectures given to some undergraduate physics majors. The set up is finite dimensional Euclidean space $\mathbb{R}^n$. The pre-requisite is quite extensive, but a significant discussion in the beginning (Section 1) covers every single thing required to go through this.

There are two goals of this lecture notes. First one is to teach Multi-variable calculus meticulously to the Physics majors who are not comfortable to use a mathematical result blindly without knowing the concepts, rather intrigued to know the actual mathematical formalisation. And the second one is to give a brief ready-made materials consisting all the concepts and results of an one semester Multi-variable course for math majors.

The explanations here are completely rigorous and always follow a well-defined set up, exactly how a typical math course should be. But most of the proofs are omitted because the purpose is to build concepts and train them to use pure mathematical concepts in problems. A lot of exercises are included in each of the sections. Some of them are Physics problems which require multi-variable calculus to be solved. Also at the end an enriched problem section is there to get a deeper insight into the subject. It shows how some new concepts of Mathematics are defined being motivated by physical phenomenon. It also shows how the proper knowledge of abstract mathematical objects becomes handy to get to know concepts of Physics.

I have tried to include everything of figures, formulas and words equally with the development of theories to reach to readers.

Section 1 covers all the pre-requisites carefully. Section 2 defines the concept of derivative on higher dimension building upon the same in one-dimensional space. Section 3 follows section 2 and defines Some important objects which are indispensable to differentiation. Section 4 and 5 on higher dimensions give idea about a lot of stuffs which we are familiar with on single variable. Next section gives criterion to find extrema of a real valued function defined on $\mathbb{R}^n$. Chapter 7 has a very crucial role in the theory of calculus. These theorems give us the proper hypothetical set up to get an explicit function of less number of variables from an implicit function and to find the inverse of a function locally. Next one is same as chapter 6 but here we find extrema given some more restrictions about the domain of the function.

Section 9 defines the concept of integrations over rectangles/boxes on higher dimensions. Next section defines a whole new concept of Lebesgue measure and gives a larger class of integrable functions. In section 11 we define integrations over any other generalised non-boxes. Section 12 discusses about the Fundamental theorem of Integral Calculus. Next one gives useful ways to compute integrals.

The last part (Section 13-20) is the most useful one for physicists. This defines integrals over parametrized curves and surfaces, defines objects like vector-fields, curl, divergence. Section $16, 19, 20$ are pretty similar. They give a relation between the integral of a function over a curve or surface and the integration of sort of its derivative over the region enclosed by the same curve or surface. The final section contains some advanced problems.

Comments and corrections should be sent to `archisman.b@niser.ac.in`

Archisman Bhattacharjee
August 25, 2020

Contents

Symbols

$=$ - equals, $\neq$ - not equals, $<$ - less than, $>$ - greater than, $\leq$ - less than or equal to, $\geq$ - greater than or equal to, $\forall$ - for all, $\exists$ - there exists, $\exists!$ -there exists unique, $\in$ - belongs to, $\ni$ - such that, $\wedge$ - and, $\vee$ - or, $\implies$ - implies, $\Leftrightarrow$/iff - if and only if, ϕ - Empty set, $\mathbb{N}$ - set of natural numbers, $\mathbb{Z}$ - set of integers, $\mathbb{Q}$ - set of rational numbers, $\mathbb{R}$ - set of real numbers, $\mathbb{C}$ - set of complex numbers. $A \subseteq B$ - A is a subset of B or equal to B, $A \cup B$ - A union B, $A \cap B$ - Intersection of A and B

1 Pre-requisites

1.1 Sets and functions

Definition 1 (Set) *A well defined collection of distinct objects*

Example : Set of real numbers, Set of countries in the world.

Union of two sets A and B is $A \cup B = \{x \mid x \in A \vee x \in B\}$.
Intersection of two sets A and B is $A \cap B = \{x \mid x \in A \wedge x \in B\}$.
Cartesian product of two sets A and B is $A \times B = \{(a, b) \mid a \in A \wedge b \in B\}$.

Definition 2 (Function) *f is a function from a set A to a set B if*

- $\forall a \in A, f(a) \in B$
- $a = b \implies f(a) = f(b)$

A is called the domain of the function f and B is the co-domain. $f(A) = \{b \in B \mid b = f(a_b) \text{ for some } a_b \in A\}$ is called the range of A or image of A. $f(a)$ is image of a and $\{a \in A \mid f(a) = b\}$ is pre-image of b.
f is one-to-one/injective function if for all x and $y \in A, x \neq y \implies f(x) \neq f(y)$.
f is surjective/onto function if for all $b \in B$ there exists $a_b \in A$ such that $f(a_b) = b$.
f is bijective function if it is both injective and surjective.

Example : $f : \mathbb{R}$ to $\{0, 1, \pi\}$ by

$$f(x) = \begin{cases} 0 & if\ x \geq 0.7 \\ \pi & if\ x < 0.7 \end{cases}$$

is a function.
$f : \mathbb{R}$ to $\mathbb{R}$ by $f(x) = e^x$ is one-to-one function.
$f : \mathbb{R}$ to $\mathbb{R}^+ \cup \{0\}$ by $f(x) = x^2$ is surjective function.
$f : \mathbb{R}$ to $\mathbb{R}$ by $f(x) = x^3$ is bijection.

1.2 Metric spaces

Definition 3 (Metric Space) *A metric space* (X, d_X) *is a set* X *equipped with a distance function* $d_X : X \times X \to \mathbb{R}^+ \cup \{0\}$ *such that*

- $\forall\, a, b \in X,\ d_X(a,b) = 0 \Leftrightarrow a = b$
- $\forall\, a, b \in X,\ d_X(a,b) = d_X(b,a)$
- $\forall\, a, b, c \in X,\ d_X(a,b) + d_X(b,c) \geq d_X(a,c)$.

Example : Set of real numbers is a metric space with the distance $d(a,b) = |a - b|$.
Take any set S, define a metric d on it by

$$d(x,y) = \begin{cases} 1 & if\ x \neq y \\ 0 & if\ x = y \end{cases}$$

Now S is a metric space with this distance function. This is called discrete metric on S.

1.3 Sequence and it's limit

A sequence on a metric space X is a function f from $\mathbb{N}$ to X. We denote the sequence as $\{a_n\}_n$ where $f(i) = a_i\ \forall\ i \in \mathbb{N}$.
Limit of a sequence $\{a_n\}$ on X is $l \in X$ if $\forall\ \varepsilon > 0\ \exists\ n_\varepsilon \in \mathbb{N} \ni d_X(a_n, l) < \varepsilon\ \forall\ n \geq n_\varepsilon$.

Example : In $\mathbb{R}$, define $x_n = \dfrac{1}{n}$. Then the limit of x_n is 0.

Series

Let $\{a_n\}_n$ be a sequence on $\mathbb{R}$. The series $\Sigma_{i \in \mathbb{N}} a_i$ is the limit of the sequence of partial sum S_n. The sequence of partial sum is defined by $S_n = \Sigma_{i-1}^{n} a_i$.
If this limit exists we say the series is convergent, else we say the series is divergent.

1.4 Limit and continuity of a function

f is a function from (X, d_X) to (Y, d_Y). Limit of f at a point $c \in X$ is $l \in Y$ (We denote it by $\lim_{x \to c} f(x) = l$) if $\forall\ \varepsilon > 0,\ \exists\ \delta_\varepsilon > 0 \ni d_X(x,c) < \delta_\varepsilon \implies d_Y(f(x), l) < \varepsilon$.
f is continuous at a point $\alpha \in X$ if $\lim_{x \to \alpha} f(x) = f(\alpha)$.

Example : $f : \mathbb{R}$ to $\mathbb{R}$ by $f(x) = \dfrac{1}{x}$. Then $\lim_{x \to 1} f(x) = 1$. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} \dfrac{sin(x)}{x} & if\ x \neq 0 \\ 1 & if\ x = 0 \end{cases}$$

Now f is continuous at 0 as $\lim_{x \to 0} \dfrac{sin(x)}{x} = 1$.

Sequential criterion of limit continuity

f is a function from (X, d_X) to (Y, d_Y). Limit of f at a point $c \in X$ is $l \in Y$ (We denote it by $\lim_{x \to c} f(x) = l$) if and only if for all sequences $\{x_n\}$, $\lim\{x_n\} = c \Rightarrow \lim\{f(x_n)\} = l$. Here $\{f(x_n)\}$ can be considered as a sequence $\{y_n\}$ on Y where $y_i = f(x_i)$ for all $i \in \mathbb{N}$. Similarly f is a function from (X, d_X) to (Y, d_Y). f is continuous at $c \in X$ if and only if for all sequences $\{x_n\}$, $\lim\{x_n\} = c \Rightarrow \lim\{f(x_n)\} = f(c)$.

1.5 Bound of sets

Archemedian Property

We don't state the property here. Rather we look at the most useful corollary of it which is equivalent to the main statement.
$\forall\, \varepsilon > 0,\ \exists\, n_\varepsilon \in \mathbb{N} \ni \dfrac{1}{n_\varepsilon} < \varepsilon$ i.e. same as saying $\forall\, \varepsilon > 0,\ \exists\, n_\varepsilon \in \mathbb{N} \ni \dfrac{1}{n} < \varepsilon\ \forall\, n \geq n_\varepsilon$.

Definition 4 (Upper bound and lower bound) *Let $S \subseteq \mathbb{R}$. $x \in \mathbb{R}$ is an upper bound of S if $s \leq x\ \forall\, s \in S$. $y \in \mathbb{R}$ is a lower bound of S if $s \geq y\ \forall\, s \in S$.*

A set $S \subseteq \mathbb{R}$ is bounded above if it has an upper bound. It is bounded below if it has a lower bound. If a set is both bounded above and bounded below, then the set is called Bounded set.

Example : The set $A = (-5, -2) \cup [0, \pi]$ is bounded above because 5.51 is an upper bound of it. The set $B = (-3.135, \infty)$ is bounded below because -100 is a lower bound of it. The set $C = [0, 1]$ is bounded because -1 and 1 are its lower and upper bound respectively.

Definition 5 (Supremum) *α is supremum of $S \subseteq \mathbb{R}$ if*

- *α is an upper bound of S.*
- *x is an upper bound of $S \implies \alpha \leq x$.*

Note that the second axiom in the definition is equivalent to $\forall\, \varepsilon > 0,\ \exists\, s \in S \ni \alpha - \varepsilon < s \leq \alpha$.

Example : $\sup A$(defined above) $= \pi$. $\sup \mathbb{Z}$ doesn't exist.

Definition 6 (Infimum) *β is infimum of $S \subseteq \mathbb{R}$ if*

- *β is an lower bound of S.*
- *y is an lower bound of $S \implies \beta \geq y$.*

Note that the second axiom in the definition is equivalent to $\forall\, \varepsilon > 0,\ \exists\, s \in S \ni \beta + \varepsilon > s \geq \beta$.

Example : $\inf \mathbb{N} = 1$. $\inf \mathbb{Q}$ doesn't exist.

1.6 Groups

Definition 7 (Binary operation) *A binary operation on a set S is a function $f : S \times S \to S$.*

Example : $f(a, b) = a + b$ on $\mathbb{Z}$.

Definition 8 (Groups) *A group $(G, \cdot)$ is a set G equipped with a binary operation $\cdot$ on it such that*

- *Associativity-* $\forall\, a, b, c \in G,\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$
- *Existence of identity -* $\exists\, e \in G \ni a \cdot e = e \cdot a = a \,\forall\, a \in G$. *($e$ is called identity of G)*
- *Existence of inverse-* $\forall\, a \in G\ \exists\, b_a \in G \ni a \cdot b_a = b_a \cdot a = e$.*($b_a$ is called inverse of a in G)*

Example : Set of Real numbers with addition, Set of all 3×3 real invertible matrices with matrix multiplication.

Definition 9 (Abelian groups) *A group $(G, \cdot)$ is abelian if the binary operation is commutative i.e.* $a \cdot b = b \cdot a \,\forall\, a, b \in G$

Example : Set of Real numbers with addition, Set of all 7×11 real matrices with addition.

Definition 10 (Field) *A field $(F, +, \times)$ is a set F equipped with two binary operations $+$ and $\times$ such that*

- $(F, +)$ *is an abelian group.*
- $(F \setminus e_{F+}, \times)$ *is an abelian group where e_{F+} is the identity of $(F, +)$.*
- $a \times (b + c) = a \times b + a \times c \,\forall\, a, b, c \in F$.

Example : $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$

1.7 Vector spaces

Definition 11 (Vector space) *A vector space V over a field F is an abelian group $(V, +)$ equipped with an action of F on V i.e. a map $* : F \times V \to V$ satisfying the following axioms*[1].

Note that $(F, +, \times)$ is a field. We write ab in place of $a \times b$. If we write $v_1 + v_2$,(where $v_1, v_2 \in V$) then this " $+$ " is of $(V, +)$. And if we write $a + b$,(where $a, b \in F$) then this " $+$ " is of $(F, +, \times)$. We also write av in place of $*(a, v)$ where $a \in F$ and $v \in V$. The identity of both $(F, +)$ and $(V, +)$ will be denoted by 0. From the context it will be understood which identity we are talking about. The identity of $(F, \times)$ will be denoted by 1. Now onwards we are going to use these simplified notations. We call elements of V as vectors and that of F as scalars.

Axioms[1]

- $(ab)v = a(bv) \; \forall \; a, b \in F \wedge v \in V$.
- $1(v) = v \; \forall \; v \in V$.
- $(a+b)v = av + bv \; \forall \; a, b \in F \wedge v \in V$.
- $a(u+v) = au + av \; \forall \; a \in F \wedge u, v \in V$.

Example : $\mathbb{R}^n$ over $\mathbb{R}$ is a vector space.

Linear independence

V is a vector space over F. The vectors $v_1, v_2, \cdots, v_n$ are said to be linearly independent if $\forall \; c_1, c_2, \cdots, c_n \in F,\; \Sigma_{i=1}^n c_i v_i = 0 \implies c_i = 0 \; \forall \; i \in \{1, 2, \cdots, n\}$.
The vectors $v_1, v_2, \cdots, v_n$ are said to be linearly dependent if $\exists \; c_1, c_2, \cdots, c_n \in F,\; \wedge \; \exists \; i_0 \in \{1, 2, \cdots, n\} \ni c_{i_0} \neq 0 \wedge \Sigma_{i=1}^n c_i v_i = 0$.

Example : $\mathbb{R}^2$ over $\mathbb{R}$ is a vector space. $(1, 0)$ and $(3, 7)$ are linearly independent. But $(1, 0)$ and $(3, 0)$ are linearly dependent.

Span

Let V be a vector space over F and $S \subseteq V$. Then $Span(S) = \{\Sigma_{i=1}^n c_i s_i \mid c_i \in F,\; s_i \in S,\; n \in \mathbb{N}\}$. Observe that $Span(S)$ is a vector space over F for all subset S of V.

Example : $Span(\pi, 0)$ in $\mathbb{R}^2$ over $\mathbb{R}$ is $x \; axis$.

Basis

Let V be a vector space over F. $B \subseteq V$ is a basis of V if

- The vectors in B are linearly independent.
- $Span(B) = V$.

The numbers of elements in B is called the dimension of V over$F$.

Example : $\{(0, 5), (-\pi, 0)\}$ is a basis of $\mathbb{R}^2$ over $\mathbb{R}$.

Linear Maps

Let V and W be vector spaces over F. $f : V \to W$ is linear if

- $f(v_1) + f(v_2) = f(v_1 + v_2) \; \forall \; v_1, v_2 \in V$.
- $cf(v) = f(cv) \; \forall \; v \in V \wedge c \in F$.

A linear functional on V is a linear map from V to F. (Note that F is a vector space over F).

Example : Consider $\mathbb{R}^2$ and $\mathbb{R}^3$ over $\mathbb{R}$. Define $f : \mathbb{R}^2$ to $\mathbb{R}^3$ by $f(x, y) = (2x + 3y, -7x, 9y)$. f is a linear map.

1.8 Inner product and norm

We won't define these rigorously here, rather we will see the properties of them. Among these properties if we take some (i.e.minimal set of axioms) they will uniquely determine/ well-define the object.

Inner product

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ or $\mathbb{C}$ is a function, called inner product. $\langle a, b\rangle$ is inner product of vectors a and b. Here are some useful properties of inner product. Note that $u, v, w \in V$ and $a \in F$ arbitrary where $F = \mathbb{R}$ or $\mathbb{C}$.

- $\langle v, v\rangle \geq 0$
- $\langle v, v\rangle = 0 \Leftrightarrow v = 0$
- $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$
- $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
- $\langle u, v\rangle = \overline{\langle v, u\rangle}$
- $\langle au, v\rangle = a\langle u, v\rangle$
- $\langle u, av\rangle = \bar{a}\langle u, v\rangle$

Norm

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. $\| \cdot \| : V \to \mathbb{R} \cup \{0\}$ is a function, called norm. For $u, v \in V$ and $a \in F$ arbitrary, it has the following properties.

- $\| v \| = 0 \Leftrightarrow v = 0$
- $\| av \| = |a| \, \| v \|$
- $\| u + v \| \leq \| u \| + \| v \|$
- $\| uv \| \leq \| u \| \| v \|$

A normed linear space is a vector space with a norm defined on it. An inner product space is a vector space with an inner product defined on it. An inner product space is a normed linear space with the norm defined by $\| v \| = \sqrt{\langle v, v\rangle}$

1.9 Hilbert space

Let (X, d_X) be a metric space. A sequence $\{a_n\}$ is convergent in X if there exists $l \in X$ such that $\lim\{a_n\} = l$.
A sequence $\{a_n\}$ is Cauchy in X if $\forall\, \varepsilon > 0\ \exists\, n_\varepsilon \in \mathbb{N} \ni d_X(a_n, a_m) < \varepsilon\ \forall\, n, m \geq n_\varepsilon$.
Note that, every convergent sequence is Cauchy. But the converse is not necessarily true. Take $X = (0, 1)$ with $d(a, b) = |a - b|$ and take $a_n = \dfrac{1}{n+1}$ for counterexample.
A metric space (X, d_X) is said to be complete it every Cauchy sequence in it has a limit $l \in X$ i.e. every Cauchy sequence is convergent in it.

Let $\mathbb{H}$ be a Vector space. It is indeed an inner product space. Observe that $\mathbb{H}$ has a natural metric on it defined by $d_{\mathbb{H}}(x, y) = \| x - y \| = \sqrt{\langle x - y, x - y\rangle}$. Here $-y$ is the additive inverse of $y \in \mathbb{H}$. If we say $\mathbb{H}$ is complete we mean this is complete as this metric space with this metric.

Definition 12 (Hilbert space) *A Hilbert space is a complete inner product space.*

Definition 13 (Banach space) *A Banach space is a complete normed linear space.*

Example : All the examples of this section will be described in the subsection $'\mathbb{R}^n$ *as a metric space'* under the next section.

1.10 Topology of metric spaces

Let (X, d_X) be a metric space. An open ball of radius r around a point $a \in X$ is $B(a, r) = \{x \in X \mid d_X(a, x) < r\}$. Similarly a closed ball of radius r around a point $a \in X$ is $\overline{B}(a, r) = \{x \in X \mid d_X(a, x) \leq r\}$.

Interior points and open sets

Let $A \subseteq X$. $a \in A$ is an interior point of A if $\exists\ \delta > 0 \ni B(a, \delta) \subset A$. A subset A of X is open if $\forall\ a \in A,\ a$ is an interior point of A. A subset B of X is closed if $X \setminus B$ is open.

Example : $B((0, 0), 1)$ is open in $\mathbb{R}^2$ and $(\frac{1}{2}, \frac{-1}{3})$ is an interior point of it. $[3, \infty)$ is closed in $\mathbb{R}$ because $(-\infty, 3)$ is open in $\mathbb{R}$.

Limit points and closed sets

Let $A \subseteq X$. $l \in X$ is a limit point of A if $\forall\ \delta > 0,\ B(l, \delta) \cap A \setminus \{l\} \neq \phi$. A subset A of X is closed, if all it's limit points belong to A. A subset B of X is open if $X \setminus B$ is closed.

Example : $\overline{B}((0, 0), 1)$ is closed in $\mathbb{R}^2$ and $(-1, 0)$ is a limit point of it. $(3, \infty)$ is open in $\mathbb{R}$ because $(-\infty, 3]$ is closed in $\mathbb{R}$.

Some related notations

$int(A)$ is the set of interior points of A. $D(A)$ is the set of limit points of A. Closure of $A = \overline{A} = A \cup D(A)$.

Example : In $\mathbb{R}$, $int((0, 1]) = (0, 1)$, $\overline{(0, 1]} = [0, 1]$ and $D(\mathbb{N}) = \phi$.

$\mathbb{R}^n$ as a metric space

There are a lot of distance functions that can give $\mathbb{R}^n$ a metric space structure. But for the future requirement we will only focus on Euclidean distance or l_2^n norm. Let $x = (x_1, x_2, \cdots, x_n)$ and $y = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$. Then $d(x, y) = (\Sigma_{i=1}^n (x_i - y_i)^2)^{\frac{1}{2}}$. So in $\mathbb{R}$ it becomes $|a - b| = d(a, b)$. As $\mathbb{R}^n$ is a vector space over $\mathbb{R}$, it becomes a Hilbert space with the Euclidean metric which can also be induced from the inner product defined by $\langle x, y\rangle = \Sigma_{i=1}^n x_i y_i$.

Bounded sets

Let (X, d_X) be a metric space. $A \subseteq X$ is called bounded if $\exists\ K \in \mathbb{R}^+ \ni d(a_1, a_2) < K\ \forall\ a_1, a_2 \in A$.

Example : Any subset of a discrete metric space.

Compact sets in $\mathbb{R}^n$

A subset A of $\mathbb{R}^n$ is compact if and only if A is closed and bounded.

Example : $\overline{B}((0,0),1) \cup \overline{B}((1,1),1)$ in $\mathbb{R}^2$

Connected sets

Let (X, d_X) be a metric space. A set $S \subseteq X$ is disconnected if there exists open sets $U$and V in X such that

- $U \neq \phi$ and $V \neq \phi$
- $U \cap V = \phi$
- $S = (U \cap S) \cup (V \cap S)$

A set is called connected if it is not disconnected.

Example : $(0,1)$ is connected in $\mathbb{R}$ and $\overline{B}((0,0),1)$ is connected in $\mathbb{R}^2$. $\mathbb{N} \times \mathbb{N}$ is disconnected in $\mathbb{R}^2$.

Part I
Differentiation

2 Differentiability on $\mathbb{R}^n$

First we recall the definition of differentiability on $\mathbb{R}$. Let U be an open set in $\mathbb{R}$. $f : U \to \mathbb{R}$ is differentiable at $a \in U$ if there exists $L \in \mathbb{R}$ such that

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

We call $L = f'(a)$ i.e. the derivative of f at the point a. So existence of derivative at a point a is same as saying $\lim_{h\to 0} \dfrac{f(a+h) - f(a)}{h}$ exists. Now let's think about $\mathbb{R}^n$ for $n > 1$. The first thing one can observe is that the expression $\dfrac{f(a+h) - f(a)}{h}$ doesn't make sense in $\mathbb{R}^n$ because a and h are elements/vectors of $\mathbb{R}^n$ and division by h is not well-defined.

The next thing that naturally comes to our mind is that we can define differentiability on $\mathbb{R}^n$ by $f'(a) = \lim_{h\to 0} \dfrac{\| f(a+h) - f(a) \|}{\| h \|}$. But in that case the new definition must be consistent with the definition of differentiability on $\mathbb{R}$. Note that according to our usual definition $f(x) = |x|$ is not differentiable at $0 \in \mathbb{R}$. The limit of $\dfrac{||0+h| - |0||}{h}$ as h approaches to 0 doesn't exist. The left and right hand limits are different. But the limit of $\dfrac{||0+h| - |0||}{|h|}$ exists and equal to 1 as h approaches to 0. Hence the newly assumed definition implies $|x|$ is differentiable at 0, which is not true. This definition is not consistent with the definition of differentiability on $\mathbb{R}$. We need to think of something else.

On $\mathbb{R}$ we have $f'(a) = \lim_{h\to 0} \dfrac{f(a+h) - f(a)}{h}$, which is same as saying $\lim_{h\to 0} \dfrac{f(a+h) - f(a) - f'(a)h}{h} = 0$. Now let us look at the object $f'(a)$ in a different way. Observe that any linear function L from $\mathbb{R}$ to $\mathbb{R}$ is of the form $L(x) = cx$ for some $c \in \mathbb{R}$. Here we can think of a linear function $Df_a : \mathbb{R} \to \mathbb{R}$ by $Df_a(x) = f'(a)x$. Now we state the definition of differentiability on $\mathbb{R}$ in the following way.

Let U be an open set in $\mathbb{R}$. $f : U \to \mathbb{R}$ is differentiable at $a \in U$ if there exists a linear map $Df_a : \mathbb{R} \to \mathbb{R}$ such that

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

The only non-trivial thing with this newly written definition is the modulus of h in its denominator. We can describe this using $\varepsilon - \delta$ definition of limit of a function. $\lim_{h\to 0} \dfrac{f(a+h) - f(a) - Df_a(h)}{h} = 0 \implies$ for all $\varepsilon > 0$ there exists $\delta > 0$ such that $|h| < \delta \implies \left| \dfrac{f(a+h) - f(a) - Df_a(h)}{h} \right| < \varepsilon \implies \left| \dfrac{f(a+h) - f(a) - Df_a(h)}{|h|} \right| < \varepsilon$. So

now we have $\lim_{h\to 0} \dfrac{f(a+h)-f(a)-Df_a(h)}{|h|} = 0.$
Our next goal is to trace the same idea in the definition of differentiability on $\mathbb{R}^n$.

Definition of differentiability on higher dimension

U is an open subset of $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$. We say f is differentiable at a point $a \in U \subseteq \mathbb{R}^n$ if there exists a linear map $Df_a : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

The linear map Df_a is called the derivative of f at a. Note that, derivative at a point is not a number any more, it's a linear map. We know that if T is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ then $T(v) = Av$ for some $m \times n$ matrix A and for all $v \in \mathbb{R}^n$(Proved later[$Exercise$ 15]). Here vectors of $\mathbb{R}^n$ are considered as column vectors or $n \times 1$ column matrix. Similarly for $\mathbb{R}^m$. Now we can visualize how the dimensions match after the matrix multiplication. i.e. $[T(v)]_{m\times 1} = [A]_{m\times n}[v]_{n\times 1}$. So now we will try to find the matrix of linear transformation Df_a. If $n = m = 1$ this is just an 1×1 matrix or number. We will talk about this matrix again when we will learn $Jacobian$. Some examples will be worked out then as well.

Proposition 14 *U is an open subset of $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$ and f is differentiable at a point $a \in U$. Then f is continuous at a.*

Proof(Hint). Since f is differentiable, $\parallel f(a+h)-f(a) \parallel = Df_{x_0}(h) + \parallel h \parallel E_{x_0}(h)$ for some $\lim_{h\to 0} E_{x_0}(h) = 0$. Now whenever $\parallel h \parallel < \delta$, we will have $\parallel h \parallel (E_{x_0}(h) + \parallel Df_{x_0} \parallel) < \varepsilon$. (Using the fact that *Norm of a linear map is finite on a finite dimensional vector space.*) [Defined below]
Norm of a linear map $T(: \mathbb{R}^n \to \mathbb{R}^m) =$

$$\parallel T \parallel = \sup_{v\in\mathbb{R}^n\setminus\{\vec{0}\}} \frac{\parallel T(v) \parallel}{\parallel v \parallel} = \sup_{v\in\mathbb{R}^n \wedge \parallel v \parallel = 1} \frac{\parallel T(v) \parallel}{\parallel v \parallel}$$

Exercise 15 *Let T is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. Show that*

1. *$T(v) = A_T v$ for some $m \times n$ matrix A_T and for all $v \in \mathbb{R}^n$ where vectors of $\mathbb{R}^n$ are considered as column vectors or $n \times 1$ column matrix. A_T is called the matrix of the linear map T.*

2. *$\parallel T(v) \parallel \leq \parallel T \parallel \parallel v \parallel$*

Hint. For the first part fix basis $\{e_i\}_{i=1}^n$ and $\{f_j\}_{j=1}^m$ of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Write $T(e_i) \in \mathbb{R}^m$ in terms of f_i s, like $T(e_i) = \Sigma_{j=1}^m a_{ji} f_j$. Then write $v = \Sigma_{i=1}^n v_i e_i$. Now use linearity of T and write $T(v)$ using coefficients a_{ij} and basis f_j. Second part is immediate from the definition of $\parallel T \parallel$.

Exercise 16 *Let f is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$. Show that f is differentiable at any point of $\mathbb{R}^n$. (Hint : $Df_a = f$)*

Let U, V, W be vector spaces over F. Then $U \times V$ is a vector space over F.(with component wise addition and scalar multiplication in each component). $f : U \times V \to W$. For $u \in U$, define $f_u : V \to W$ by $f_u(v) = f(u, v)$. Similarly define $f_v : U \to W$ for all $v \in V$. Now f is called bilinear if f_u and f_v are linear for all $u \in U$ and $v \in V$.

Exercise 17 *Let f is a bilinear function from $\mathbb{R}^a \times \mathbb{R}^b$ to $\mathbb{R}^c$. Show that f is differentiable at any point of $\mathbb{R}^a \times \mathbb{R}^b$. (Hint : $Df_{(a,b)}(t_1, t_2) = f(a, t_2) + f(t_1, b + t_2)$)*

Exercise 18 *Norm of a linear map is finite on a finite dimensional vector space.*

Hint. An n dimensional vector space is same as $\mathbb{R}^n$. So without loss of generality assume $T : \mathbb{R}^n \to \mathbb{R}^m$. Then use the fact that $T(v) = Av$ where A is a $m \times n$ matrix. Try to see $\| T \|$ in terms of entries of A. As A has finitely many entries take the maximum among them and find some bound for the norm.

3 Directional and partial derivative, Jacobian

3.1 Directional derivative

Assume when h approaches to 0 the quantity $\dfrac{f(a+h) - f(a)}{h}$ has a limit in $\mathbb{R}$. We can talk about left hand and right hand limit depending on $h > 0$ or $h < 0$. In $\mathbb{R}$ we can approach a only through two different paths. But if $a \in \mathbb{R}^n$ for some $n > 1$, the situation is different. There we can approach a through infinitely many different paths.

There can be a case when $\lim_{h\to 0} \dfrac{f(a+h) - f(a)}{h}$ doesn't exist but left and right hand limit exist separately. If and only if they are equal, we have the overall existence of limit at a. Similarly in $\mathbb{R}^n$ we can define a notion of directional derivative for each of the infinitely many directions like left and right derivative on $\mathbb{R}$.

Definition 19 (Directional derivative) *Let U be an open subset of $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$ and $x_0 \in U$. Let $u \in \mathbb{R}^n$ such that $\| u \| = 1$. Then the derivative of f at x_0 to the direction of u is $D_u f(x_0) = \lim_{h\to 0} \dfrac{f(x_0 + hu) - f(x_0)}{h}$ if the limit exists.*

Note that $D_u f(x_0) = Df_{x_0}(u)$ i.e. the directional derivative of f at x_0 to the direction of u is same as the value of derivative of f at x_0 evaluated at u. This is an important result we will need later. Proving this is easy by the straightforward use of definitions of derivative and directional derivative.

3.2 Partial derivative

As $\mathbb{R}^n$ is a vector space of dimension n, it has a basis $e_1, e_2, \cdots e_n$. Here $e_1 = (1, 0, \cdots, 0)$, $e_2 = (0, 1, 0, \cdots, 0)$ and so on.
Let U be an open subset of $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$ and $x_0 \in U$. Partial derivatives of f at x_0 are nothing but $D_{e_1} f(x_0), D_{e_2} f(x_0), \cdots, D_{e_n} f(x_0)$. We denote $D_{e_k} f(x_0)$ by $\dfrac{\partial f}{\partial x_k}(x_0)$.

Example

Define $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x, y) = 2x + 3y^2$.
Now $D_{e_1} f(x_0, y_0) = \lim_{h\to 0} \dfrac{f[(x_0, y_0) + h(1, 0))] - f(x_0, y_0)}{h} = \lim_{h\to 0} \dfrac{2h}{h} = 2$. Hence we have $\dfrac{\partial f}{\partial x}(x_0, y_0) = 2$. The object $\dfrac{\partial f}{\partial x_k}$ can be informally considered as derivative of f with respect to variable x_k where other variables are constant in the expression of f. Similarly, one can check $\dfrac{\partial f}{\partial x}(x_0, y_0) = 6y_0$ for this example.

3.3 Jacobian

Lemma 20 *Let V is a finite dimensional inner product space. Let $e_1, e_2, \cdots, e_n$ be an orthonormal basis of V. i.e. $\langle e_i, e_j\rangle = 0 \ \forall\ i \neq j$ and $\langle e_i, e_i\rangle = 1\ \forall\ i$. Then for all vector $v \in V$, $v = \Sigma_{i=1}^n \langle v, e_i\rangle e_i$.*

Proof. As $\{e_i\}_{i=1}^n$ is a basis of V, let $v = \Sigma_{i=1}^n v_i e_i$. Now we see $\langle v, e_i\rangle = v_i$. ∎

Now onwards whenever we will talk about $\mathbb{R}^n$, we will assume the standard orthonormal basis (denoted by $\{e_i^{(n)}\}_{i=1}^n$). Here $e_i^{(n)}$ is an n-tuple with 1 at i^{th} place and 0 at the rest of the places. If $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(x) \in \mathbb{R}^m$ can be written as $(f_1(x), f_2(x), \cdots, f_m(x))$. Now we define natural projections $\pi_i : \mathbb{R}^m \to \mathbb{R}$ by $v \mapsto \langle v, e_i^{(m)}\rangle$ which is same as saying $\pi_i(a_1, a_2, \cdots, a_m) = a_i$. So f_i s can be considered as functions from $\mathbb{R}^n$ to $\mathbb{R}$ by $f_i = \pi_i \circ f$. Assume f is differentiable. Recalling the technique used in $Exercise$ 15 and using $Lemma$ 20 we understand the entries of the matrix of the linear map Df_{x_O} and the following is the matrix of it.

$$A = \begin{bmatrix} \langle Df_{x_0}(e_1^{(n)}), e_1^{(m)}\rangle & \langle Df_{x_0}(e_2^{(n)}), e_1^{(m)}\rangle & \cdots & \langle Df_{x_0}(e_n^{(n)}), e_1^{(m)}\rangle \\ \langle Df_{x_0}(e_1^{(n)}), e_2^{(m)}\rangle & \cdots & & \langle Df_{x_0}(e_n^{(n)}), e_2^{(m)}\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle Df_{x_0}(e_1^{(n)}), e_m^{(m)}\rangle & \cdots & & \langle Df_{x_0}(e_n^{(n)}), e_m^{(m)}\rangle \end{bmatrix}$$

Lemma 21 *With everything well-defined, $\langle Df_{x_0}(e_i^{(n)}), e_j^{(m)}\rangle = \dfrac{\partial f_j}{\partial x_i}(x_0)$*

Proof(Hint).
Use the definition of partial derivative and the fact that $D_u f(x_0) = Df_{x_0}(u)$.

Let U be an open subset of $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$. Let f is differentiable at a point $a \in U$. Jacobian of f is a $m \times n$ matrix such that evaluation of this matrix at a gives the matrix of the linear map Df_a. We denote it by $\mathcal{J}f$ and evaluation of it at a by $\mathcal{J}f(a)$. By evaluation we mean the entry $\dfrac{\partial f_j}{\partial x_i}$ evaluated at a is $\dfrac{\partial f_j}{\partial x_i}(a)$.

$$\mathcal{J}(f) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & & & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & & \vdots \\ & & & \ddots & \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_{n-1}} & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

If $m = 1$ i.e. $f : \mathbb{R}^n \to \mathbb{R}$, we denote $\mathcal{J}f$ by ∇f (called gradient of f) and

$$\nabla f = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}]$$

Example

Define $f : \mathbb{R}^2 \to \mathbb{R}^3$ by $f(x, y) = (x, y, x^2 + y^2)$. Let $h = (h_1, h_2)$. Fix any point $(x_0, y_0) \in \mathbb{R}^2$.

So $\mathcal{J}f(x_0, y_0) = \begin{bmatrix} \frac{\partial}{\partial x}(x) & \frac{\partial}{\partial y}(x) \\ \frac{\partial}{\partial x}(y) & \frac{\partial}{\partial y}(y) \\ \frac{\partial}{\partial x}(x^2+y^2) & \frac{\partial}{\partial y}(x^2+y^2) \end{bmatrix}(x_0, y_0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2x_0 & 2y_0 \end{bmatrix}$.

Now, $\lim_{h\to 0} \dfrac{f((x_0, y_0) + (h_1, h_2) - f(x_0, y_0) - Df_{(x_0,y_0)}(h_1, h_2)}{\sqrt{h_1^2 + h_2^2}} =$

$\lim_{h\to 0} \dfrac{\begin{pmatrix} h_1 \\ h_2 \\ 2x_0h_1 + 2y_0h_2 + h_1^2 + h_2^2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2x_0 & 2y_0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}}{\sqrt{h_1^2 + h_2^2}} = \lim_{h\to 0} \begin{bmatrix} 0 \\ 0 \\ \sqrt{h_1^2 + h_2^2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

This shows us how understanding the matrix of the linear map using $Jacobian$ helps us to show the differentiability of a given function.

Proposition 22 *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a, all it's partial derivatives $\frac{\partial f}{\partial x_i}(a)$ exist.*

Proof : Trivial and left as an exercise to the reader. ■

Note that the converse of $Proposition$ 22 is not true in general. Following is the counter example.

Exercise 23 *Define $f : \mathbb{R}^2$ to $\mathbb{R}$ by*

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

Prove that all the partial derivatives of f at $(0, 0)$ exist but f is not differentiable at $(0, 0)$.

Hint. Showing the existence of partial derivatives is straightforward. For showing it's not differentiable assume on the contrary it is. Then there exists a linear map $l : \mathbb{R}^2$ to $\mathbb{R}$ such that

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

. Here $a = (0,0)$. Let $h = (h_1, h_2)$. So we have $f(h_1,h_2) = f(0,0) + l(h_1,h_2) + E(h_1,h_2)\sqrt{h_1^2 + h_2^2}(*)$ for some $\lim_{h\to 0} E(h) = 0$. Use equation $(*)$ to calculate the values of $l(h,0), l(0,h), l(h,h)$. Now use linearity of l and take limit h approaches to 0 both side. Reach a contradiction.

The next theorem gives us a sufficient condition for which existence of all partial derivatives at a point implies differentiability at that point.

Theorem 24 *U is a open set in $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$ is differentiable at $x_0 \in U$. All the partial derivatives $\dfrac{\partial f}{\partial x_i}$ of f exist in some small neighbourhood of x_0. Out of n many partial derivatives at x_0, any $n-1$ of these are continuous at x_0. i.e. $\exists\, i_1, i_2, \cdots, i_{n-1} \in \{1, 2, \cdots, n\} \ni \dfrac{\partial f}{\partial x_{i_1}}, \dfrac{\partial f}{\partial x_{i_2}}, \cdots, \dfrac{\partial f}{\partial x_{i_{n-1}}}$ are continuous at x_0. Then f is differentiable at x_0.*

Proof(Hint): Let $\{e_i\}_{i=1}^n$ be the usual orthonormal basis of $\mathbb{R}^n$. WLOG $m = 1$ (It will be evident from the method of the proof that the same way we can prove for $m > 1$). WLOG assume $\dfrac{\partial f}{\partial x_i}$ is continuous for $i \in \{2, 3, \cdots, n\}$. (If $\dfrac{\partial f}{\partial x_i}$ is not continuous we take a new basis $(f_1, f_2, \cdots, f_n)$ where $f_1 = e_i$ and rest $n-1$ of e_i s are f_2 to f_n). So we have to show

$$\lim_{H\to 0} \frac{f(x_0 + H) - f(x_0) - \mathcal{J}f(x_0)(H)}{\parallel H \parallel} = 0$$

. Let $u = \dfrac{H}{\parallel H \parallel}$ and $\parallel H \parallel = h$. So $H = hu$ and assume $u = (u_1, u_2, \cdots, u_n)$. Write

$$f(x_0 + H) - f(x_0) = \Sigma_{k=1}^n (f(x_0 + \Sigma_{i=1}^k h u_i e_i) - f(x_0 + \Sigma_{i=1}^{k-1} h u_i e_i)) \cdots (*)$$

and

$$\mathcal{J}f(x_0)(H) = \Sigma_{k=1}^n h u_i \frac{\partial f}{\partial x_i}(x_0) \cdots (**)$$

Now to calculate $f(x_0 + H) - f(x_0) - \mathcal{J}f(x_0)(H)$ in numerator use the summand wise subtraction of $(*) - (**)$. Then take limit h goes to 0 and use continuity of $\dfrac{\partial f}{\partial x_i}$ at x_0.

4 Chain rule, Mean value theorem

In $\mathbb{R}$ we use a formula $(f \circ g)'(x) = f'(g(x))g'(x)$ if f and g are differentiable on $\mathbb{R}$. Now we will formalise and generalise this concept in higher dimension.

4.1 Chain rule

U and V are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Let $g : U \to \mathbb{R}^m$ and $g(U) \subseteq V$. Now $f : V \to \mathbb{R}^k$. So we can now define $\psi = f \circ g : U \to \mathbb{R}^k$. If g is differentiable at $x_0 \in U$ and f is differentiable at $g(x_0) \in V$, then ψ is differentiable at $x_0 \in U$ and

$$D\psi_{x_0} = Df_{g(x_0)} Dg_{x_0}$$

Proof : Trivial using the definition of derivative or the concepts of $Jacobian$. ■

4.2 Mean value theorem (MVT)

MVT on $\mathbb{R}$

Let $f : (a,b) \to \mathbb{R}$ be differentiable on (a,b). Then for all $x_1, x_2 \in (a,b)$ there exists $c \in (x_1, x_2)$ such that $f(x_1) - f(x_2) = f'(c)(x_1 - x_2)$.
Note that in $\mathbb{R}$ if we say c is in between x_1 and x_2, it makes sense. But in $\mathbb{R}^n$ we need something analogous so that this will be well defined.

MVT on $\mathbb{R}^n$

Let U be open in $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$ is differentiable on U. Let $x, y \in U$ such that $tx + (1-t)y \in U$ for all $t \in [0,1]$. This essentially says that the line segment joining x and y lies inside U. Then for all $v \in \mathbb{R}^m$ there exists $t_v \in [0,1]$ such that $\langle v, f(y) - f(x)\rangle = \langle v, Df_{z_v}(y-x)\rangle$ where $z_v = t_v x + (1-t_v)y$ i.e. there exists a point z_v in the line segment joining x and y such that the above equality holds.

Proof(Hint) : Using open set U find a suitable interval $\mathcal{I}$ on $\mathbb{R}$ so that if $x, y \in \mathcal{I}$ then $tx + (1-t)y \in \mathcal{I} \ \forall \ t \in [0,1]$. The function $F : [0,1] \to \mathbb{R}$ by $F(t) = \langle v, f(x + t(y-x))\rangle$ is well-defined. Now using chain rule observe how DF will look like in terms of Df and then apply MVT (on $\mathbb{R}$) to the function F and for the points 0 and 1.

Exercise 25 *Define $f : \mathbb{R}^2 \to \mathbb{R}^3$ by $(x,y) \mapsto (x, y^2, x+3y)$. Show f is differentiable on $B((0,0),2)$ and find a suitable z_v (Notation from the previous section) for $x = (0,1)$ and $y = (1,0)$ (**Hint :** For first part use linearity of f and $exercise$ 16, second part is just calculation)*

Theorem 26 *Let U be open and connected in $\mathbb{R}^n$. $f : U \to \mathbb{R}^m$ is differentiable on U and for all $x \in U$, $Df_x = 0$ if and only if f is constant.*

Proof(Hint) : Showing the backward direction is straightforward. The forward direction is an application of MVT and the fact ($exercise$ 27). Fix $x_0 \in U$ and consider $A = \{x \in U \mid f(x) = f(x_0)\}$. Show that A is both open and closed subset of U and from $exercise$ 27 we have $A = U$.

Exercise 27 *Let X be a metric space with some distance function. $A \subseteq X$. A is a closed set as well as an open set. Show that $A = \phi$ or X. (This is an advanced exercise, if you are unable to prove it, take this statement for granted)*

5 Higher order derivatives, Taylor's theorem

5.1 Higher order derivatives

First talking about $\mathbb{R}$ we see, there is a meaning of $f'(c)$ for a point $c \in \mathbb{R}$. But now we will consider a function $f' : U \to \mathbb{R}$ by $x \mapsto f'(x)$ where U is the set of points where f is differentiable. After that we talk about differentiability of the function f' to get f''. Let $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ be the set of all linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. This can be identified with set of all $m \times n$ matrices because every linear transformation has it's matrix and for every matrix A, we have a linear transformation $v \mapsto Av$. Again this $m \times n$ matrices can be identified with $\mathbb{R}^{mn}$ by an obvious map i.e. writing down all the entries of the matrix one by one as a mn tuple.

Let U is open in $\mathbb{R}^n$ and $f : U$ to $\mathbb{R}^m$ is differentiable on U. So now onwards for such a function f, and for an element $a \in U$, we look at the object Df_a as an element of $\mathbb{R}^{mn}$. (by the identification described earlier). We define $Df : U \to \mathbb{R}^{mn}$ by $Df(x) = Df_x$. Now we can talk about the differentiability of that function. Note that from now we will use the notation Df_a and $Df(a)$ simultaneously depending on what the context demands. Both of them are the same object represented by two different notations.

So second derivative of f at a is derivative of Df at a if it exists. i.e. $D^2 f_a = D(Df)_a$ In this way we can define $D^{n+1} f_a = D(D^n f)_a$.

5.2 Hessian

Now we can also define second order partial derivatives by

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial(\frac{\partial f}{\partial x_j})}{\partial x_i} = \frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})$$

Let U is open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is differentiable at a. Then the $Hessian$ of f at $a = \mathcal{H}f(a)$ is the following matrix evaluated at a. By evaluation we mean the entry $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ gives $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$.

$$\mathcal{H}f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & & & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & & \vdots \\ & & & \ddots & \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_{n-1} \partial x_n} & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

5.3 Commutativity of second order partial derivatives

For all $i \neq j$ we have $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{\partial}{\partial x_i}(\dfrac{\partial f}{\partial x_j})$ and $\dfrac{\partial^2 f}{\partial x_j \partial x_i} = \dfrac{\partial}{\partial x_j}(\dfrac{\partial f}{\partial x_i})$. Now the question natural to ask is whether they are equal. The answer is obviously no. But it is highly non-trivial. The following exercise is the counterexample for the commutativity of second order partial derivatives.

Exercise 28 *$f : \mathbb{R}^2 \to \mathbb{R}$ by*

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

*Show that $\dfrac{\partial}{\partial x}(\dfrac{\partial f}{\partial y})(0,0) \neq \dfrac{\partial}{\partial y}(\dfrac{\partial f}{\partial x})(0,0)$ (**Hint.** Calculate to show one of them is 1 and another one is -1).*

Now through next two theorems state what condition make them commute. Note that the proofs are purely analytical and not that trivial. We have omitted the proofs. Interested readers may try to finish it on their own. Also we will use the notation $\dfrac{\partial^2 f}{\partial x_i^2}$ for $\dfrac{\partial^2 f}{\partial x_i \partial x_i}$.

Theorem 29 (Claircut's theorem) *Let U is open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is differentiable at x_0. Assume $i \neq j$. In some open ball around x_0 $\dfrac{\partial f}{\partial x_i}, \dfrac{\partial f}{\partial x_j}$ and $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ exist. And $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ is continuous at x_0. Then $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(x_0)$ exists and $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(x_0) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$*

Theorem 30 *Let U is open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is differentiable at x_0. Assume $i \neq j$. In some open ball around x_0 $\dfrac{\partial f}{\partial x_i}, \dfrac{\partial f}{\partial x_j}$ exist and they are differentiable at x_0. Then $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$ and $\dfrac{\partial^2 f}{\partial x_j \partial x_j i}(x_0)$ exist and $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(x_0)$.*

Exercise 31 *Define $f : \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x,y) = (-y,x)$. Calculate $\dfrac{\partial^2 f}{\partial x \partial y}(1,2)$ and $\dfrac{\partial^2 f}{\partial y \partial x}(1,2)$. Also calculate $D^2 f_{(1,2)}$ if it exists.*

5.4 Taylor's theorem

Taylor's theorem on $\mathbb{R}$

U is open in $\mathbb{R}$. f is differentiable upto order $\leq k$ on U. Let $x, y \in U$ such that $[x,y] \subset U$. Then there exists $z \in [x,y]$ such that $f(y) - f(x)$ =

$$(y-x)f'(x) + \frac{(y-x)^2}{2} f''(x) + \frac{(y-x)^3}{3!} f'''(x) + \cdots + \frac{(y-x)^{k-1}}{(k-1)!} f^{(k-1)}(x) + \frac{(y-x)^k}{k!} f^{(k)}(z)$$

Some useful notations

U is open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$. Let $t = (t_1, t_2, \cdots, t_n)^{tr} \in \mathbb{R}^n$. By writing tr we mean the vector is not a row, rather a column. We have already seen that we always consider vectors of $\mathbb{R}^n$ as $n \times 1$ column matrix. Now for $x_0 \in \mathbb{R}^n$ we denote as follows (if they exists)

$$f'(x_0, t) = \nabla f(x_0)t = \Sigma_{i=1}^n \frac{\partial f}{\partial x_i}(x_0)t_i$$

$$f''(x_0, t) = \Sigma_{i=1}^n \Sigma_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)t_j t_i$$

$$f'''(x_0, t) = \Sigma_{i=1}^n \Sigma_{j=1}^n \Sigma_{k=1}^n \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x_0)t_k t_j t_i$$

In such way we can have a notion of $f^k(x_0, t)$. Next thing we define formally is the line joining two vectors x and $y \in \mathbb{R}^n$. Define $[x, y] = \{\lambda x + (1 - \lambda)y \mid \lambda \in [0, 1]\}$.

Taylor's theorem on $\mathbb{R}^n$

U is open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ such that on U all its partial derivatives exist upto order $< k$ and they are differentiable. Let $x, y \in U$ such that $[x, y] \in U$. Then there exists $z \in [x, y]$ such that $f(y) - f(x) =$

$$f'(x, y-x) + \frac{1}{2}f''(x, y-x) + \frac{1}{3!}f'''(x, y-x) + \cdots + \frac{1}{(k-1)!}f^{(k-1)}(x, y-x) + \frac{1}{k!}f^{(k)}(z, y-x)$$

Proof (Hint): Using open set U find a suitable interval $\mathcal{I}$ on $\mathbb{R}$ so that if $x, y \in \mathcal{I}$ then $tx + (1-t)y \in \mathcal{I}\ \forall\ t \in [0, 1]$. The function $F : [0, 1] \to \mathbb{R}$ by $F(t) = f(x + t(y - x)$ is well-defined. Now observe how $f^{(n)}(x, y - x)$ will look like in terms of $f^{(n)}(t)$ and then apply Taylor's theorem (on $\mathbb{R}$) to the function F and for the points 0 and 1.

Exercise 32 *Define $f : \mathbb{R}^3$ to $\mathbb{R}$ by $f(x, y, z) = sin(x) + e^y + z$. Show that all order partial derivatives of f exist. And expand this around $(0, 0, 0)$.*

6 Maxima, Minima

Exercise 33 *U is open in $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ is differentiable on U. $a \in U$. Assume there exists $\delta > 0$ such that $f(x) \geq f(a)$ for all $x \in B(a, \delta) \subseteq U$. Prove that $Df(a) = 0$. (i.e. the derivative at a local minima of a function is 0)*

Hint : Look at the Jacobian or gradient of f. We show each entry of them must be 0. f is differentiable hence partial derivatives exist. So both the left hand and right hand limit of the quantity $\dfrac{f(a + he_i^{(n)}) - f(a)}{h}$ as h approaches to 0 must be equal. But for $h > 0$ this is > 0 and for $h < 0$ this is < 0. So the limit must be 0.

In this chapter we will learn to use some similar techniques for finding maxima, minima we use in $\mathbb{R}$. From $Exercise$ 33 it is evident that the derivative of a function at its local maxima and minima is 0.

Definition 34 (The class $\mathcal{C}^k$) *$f : \mathbb{R}^n \to \mathbb{R}$. We say f is of class $\mathcal{C}^k$ or $f \in \mathcal{C}^k(\mathbb{R}^n)$ if all the partial derivatives of f exist upto order $\leq k$ and they are continuous.*

Exercise 35 *Prove or disprove : If $f : \mathbb{R} \to \mathbb{R}$ is differentiable upto order $\leq k$, then $f \in \mathcal{C}^k(\mathbb{R})$*

Hint : Give a counter example by defining

$$f(x) = \begin{cases} x^{k+1} sin(\dfrac{1}{x}) & if\ x \neq 0 \\ 0 & if\ x = 0 \end{cases}$$

Note that the same thing can be shown for $f : \mathbb{R}^n$ to $\mathbb{R}$ by taking this example in one component and taking the rest all to be 0.

Definition 36 (Smooth function) *$f : \mathbb{R}^n \to \mathbb{R}$. We say f is smooth if $f \in \mathcal{C}^k(\mathbb{R}^n)$ for all $k \in \mathbb{N}$.*

Our discussions of maxima, minima are limited to all those functions which are of class $\mathcal{C}^2$. U is open in $\mathbb{R}^n$. $f : U \to \mathbb{R}$. We say f has a local minima at $a \in U$ if there exists $\delta > 0$ such that $f(x) \geq f(a)$ for all $x \in B(a, \delta) \subseteq U$. Similarly f has a local maxima at $b \in U$ if there exists $\kappa > 0$ such that $f(x) \leq f(b)$ for all $x \in B(b, \kappa) \subseteq U$. $x_0 \in U$ is a saddle point of f if for all $\varepsilon > 0$ there exists $y_\varepsilon, z_\varepsilon \in B(x_0, \varepsilon)$ such that $f(y_\varepsilon) > f(x_0)$ and $f(z_\varepsilon) < f(x_0)$. $x_0 \in U$ is a critical/stationary point of f if $Df(x_0) = 0$.

Let $\mathcal{H}f$ be the Hessian of f, for any $h \in \mathbb{R}^n$ define

$$\mathcal{Q}_{x_0}(h) = h^{tr}[\mathcal{H}f(x_0)]h$$

where h is a column vector of size n and h^{tr} is a row vector of size n.
Recall the maxima-minima criterion in $\mathbb{R}$

- $f'(a) = 0$ and $f''(a) > 0 \Rightarrow f$ has a local minima at $a \in \mathbb{R}$.
- $f'(a) = 0$ and $f''(a) < 0 \Rightarrow f$ has a local maxima at $a \in \mathbb{R}$.

Now we proceed to find some similar criterion in $\mathbb{R}^n$

Theorem 37 *U is open in $\mathbb{R}^n$. $f : U \to \mathbb{R}$. There exists $x_0 \in U$ such that $Df(x_0) = 0$. Then*

- *$\mathcal{Q}_{x_0}(h) > 0$ for all $h \in \mathbb{R}^n \Rightarrow x_0$ is a local minima of f.*
- *$\mathcal{Q}_{x_0}(h) < 0$ for all $h \in \mathbb{R}^n \Rightarrow x_0$ is a local maxima of f.*
- *There exists $h_1, h_2 \in \mathbb{R}^n$ such that $\mathcal{Q}_{x_0}(h_1) > 0$ and $\mathcal{Q}_{x_0}(h_2) < 0 \Rightarrow x_0$ is a saddle point of f.*

Now we will find some other useful criterion as well. Let T be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. We say $\lambda \in \mathbb{R}$ is an eigen value of T if there exists $v \in \mathbb{R}^n$ such that $T(v) = \lambda v$. v is called the eigen vector of T corresponding to the eigen value λ. Now as set of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$ i.e. $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ is nothing but set of $n \times n$ real matrices, we can define the eigen value and eigen vector of a matrix in the same way. (Matrix A corresponds to the linear transformation $T(v) = Av$). We define determinant of a linear transformation/a matrix by $det(A)$ = product of all eigen values of A. Now we define two more quantities which will help later to compute maxima and minima of a given function

$$\lambda_{min}^{(x_0)} = \inf_{h \in \mathbb{R}^n \wedge \|h\|=1} \mathcal{Q}_{x_0}(h) \quad and \quad \lambda_{max}^{(x_0)} = \sup_{h \in \mathbb{R}^n \wedge \|h\|=1} \mathcal{Q}_{x_0}(h)$$

Lemma 38 *$\lambda_{min}^{(x_0)}$ is the minimum eigen value of the matrix $\mathcal{H}f(x_0)$ and $\lambda_{max}^{(x_0)}$ is the maximum eigen value of the matrix $\mathcal{H}f(x_0)$*

Since $f \in \mathcal{C}^2(\mathbb{R}^n)$ we have the function $\mathcal{Q}_{x_0} : \mathbb{R}^n \to \mathbb{R}$ by $h \mapsto \mathcal{Q}_{x_0}(h)$ continuous. Now it is evident that $\mathcal{Q}_{x_0}(h) > 0$ for all $h \in \mathbb{R}^n \Leftrightarrow \lambda_{min}^{(x_0)} > 0$.

From the definition of $\lambda_{min}^{(x_0)}$ we just get that this quantity is ≥ 0. To justify the fact that $\lambda_{min}^{(x_0)} \neq 0$ we need to use the continuity of $\mathcal{Q}_{x_0}$. We see $\{h \in \mathbb{R}^n :\| h \|= 1\}$ is compact in $\mathbb{R}^n$. So continuous image of a compact set is compact in $\mathbb{R}$. Hence the infimum will belong to the set of images. ($Exercise$ 39)

Similarly we have $\mathcal{Q}_{x_0}(h) < 0$ for all $h \in \mathbb{R}^n \Leftrightarrow \lambda_{max}^{(x_0)} < 0$. Also $\exists\ h_1, h_2 \in \mathbb{R}^n$ such that $\mathcal{Q}_{x_0}(h_1) > 0$ and $\mathcal{Q}_{x_0}(h_2) < 0 \Leftrightarrow \lambda_{min}^{(x_0)} < 0 < \lambda_{max}^{(x_0)}$.

Exercise 39 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. If K is a compact subset of $\mathbb{R}^n$, show that $f(K)$ is compact subset of $\mathbb{R}$. If F is a compact subset of $\mathbb{R}$, show that $\sup F,\ \inf F \in F$. (**Hint :** Compact sets are closed and bounded and vice-versa. Then use the definition (of closed set) which uses limit point)*

Theorem 40 *(Restating maxima-minima using eigen values of Hessian) U is open in $\mathbb{R}^n$. $f : U \to \mathbb{R}$. There exists $x_0 \in U$ such that $Df(x_0) = 0$. $\lambda_{min}^{(x_0)}$ and $\lambda_{max}^{(x_0)}$ are the minimum eigen value and maximum eigen value of the matrix $\mathcal{H}f(x_0)$ respectively. Then*

- $\lambda_{min}^{(x_0)} > 0 \Rightarrow x_0$ *is a local minima of f.*
- $\lambda_{max}^{(x_0)} < 0 \Rightarrow x_0$ *is a local maxima of f.*
- $\lambda_{min}^{(x_0)} < 0 < \lambda_{max}^{(x_0)} \Rightarrow x_0$ *is a saddle point of f.*

Concept of principal minor

Let A be an $n \times n$ matrix. For $1 \leq k \leq n$ we define its k^{th} principal minor A_k by taking the first k rows and columns of A. i.e. if $A = (a_{ij})_{n \times n}$ then A_k is a $k \times k$ matrix defined by $A_k = (a_{ij})_{k\times k}$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & & a_{1n} \\ a_{21} & \cdots & & & a_{2n} \\ \vdots & \vdots & \ddots & & \vdots \\ & & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{(n)(n-1)} & a_{nn} \end{bmatrix}_{n\times n} \quad A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & \cdots & & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}_{k\times k}$$

Theorem 41 *U is open in $\mathbb{R}^n$. $f : U \to \mathbb{R}$. There exists $x_0 \in U$ such that $Df(x_0) = 0$. Let $\mathcal{A}$ denote the matrix $\mathcal{H}f(x_0)$. For any matrix $\mathbb{M}$, $\mathbb{M}_k$ denotes the k^{th} principal minor of $\mathbb{M}$ and $det(\mathbb{M})$ denotes the determinant of it. Let $\breve{n}$ denote the set of all natural numbers $\leq n$. Then*

- *There exists $k \in \breve{n/2}$ such that $det(\mathcal{A}_{2k}) < 0 \Leftrightarrow x_0$ is a saddle point of f.*
- *For all $k \in \breve{n}$, $det(\mathcal{A}_k) > 0 \Leftrightarrow x_0$ is a local minima of f.*
- *For all $k \in \breve{n}$, $(-1)^k det(\mathcal{A}_k) > 0 \Leftrightarrow x_0$ is a local maxima of f.*
- *$det(\mathcal{A}_n) = 0 \Leftrightarrow x_0$ is a degenerate critical point.*

Exercise 42 *For each of the following functions find all its critical points, local maxima, local minima and saddle points using any of $theorems - 37, 40, 41$ to your accordance. If you find a degenerate critical point, these theorems won't be applicable any more. In that case use some basic analysis or inequalities to find whether it is maxima or minima or saddle point.*

- $f_1 : \mathbb{R}^2 \to R$ by $f_1(x, y) = x^4 + x^2y + y^2$
- $f_2 : \mathbb{R}^2 \to R$ by $f_2(x, y) = 12x^3 + y^3 + 12x^2y - 75y$
- $f_3 : \mathbb{R}^3 \to R$ by $f_3(x, y, z) = 3x^2y - yz^2 - 4xz + 7$
- $f_4 : \mathbb{R}^2 \to R$ by $f_4(x, y) = x^3 + y^2 - 6xy$

7 Implicit and Inverse function theorem

Before starting this section we will write two equations which can be seen as non-rigorous, not well defined and nonsense to us. But it will seem so natural that it will be a true statement for an intuitionist.

$$\frac{\partial y}{\partial x}\frac{\partial x}{\partial y} = 1 \quad and \quad \frac{\partial y}{\partial x} = \frac{\frac{\partial z}{\partial x}}{\frac{\partial z}{\partial y}}$$

Now we will proceed to the theory to see how can we formalise this intuition.

7.1 Implicit function theorem

We will always identify $\mathbb{R}^n \times \mathbb{R}^m$ with $\mathbb{R}^{n+m}$ for simplicity. i.e.
$((a_1, a_2, \cdots, a_n), (b_1, b_2, \cdots, b_m))$ can be regarded as $(a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_m)$.

Theorem 43 *Let Ω be an open subset of $\mathbb{R}^n \times \mathbb{R}^m$. $F : \Omega \to \mathbb{R}^m$ is of class $\mathcal{C}^1$. There exists $(x_0, y_0) \in \Omega$ such that $F(x_0, y_0) = 0$.*
If $\{e_i^{(n)}\}_{i=1}^n$ and $\{e_i^{(m)}\}_{i=1}^m$ forms a basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, then $\{\tilde{e}_i^{(n)}\}_{i=1}^n \cup \{\tilde{e}_i^{(m)}\}_{i=1}^m$ forms a basis of $\mathbb{R}^n \times \mathbb{R}^m$, where $\tilde{e}_i^{(n)}$ is $e_i^{(n)}$ followed by m many zeroes and $\tilde{e}_i^{(m)}$ is $e_i^{(m)}$ after n many zeroes. This is same as saying $\{e_i^{(n+m)}\}_{i=1}^{n+m}$ forms a basis of $\mathbb{R}^{n+m}$.
Note that we denote $D_{\tilde{e}_j^{(n)}}F(x_0, y_0)$ by $\dfrac{\partial F}{\partial x_j}(x_0, y_0)$ and $D_{\tilde{e}_j^{(m)}}F(x_0, y_0)$ by $\dfrac{\partial F}{\partial y_j}(x_0, y_0)$. One can informally think that we are naming the variables of $\mathbb{R}^n \times \mathbb{R}^m$ e.g. the variables of $\mathbb{R}^{n+m}$ as $x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m$.
Let the $m \times m$ matrix $\mathcal{M}(x_0, y_0)$ is invertible where the ij^{th} entry of this is $\dfrac{\partial F_i}{\partial y_j}(x_0, y_0)$.
A square matrix M is invertible if there exists another matrix N such that $MN = NM = I$ where I is the identity matrix of the same order.

Then there exists an open set $U \in \mathbb{R}^n$ containing x_0 and an open set $\tilde{U} \in \mathbb{R}^m$ containing y_0 such that

- *For all $x \in U$ there is unique $y_x = f(x) \in \tilde{U}$ with $F(x, f(x)) = 0$.*
- *$f(x_0) = y_0$ i.e. y_0 can be written explicitly as a functional value of x_0.*
- *$\mathcal{J}f(x) = -[\mathcal{M}(x, f(x))]^{-1}_{m\times m}[a_{ij}(x, f(x))]_{m\times n}$ where $a_{ij}(x, f(x)) = \dfrac{\partial F_i}{\partial x_j}(x, f(x))$ for all $x \in U$.*

Proof(Hint)

The proof of the first part i.e. the existence of such f is highly non-trivial.
So the proof will be omitted. Those who are interested can try to think about it.
But assuming the existence of such f we prove the last part.
Define $\psi : U \to \mathbb{R}^n \times \mathbb{R}^m$ by $x \mapsto (x, f(x))$. Now $g = F \circ \psi$. Use chain rule on g.
We have $Dg(x) = DF(\psi(x))D\psi(x)(equation\ 1)$.
Observe that $\mathcal{J}\psi = \begin{bmatrix} I_{n\times n} \\ ---------- \\ \mathcal{J}f_{m\times n} \end{bmatrix}_{(n+m)\times n}$, $\mathcal{J}F = \begin{vmatrix} (\dfrac{\partial F}{\partial x_i})_{m\times n} & (\dfrac{\partial F}{\partial y_i})_{m\times m} \end{vmatrix}_{m\times(n+m)}$.
Also observe that $\mathcal{J}g = \begin{bmatrix} 0 \end{bmatrix}_{m\times n}$. Put these values in $equation\ 1$ and obtain the final required equality.

7.2 Inverse function theorem

Theorem 44 *Let Ω be an open subset of $\mathbb{R}^n$ and $F : \Omega \to \mathbb{R}^n$ is of class $\mathcal{C}^1$. There exists $x_0 \in \Omega$ such that $\mathcal{J}F(x_0)$ is invertible. Then there exists*

- *An open set $U \in \mathbb{R}^n$ containing x_0*
- *An open set $V \in \mathbb{R}^n$ containing $F(x_0)$*

- *A function $G : V \to U$ of class $\mathcal{C}^1$*
- *$G(F(x)) = x$ for all $x \in U$*

such that the following are satisfied

- *$F(G(y)) = y$ for all $y \in V$*
- *$\mathcal{J}G(y) = [\mathcal{J}F(G(y))]^{-1}$ for all $y \in V$*

Proof (Hint) : For existence of G use Implicit function theorem defining another map $\Phi : U \times \mathbb{R}^n \to \mathbb{R}^n$ by $\Phi(x, y) = F(x) - y$. Now use chain rule on the identity map $G \circ F$ on U.

Definition 45 (Diffeomorphism) *Let U and V be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. A differentiable map $f : U \to V$ is called diffeomorphism if f is a bijection and $f^{-1} : V \to U$ is also differentiable. Two sets are called diffeomorphic if there exists a diffeomorphism from one of them to another.*

Exercise 46 *Prove or disprove : There is a diffeomorphism f from some open subset of $\mathbb{R}^n$ to an open subset of $\mathbb{R}^n$ such that $det[\mathcal{J}f(x)] = 0$ for some x in its domain. You can assume $det(AB) = det(A)det(B)$.*

We won't solve a lot of problems in this section. Rather we will see some applications of these two theorems to Physics. Note that, we are not going to take care of any notation or terminology used in the next subsection (7.3). Our purpose is not to rigorously focus on the topics described in next subsection. Rather our goal is to show how these two theorems are used. Finding the meaning or definition of the things used there is left as an exercise to the reader. Readers who are already aware of the problems they will relate better to the use of these theorems.

7.3 Some applications to Physics

The Implicit Function Theorem is frequently used in mechanics. For example, in the construction of canonical transformations in analytical mechanics.

For another example, take the motion of a point mass m, subject to a force law f generated by an infinitely differentiable potential V satisfying the condition

$$DV(\xi) \neq 0 \quad for \quad each \quad \xi \neq 0$$

The existence of an infinitely differentiable inverse function is guaranteed constructively by our Inverse Function Theorem. This situation arises in the solution of certain Cauchy problems, such as the damped pendulum

$$m\ddot{x} + \lambda\dot{x} + ksin(x) = \gamma f(t) \quad (t \in \mathbb{R}^+) \cdots\cdots (*)$$

where $\lambda \in \mathbb{R}$; λ, m, k are positive ; $\lambda^2 \neq 4km$ and periodic and infinitely differentiable. If the oscillations are small, then the damped linear oscillator admits isochronous periodic motions with forcing term. More precisely, if f is infinitely differentiable with period $T > 0$, then, provided γ is small enough, there exists a periodic motion, with period T, satisfying $(*)$. Together with $(*)$ let us consider the linearised equation

$$m\ddot{x} + \lambda\dot{x} + kx = \gamma f \cdots\cdots (**)$$

which admits a periodic solution $\tilde{x}$ isochronous with f. We look for a periodic solution of $(*)$ of the form

$$x(t) = \gamma\tilde{x}(t) + y(t) \ \ (t \in \mathbb{R}^+)$$

with initial data

$$y(0) = \varepsilon \ ; \ \dot{y}(0) = \eta \cdots\cdots (***)$$

We set

$$x(T) = \gamma\tilde{x}(T) + a(\varepsilon, \eta, \gamma) \ \ and \ \ \dot{x}(T) = \gamma\dot{\tilde{x}}(T) + b(\varepsilon, \eta, \gamma)$$

Since $\tilde{x}(0) = \tilde{x}(T)$ and $\dot{\tilde{x}}(0) = \dot{\tilde{x}}(T)$ by the periodicity of $\tilde{x}$ the condition that $(*)$ admits a periodic solution with period T can be restated as $a(\varepsilon, \eta, \gamma) = \varepsilon$ and $b(\varepsilon, \eta, \gamma) = \eta$. But the solvability of the equations

$$\begin{cases} a(\varepsilon, \eta, \gamma) = \varepsilon \\ b(\varepsilon, \eta, \gamma) = \eta \end{cases}$$

is equivalent to the existence of a periodic solution of $(*)$ with period T. We now have the problem of expressing ε and η as functions of a sufficiently small γ.
This is possible if the Jacobian determinant of partial derivatives with respect to ε and η is non zero. The computation of these partial derivatives is based on the equations

$$\begin{cases} a(\varepsilon, \eta, \gamma) = y(T) \\ b(\varepsilon, \eta, \gamma) = \dot{y}(T) \end{cases}$$

where $y(t)$ is a solution of the Cauchy problem

$$m\ddot{y}(t) + \lambda\dot{y}(t) + ky(t) = k(\gamma\tilde{x}(t) + y(t) - sin(\gamma\tilde{x}(t) + y(t)))$$

with initial conditions $(***)$. Finally, it is easy to see that the corresponding Jacobian determinant equals

$$(e^{a_+T} - 1)(e^{a_-T} - 1) \neq 0$$

where a_+ and a_- are respectively, the positive and negative parts of a. So the Cauchy problem has a constructive solution given by our Implicit Function Theorem.

8 Lagrange Multiplier

This is a fairly small section, may be the smallest. We state one simple result and see applications. In maxima-minima section we learnt to find local maxima and minima of a $\mathcal{C}^1$ function. But now we will find maxima-minima of a function given some constraints. For example, in $\mathbb{R}^2$ let we have to find the minimum value of $x+2y$ given the constraint $xy = 3$. So we need to find the minima of $f(x, y) = x + 2y$ restricted to the set $\{(x, y) \in \mathbb{R}^2 \mid xy - 3 = 0\}$. Now we state the theorem, then methods, then we apply this to various interesting problems.

Theorem 47 *$f : \mathbb{R}^n \to \mathbb{R}$ is of class $\mathcal{C}^1$. $m \leq n$. $g : \mathbb{R}^n \to \mathbb{R}^m$ is also of class $\mathcal{C}^1$. $\mathfrak{M} = \{x \in \mathbb{R}^n \mid g(x) = 0\}$. (Note that $\mathfrak{M}$ is a subset of $\mathbb{R}^n$). Let a be a critical point of $f\mid_{\mathfrak{M}}$. i.e. $a \in \mathfrak{M}$ such that $Df(a) = 0$. Then*

$$\nabla f(a) = \Sigma_{i=1}^m \lambda_i \nabla g_i(a) \ \ for \ \ some \ \ \lambda_i \in \mathbb{R}$$

The proof of this theorem requires a lot of new concepts from differential geometry and linear algebra. We don't focus on that but interested readers may try to think about it after figuring out details or just the definition of some concepts like level set, n-Surface, Vector field, tangent field, parametrized curve, orthogonal complement of a subspace of a vector space etc. We don't discuss those here. We try to find out the proper method to apply it.

The set up of our problem will be the following. Or given any problem we need to create the following set up first.
We are given a $f : \mathbb{R}^n \to \mathbb{R}$ and we have to find maximum or minimum value of f when the domain is restricted to $\{x \in \mathbb{R}^n \mid g(x) = 0\}$.So we also have a $g : \mathbb{R}^n \to \mathbb{R}^m$ and $m \leq n$.
The method/algorithm we follow to solve the problem is the following. We find the $1 \times n$ row vector ∇f. Similarly find the $1 \times n$ row vectors ∇g_i for all $1 \leq i \leq m$. We put these in the equation $\nabla f = \Sigma_{i=1}^{m} \lambda_i \nabla g_i$. Now from n many components we find n many equations. We solve them by eliminating λ_i s. And find the point a such that $Df(a) = 0$. We have a critical point. Now we use our usual methods from section 6 to find whether it's a maxima, minima or saddle point.

Problems

Problem 48 *Prove AM - GM inequality. Let $x_1, x_2, \cdots, x_n$ be positive real numbers. Then*

$$\frac{\Sigma_{i=1}^{n} x_i}{n} \geq (\Pi_{i=1}^{n} x_i)^{\frac{1}{n}}$$

Hint : As x_i s are positive real numbers, we have $x_i = y_i^2$ for some $y_i \in \mathbb{R}^+$. Define $f : \mathbb{R}^n \to \mathbb{R}$ by $(y_1, y_2, \cdots, y_n) \mapsto y_1^2 y_2^2 \cdots y_n^2$ and $g : \mathbb{R}^n \to \mathbb{R}$ by $(y_1, y_2, \cdots, y_n) \mapsto \Sigma_{i=1}^{n} y_i^2$.

Problem 49 *Find the shortest distance from the ellipse $x^2 + 2y^2 = 2$ to the line $x + y = 2$. Hint : Define $f : \mathbb{R}^4 \to \mathbb{R}$ by $(x, y, u, v) \mapsto (u - x)^2 + (v - y)^2$ and $g : \mathbb{R}^2 \to \mathbb{R}$ by $(x, y) \mapsto (x^2 + 2y^2 - 2, x + y - 2)$.*

Problem 50 *The temperature at the point (x, y, z) in the $3 - dimentional$ space is given by $f(x, y, z) = xy + z^2$. Find the hottest and coolest point on the sphere $x^2 + y^2 + z^2 = 2z$.*

Problem 51 *Find the norm of the matrix $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$*
Hint : Find the maxima of $\dfrac{\parallel Av \parallel}{\parallel v \parallel}$ given $\parallel v \parallel = 1$ where $v \in \mathbb{R}^2$.

Problem 52 *For a given cylindrical can of unit volume, what should be its radius and length for it to have the least surface area ?*

Part II
Integration

9 Integration on $\mathbb{R}^n$

9.1 Integration on $\mathbb{R}$

First we look at the analytical meaning of integration (namely Riemann Integration) on $\mathbb{R}$.

Definition 53 (Partition) *Let $[a, b] \subset \mathbb{R}$. A partition $P[a, b]$ is a finite collections of points of $[a, b]$ in increasing order with the endpoints a and b i.e. $P[a, b] : a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ for some $n \in \mathbb{N}$ and $t_i \in [a, b]$.*

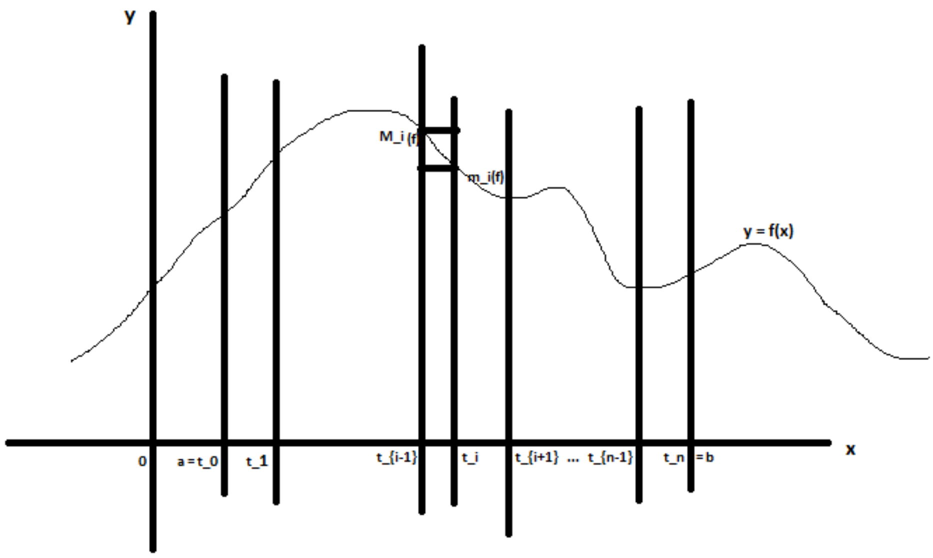

Figure 1: Partition of [a,b]

Let $f : [a, b] \to \mathbb{R}$ be a bounded function. P is a partition of the interval $[a, b]$ and $P : a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$. For all $1 \leq i \leq n$ define

$$m_i(f) = \inf_{x \in [t_{i-1}, t_i]} f(x) \quad and \quad M_i(f) = \sup_{x \in [t_{i-1}, t_i]} f(x)$$

So we can think of rectangular strips of length $t_i - t_{i-1}$ for all i. Now if we take the height of this strip as $m_i(f)$ it gives us a lower value of the area than the actual area under the curve $y = f(x)$ in the region $[t_{i-1}, t_i]$. Similarly, if we take the height of this strip as $M_i(f)$ it gives us a higher value of the area than the actual area under the curve $y = f(x)$ in the region $[t_{i-1}, t_i]$. Now we analytically try to minimise these errors in the value of area, take sum over all the areas in a partition, find the area under the curve

$y = f(x)$ in $[a, b]$. And we denote that by $\int_a^b f(x)dx$.
For the partition P of $[a, b]$ we now define two more things.

- Lower sum of f determined by $P = \mathcal{L}(f, P) = \Sigma_{i=1}^n m_i(f)(t_i - t_{i-1})$
- Upper sum of f determined by $P = \mathcal{U}(f, P) = \Sigma_{i=1}^n M_i(f)(t_i - t_{i-1})$

Till now we defined things for a fixed partition P of $[a, b]$. But there can be a lot of different partitions of $[a, b]$ as well. Let $\mathbb{P}[a, b]$ denote the set of all partitions of the interval $[a, b]$. Then we define two new things.

- Lower integral of f over $[a, b] = \int_{\frown a}^b f = \sup_{P \in \mathbb{P}[a,b]} \mathcal{L}(f, P)$
- Upper integral of f over $[a, b] = \int_a^{\smile b} f = \inf_{P \in \mathbb{P}[a,b]} \mathcal{U}(f, P)$

Definition 54 *$f : [a, b] \to \mathbb{R}$ is a bounded function. We say f is integrable over $[a, b]$ if $\int_{\frown a}^b f = \int_a^{\smile b} f$. In that case we also say*

$$\int_a^b f(x)dx = \int_{\frown a}^b f = \int_a^{\smile b} f$$

For example we can take zero function, constant function and identity function. In all of these cases calculating $\mathcal{L}(f, P)$ and $\mathcal{U}(f, P)$ is pretty easy for a particular partition P of $[a, b]$. One can try to find lower and upper integrals from them and show that they are equal. But this definition is not always helpful in terms of proving a function is integrable. So we look at some other equivalent criterion which can be proved from this basic definition ($definition$ 54). But we will consider them too as definitions and use if necessary. Interested readers may try to prove the equivalences of the definitions.

Definition of integrability

Definition 55 (Mesh of a partition) *Let P be a partition of the interval $[a, b]$ and $P : a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$. Mesh of this partition $= \| P \| = \sup_{i=1}^n (t_i - t_{i-1})$.*

Note that all the criterion we are going to discuss here are theorems itself and all of them are equivalent. But for our convenience we will consider them as definition. We follow all the notations already described in this section. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. f is integrable over $[a, b]$ if

- $\varepsilon - \delta$ **Criterion :** There exists $\mathfrak{I} \in \mathbb{R}$ such that for all $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that
$$\forall\ P \in \mathbb{P}[a, b]\ ,\ \| P \| < \delta_\varepsilon \implies | \mathcal{L}(f, P) - \mathfrak{I} | < \varepsilon$$
and $\int_a^b f(x)dx = \mathfrak{I}$.
- **Darboux Criterion :** Already mentioned in $Definition$ 54.
- **Cauchy Criterion :** For all $\varepsilon > 0$ there exists $P_\varepsilon \in \mathbb{P}[a, b]$ such that
$$\mathcal{U}(f, P_\varepsilon) - \mathcal{L}(f, P_\varepsilon) < \varepsilon$$

Riemann sum Criterion :

First we define Riemann Sum of f determined by a partition $P : a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$. For any choice of $t_i^* \in [t_{i-1}, t_i]$ we define $\mathcal{R}(f, P) = \Sigma_{i=1}^n f(t_i^*)(t_i - t_{i-1})$. Till now this Riemann sum is not well defined. There can be many values of it depending on the choice of t_i^* s. Now f is integrable over $[a, b]$ if there exists $\mathfrak{I} \in \mathbb{R}$ such that for all $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that

$$\forall\ P \in \mathbb{P}[a, b]\ ,\ \parallel P \parallel < \delta_\varepsilon \implies \mid \mathcal{R}(f, P) - \mathfrak{I} \mid < \varepsilon$$

for any choice of $\mathcal{R}(f, P)$ s. and $\int_a^b f(x)dx = \mathfrak{I}$. So now one can show that if f is integrable over $[a, b]$ and $\parallel P \parallel \to 0$ i.e. if $n \to \infty$ the quantity $\mathcal{R}(f, P)$ is a unique number irrespective of the choice of t_i^* s and $\int_a^b f(x)dx = \lim_{n\to\infty} \mathcal{R}(f, P)$.

Exercise 56 *$f : [a, b] \to \mathbb{R}$ is continuous. Prove that f is integrable over $[a, b]$. Also prove if $g : [a, b] \to \mathbb{R}$ is differentiable everywhere then g is integrable over $[a, b]$.* ***(Hint :*** *For first part use epsilon-delta definition of continuity and any of your favourite definition of integrability which uses epsilon-delta. For second part use the fact that differentiable functions are continuous)*

9.2 Integration on $\mathbb{R}^n$

In the previous subsection we talked about integrations on $\mathbb{R}$ over intervals. Similarly we will define rectangles/boxes in $\mathbb{R}^n$ analogous to intervals in $\mathbb{R}$. A rectangle/box $\Theta \in \mathbb{R}^n$ is a set $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ where $[a_i, b_i]$ is an interval in $\mathbb{R}$ for all $1 \le i \le n$. In this section we will discuss integrations on $\mathbb{R}^n$ over boxes/rectangles. Note that in $\mathbb{R}^2$ boxes are actual rectangles. In $\mathbb{R}^3$ they are cuboids. In general we will always call them rectangles in $\mathbb{R}^n$.

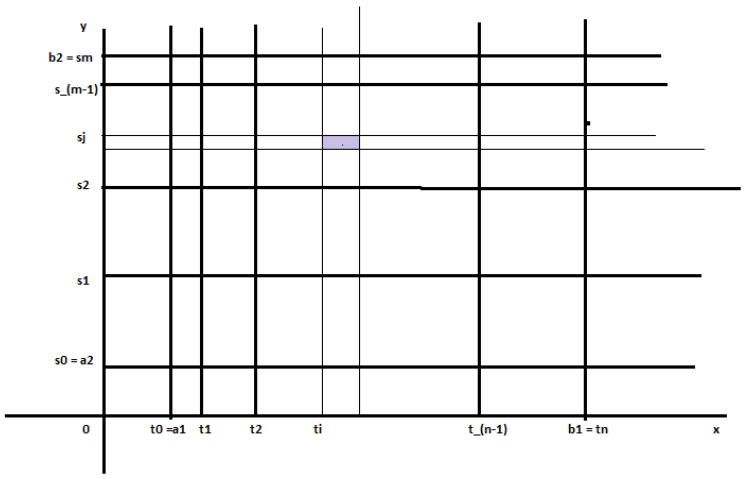

Figure 2: Sub-rectangles in $\mathbb{R}^2$

As we had partition in $\mathbb{R}$, we should be able to find something analogous here. Before that we look at the partition in $\mathbb{R}$ once again. Let $P \in \mathbb{P}[a, b]$ and $P : a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$. We call each small intervals $[t_{i-1}, t_i]$ as a subinterval T_i of

P. So another way to define partition P is collection of subintervals $\{T_i\}_{i=1}^n$ where end point of T_{i-1} is the initial point of T_i and initial point of T_1 is a, endpoint of T_n is b.

Let $\Theta = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ be a rectangle in $\mathbb{R}^n$. Now a partition P of Θ is of the form $P_1 \times P_2 \times \cdots \times P_n$ where $P_i \in \mathbb{P}[a_i, b_i]$ is a partition of $[a_i, b_i]$ in $\mathbb{R}$ for all $1 \leq i \leq n$. So a sub-box/ sub-rectangle R of P is of the form $R_1 \times R_2 \times \cdots \times R_n$ where R_i is a subinterval of P_i for all $1 \leq i \leq n$.

In figure-2 the shaded portion is a sub-rectangle of P of $\mathbb{R}^2$. Note that $R = T_{i+1} \times S_j$ where T_{j+1} is a subinterval of $\mathbb{R}$ and S_j is also a subinterval of $\mathbb{R}$.

If a subinterval R_i is of the form $[c, d]$ in $\mathbb{R}$ we say it's length is $d - c$. Similarly if a sub-rectangle R in $\mathbb{R}^n$ is of the form $[c_1, d_1] \times [c_2, d_2] \times \cdots \times [c_n, d_n]$ then it's volume is defined as $\mathcal{V}(R) = (d_1 - c_1)(d_2 - c_2) \cdots (d_n - c_n)$.

Now we are ready to define integration on higher dimension. Let Θ be a rectangle in $\mathbb{R}^n$ and $f : \Theta \to \mathbb{R}$ be a bounded function. Now $P = \{R_i\}_{i=1}^k$ be a partition of Θ where R_i s are sub-rectangles of the partition P. Then for all $1 \leq i \leq k$ define

$$m_i(f) = \inf_{x \in R_i} f(x) \quad and \quad M_i(f) = \sup_{x \in R_i} f(x)$$

Similarly the way we defined on $\mathbb{R}$ here also we define

- Lower sum of f determined by $P = \mathcal{L}(f, P) = \Sigma_{i=1}^k m_i(f)\mathcal{V}(R_i)$
- Upper sum of f determined by $P = \mathcal{U}(f, P) = \Sigma_{i=1}^k M_i(f)\mathcal{V}(R_i)$

And also we have the concepts of lower and upper integrals in the same way. Let $\mathbb{P}(\Theta)$ denote the set of al partitions of Θ. Then

- Lower integral of f over $\Theta = \int_{\Theta \frown} f = \sup_{P \in \mathbb{P}(\Theta)} \mathcal{L}(f, P)$
- Upper integral of f over $\Theta = \int_{\Theta}^{\smile} f = \inf_{P \in \mathbb{P}(\Theta)} \mathcal{U}(f, P)$

Here we define the **mesh of a partition** P (where $P = \{R_i\}_{i=1}^k$) as

$$\parallel P \parallel = \sup_{i=1}^{k} \mathcal{V}(R_i)$$

. Now with this set up we are ready to define the definitions and equivalent theorems of integrability on $\mathbb{R}^n$. We finish the whole theory of integration on $\mathbb{R}^n$ by stating the fact that all the theorems and definitions of the previous section holds with a small change i.e. this new meaning of lower and upper sum, lower and upper integrals and mesh of a partition. Our goal is now to go back and see $Definition$ 54 once again followed by $epsilon - delta$, $Cauchy$, $Darboux$ and $Riemann\ sum$ criterion. Interested readers may write the same thing again on $\mathbb{R}^n$. We avoid the copy-paste job. Note that we denote the integration of f over Θ by $\int_\Theta f$.

Exercise 57 $\Theta = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ *and following functions are defined from* Θ *to* $\mathbb{R}$. *Check whether they are integrable over* Θ *and if so calculate* $\int_\Theta f_i$ *for all of the following problems.*

- $$f_1(x,y) = \begin{cases} 1 & if \ \ (x,y) = (0,0) \\ 0 & otherwise \end{cases}$$

- $$f_2(x,y) = \begin{cases} 1 & if \ \ x \in \mathbb{Q}, \ y \in \mathbb{Q} \\ 0 & otherwise \end{cases}$$

- $$f_3(x,y) = \begin{cases} 1 & if \ \ x = 0, \ y \in \mathbb{Q} \\ 0 & otherwise \end{cases}$$

- $$f_4(x,y) = \begin{cases} 1 & if \ \ x = y \\ 0 & if \ \ x \neq y \end{cases}$$

Continuity implies Integrability

Statement : $f : \Theta \to \mathbb{R}$ is a continuous function where Θ is a rectangle in $\mathbb{R}^n$. Then f is integrable over Θ.
In $exercise$ 56 it was proved for $n = 1$. Now we are in a more general set up. This can be proved using the hints of the $exercise$ 56 and the following fact.
We define a stronger version of continuity namely uniform continuity. Let $f : A \to \mathbb{R}$ where $A \subseteq \mathbb{R}^n$. We say f is uniformly continuous on A if for all $\varepsilon > 0$ there exists δ such that $\forall \ x, y \in A, \| \ x - y \ \| < \delta \implies | \ f(x) - f(y) \ | < \varepsilon$. So the difference of this definition from the usual definition of continuity is that the value of δ here is independent of the point where we are considering the continuity of the function.
Now we state another important result. Let A be a compact subset of $\mathbb{R}^n$. If $f : A \to \mathbb{R}$ is continuous on A, then f is uniformly continuous on A.
Now our closed rectangle Θ is compact in $\mathbb{R}^n$. So use the fact f is uniformly continuous and use the hints given.

More on partitions

We didn't study some properties of partitions in $\mathbb{R}$ so that we can do them now in a more general set up. We first define **Refinement of a Partition**.
First we define our partition of a rectangle Θ as a collection of elements of Θ in a particular order. For example, in $\mathbb{R}$, partition of $[a,b]$ is P and $P : a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$. i.e. P is collection of t_i s. In this set up we define refinement as following. Let P and Q be two different partitions of Θ. We say Q is a refinement of P if $P \subseteq Q$. It's just that we obtain Q by some more fine partitions of $[a,b]$ than what we already had in P. Sometimes we also use the phrase that Q is finer than P.
Now we define our partition of a rectangle Θ as a collection of sub-rectangles in Θ. So in this set up Q is finer than P if and only if all the sub-rectangles of Q are sub-rectangle of some sub-rectangles of P. Now we see some lemmas/propositions in the form of exercises.

Exercise 58 *Θ is a rectangle in $\mathbb{R}^n$ and $f : \Theta \to \mathbb{R}$ be a bounded function. Let $P, Q \in \mathbb{P}(\Theta)$ be two arbitrary partitions. Then*

- *Prove that $\mathcal{L}(f,P) \leq \mathcal{U}(f,Q)$*
- *If Q is a refinement of P show that $\mathcal{L}(f,P) \leq \mathcal{L}(f,Q)$ and $\mathcal{U}(f,P) \geq \mathcal{U}(f,Q)$*

Hint : *Follows from definition*

This exercise gives an intuition that if we keep making the partition finer and finer the value of $\mathcal{L}(f,P)$ s increase and approach towards it's supremum i.e. $\int f$. Similarly, making the partition finer decreases the value of $\mathcal{U}(f,P)$ s and reach $\int f$ slowly. Now making a partition is finer and finer means making the mesh of the partition smaller and smaller. Recall the $Riemann\ sum\ criterion$. We do exactly the same thing i.e. $n \to \infty$ (making the partition finer in limiting sense) or $\parallel P \parallel \to 0$ (making the mesh smaller in limiting sense).

Special case of Riemann sum

Our next and final goal of this section is to review a special case of Riemann sum criterion which will be useful later. We will restrict our discussion upto three variables. The concept is same for n variables.

Let $f : [a,b] \to \mathbb{R}$ be bounded. Let $n \in \mathbb{N}$. Define $\Delta x = \dfrac{b-a}{n}$. Our partition is $a, a+\Delta x, a+2\Delta x, \cdots, a+n\Delta x = b$. Hence the mesh of the partition is Δx. And now we choose a representative from each of these n many intervals. For simplicity let's choose the midpoint of the interval. Let $\overline{x_i}$ denote the representative from the i^{th} interval. So, for all $1 \leq i \leq n$ we have $\overline{x_i} = a + (i - \frac{1}{2})\Delta x$. Now we define our Riemann sum explicitly by $\mathcal{R}(f) = \Sigma_{i=1}^{n} f(\overline{x_i})\Delta x$. If f is integrable over $[a,b]$, from Riemann sum criterion we have

$$\int_a^b f(x)dx = \lim_{n\to\infty} \Sigma_{i=1}^{n} f(\overline{x_i})\Delta x$$

.

Now in $\mathbb{R}^2$ let $A = [a,b] \times [c,d]$ be a rectangle. $f : A \to \mathbb{R}$ is integrable over A. Let $m, n \in \mathbb{N}$. Define $\Delta A = (\dfrac{b-a}{m})(\dfrac{d-c}{n})$. So we have partition P of $[a,b]$ as $a, a+\Delta x, a+2\Delta x, \cdots, a+m\Delta x = b$ where $\Delta x = \dfrac{b-a}{m}$. We also have partition Q of $[c,d]$ as $c, c+\Delta y, c+2\Delta y, \cdots, c+n\Delta y = d$ where $\Delta y = \dfrac{d-c}{n}$. So, $\Delta A = \Delta x \Delta y$. In the same manner we take midpoints $\overline{x_i}$ of intervals of P and midpoints $\overline{y_i}$ of intervals of Q. Now from our Riemann Sum criterion we have (follow some new notations)

$$\int_A f dA = \int_c^d \int_a^b f(x,y)dxdy = \lim_{m,n\to\infty} \Sigma_{j=1}^{n} \Sigma_{i=1}^{m} f(\overline{x_i}, \overline{y_j})\Delta x \Delta y$$

Now in $\mathbb{R}^3$ with a new variable z, let $V = [a,b] \times [c,d] \times [e,f]$ be a cuboid and $g : V \to \mathbb{R}$ be integrable over V. So $\Delta V = \Delta x \Delta y \Delta z$. And we have

$$\int_V g dV = \int_e^f \int_c^d \int_a^b g(x,y,z)dxdydz = \lim_{m,n,p\to\infty} \Sigma_{k=1}^{p} \Sigma_{j=1}^{n} \Sigma_{i=1}^{m} g(\overline{x_i}, \overline{y_j}, \overline{z_k})\Delta x \Delta y \Delta z$$

10 Measure on $\mathbb{R}^n$

Measure theory is an important tool of analysis. It gives rise to a complete new concept. We can define integrations in a more general set-up after knowing something called

$Lebesgue\ measure$. But here our goal is not to develop the theories of Lebesgue integration. We just learn something about $Lebesgue\ measure$ and see the Lebesgue criterion of Riemann integrability. From the last section we know that continuous functions are integrable. But after knowing Lebesgue criterion we will find a larger class of integrable functions. Before that we define some basic concepts of sets.

- A set A is finite if there exists a bijection $f : A \to \{1, 2, 3, \cdots, n\}$ for some $n \in \mathbb{N}$.
- A set is infinite if it is not finite.
- A set A is countably infinite if there exists a bijection $g : A \to \mathbb{N}$.
- A set is countable if it is finite or countably infinite.
- A set is uncountable if it is not countable.

10.1 Measure on $\mathbb{R}$

Let's think about intervals in $\mathbb{R}$. If we have (a, b) or $[a, b]$ or $(a, b]$, we can think of a small line segment joining points a and b on real axis. And the length of this line segment is nothing but $b - a$. So there are some subsets of $\mathbb{R}$ whose lengths can be defined. Till now we have seen a very small class of sets namely intervals whose length can be defined. But the next question naturally comes to our mind is whether we can define length of any arbitrary subset of $\mathbb{R}$. And this encourages us to define a concept of measure/length/Lebesgue measure. To define measure of any subset we use what we already know i.e. length of intervals.

First we define an open cover of a subset A of $\mathbb{R}$. Let $\{U_i\}_{i\in I}$ be collections of open subsets of $\mathbb{R}$ where I is an arbitrary index set. It can be finite or infinite, countable or uncountable. We say $U = \cup_{i\in I} U_i$ an open cover of A if $A \subseteq U$. For example $\cup_{n\in\mathbb{N}}(n - \frac{1}{3}, n + \frac{1}{3})$ is an open cover of $\mathbb{N}$.

For simplicity now we will take our U_i as open intervals which are just special open sets and we will denote them by I_i s. So if our open set is $I_i = (a, b)$, we define length of it $l(I_i) = b - a$. Now let $A \subseteq \mathbb{R}$ and we denote the set of all open covers of A by $\mathcal{O}(A)$. So if $T \in \mathcal{O}(A)$ we have $T = \cup_{i\in I} I_i$ where I_i s are open intervals of $\mathbb{R}$ and $A \subseteq T$.

Now for open cover $T = \cup_{i\in I} I_i$ we think of an sum (this can be finite sum or infinite series) $\Sigma_{i\in I} l(I_i)$. Note that this sum may exist or may not. We now define (with an informal notation) $m(T) = \Sigma_{i\in I} l(I_i)$. Either $m(T)$ doesn't exist or $m(T) \in \mathbb{R}$ if it exists. Our final goal is to find a suitable cover T of A such that this quantity $m(T)$ can analytically measure the length of the set A.

We are now ready to define a new concept. Let $A \subseteq \mathbb{R}$. The Lebesgue measure of A is defined as

$$m^*(A) = \inf_{T\in\mathcal{O}(A)} m(T)$$

Now we don't solve any problems or don't see any example, rather we directly jump to define it on $\mathbb{R}^n$. After that we will solve problems and see examples.

10.2 Measure on $\mathbb{R}^n$

Let $A \subseteq \mathbb{R}^n$. The definition of open cover is same as $\mathbb{R}$. But we don't have intervals any more. So in place of I_i s on $\mathbb{R}$ we now take R_i s on $\mathbb{R}^n$ where R_i s are open rectangles in $\mathbb{R}^n$. We say R is an open rectangle of $\mathbb{R}^n$ if $R = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$ where (a_i, b_i) is an open interval of $\mathbb{R}$ for all $1 \leq i \leq n$.

Now as we have a concept of length $l(I)$ for an interval I, here we have a concept of generalised volume $V(R)$ of an rectangle R which is defined as following. If $R = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$ we define $V(R) = (b_1 - a_1)(b_2 - a_2)\cdots(b_n - a_n)$. Note that this quantity represents the area in $\mathbb{R}^2$ and volume in $\mathbb{R}^3$. In general we will call them volume.

Also our definition of $m(T)$ is also slightly different but the concept is similar. Let $T = \cup_{i\in I} R_i$. We now define (with an informal notation) $m(T) = \Sigma_{i\in I} V(R_i)$. Either $m(T)$ doesn't exist or $m(T) \in \mathbb{R}$ if it exists. So now we are all set to define Lebesgue measure in higher dimension.

Let $A \subseteq \mathbb{R}^n$. The Lebesgue measure of A is defined as

$$m^*(A) = \inf_{T\in\mathcal{O}(A)} m(T)$$

10.3 Basic properties of Lebesgue measure

0. Any open set has non-zero Lebesgue measure.
$\Rightarrow$ Open sets contain open intervals or open rectangles which has positive length/volume.

1. Let $A \subseteq \mathbb{R}^n$. We say A has measure 0 in $\mathbb{R}^n$ if for all $\varepsilon > 0$ there exists an open cover $\Theta \in \mathcal{O}(A)$ such that $m(\Theta) < \varepsilon$. In general A has measure α in $\mathbb{R}^n$ i.e. $m^*(A) = \alpha$ if for all $\varepsilon > 0$ there exists an open cover $\Theta \in \mathcal{O}(A)$ such that $\alpha \leq m(\Theta) < \alpha + \varepsilon$.
$\Rightarrow$ This directly follows from the definition of infimum.

2. $B \subseteq A \subseteq \mathbb{R}^n$. Then $m^*(B) \leq m^*(A)$.
$\Rightarrow$ Note that if Θ is an open cover of A it is also an open cover of B. So $\mathcal{O}(A) \subseteq \mathcal{O}(B)$. Which implies $\inf_{T\in\mathcal{O}(B)} m(T) \leq \inf_{T\in\mathcal{O}(A)} m(T)$.

3. Let $A \subseteq \mathbb{R}^n$ and $A = \cup_{i\in I} A_i$ for some countable index I. $m^*(A_i) = 0$ for all $i \in I$. Then $m^*(A) = 0$.
$\Rightarrow$ We can index the elements of I as $1, 2, 3, \cdots$ because set of natural numbers is countable and has a bijection with I. Start with an arbitrary ε. Now use property 1 for A_1 with $\frac{\varepsilon}{2}$. Again use property 1 for A_2 with $\frac{\varepsilon}{4}$. In a word, use property 1 for A_i with $\frac{\varepsilon}{2^i}$. Now the union of corresponding covers of A_i s will be a cover of A. Finally use the fact that $\Sigma_{i\in\mathbb{N}} \frac{\varepsilon}{2^i} = \varepsilon$ to show $m^*(A) = 0$.

4. For a subset A of $\mathbb{R}^n$ we define Boundary $Bdy(A) = \overline{A} \setminus int(A)$ where $\overline{A}$ is the closure of A and $int(A)$ is the interior of A.
Let Θ be an open rectangle in $\mathbb{R}^n$. Then $m^*(Bdy(\Theta)) = 0$ but $m^*(\Theta) \neq 0$.
$\Rightarrow$ Exercise.

5. A is a finite subset of $\mathbb{R}^n$. Then $m^*(A) = 0$
$\Rightarrow$ Let $A = \{a_1, a_2, \cdots, a_k\}$. So any a_i is of the form $(a_{i1}, a_{i2}, \cdots, a_{in})$ because $a_i \in \mathbb{R}^n$ for all $1 \leq i \leq k$. Start with an arbitrary ε. Consider an open rectangle $R_i =$

$(a_{i1}-\frac{1}{4}(\frac{\varepsilon}{k})^{1/n}, a_{i1}+\frac{1}{4}(\frac{\varepsilon}{k})^{1/n})\times(a_{i2}-\frac{1}{4}(\frac{\varepsilon}{k})^{1/n}, a_{i2}+\frac{1}{4}(\frac{\varepsilon}{k})^{1/n})\times\cdots\times(a_{in}-\frac{1}{4}(\frac{\varepsilon}{k})^{1/n}, a_{in}+\frac{1}{4}(\frac{\varepsilon}{k})^{1/n})$. So $V(R_i) = \frac{\varepsilon}{2^n k}$. Now take an open cover of A by $\Theta = \cup_{i=1}^{k} R_i$. We have $m(\Theta) = \frac{\varepsilon}{2^n} < \varepsilon$. Now use property 1.

6. A is a countable subset of $\mathbb{R}^n$. Then $m^*(A) = 0$
$\Rightarrow$ Let $A = \{a_1, a_2, \cdots\}$. So any a_i is of the form $(a_{i1}, a_{i2}, \cdots, a_{in})$ because $a_i \in \mathbb{R}^n$ for all $1 \leq i \leq k$. Start with an arbitrary ε. Consider an open rectangle $R_i = (a_{i1} - \frac{1}{4}(\frac{\varepsilon}{2^i})^{1/n}, a_{i1}+\frac{1}{4}(\frac{\varepsilon}{2^i})^{1/n})\times(a_{i2}-\frac{1}{4}(\frac{\varepsilon}{2^i})^{1/n}, a_{i2}+\frac{1}{4}(\frac{\varepsilon}{2^i})^{1/n})\times\cdots\times(a_{in}-\frac{1}{4}(\frac{\varepsilon}{2^i})^{1/n}, a_{in}+\frac{1}{4}(\frac{\varepsilon}{2^i})^{1/n})$. So $V(R_i) = \frac{\varepsilon}{2^n 2^i}$. Now take an open cover of A by $\Theta = \cup_{i\in\mathbb{N}} R_i$. We have $m(\Theta) = \frac{\varepsilon}{2^n}(\Sigma_{i\in\mathbb{N}}\frac{1}{2^i}) = \frac{\varepsilon}{2^n} < \varepsilon$. Now use property 1.

7. Prove or disprove. A is an uncountable subset of $\mathbb{R}^n$. Then $m^*(A) \neq 0$
$\Rightarrow$ Cantor set is the counter-example.

Cantor Set

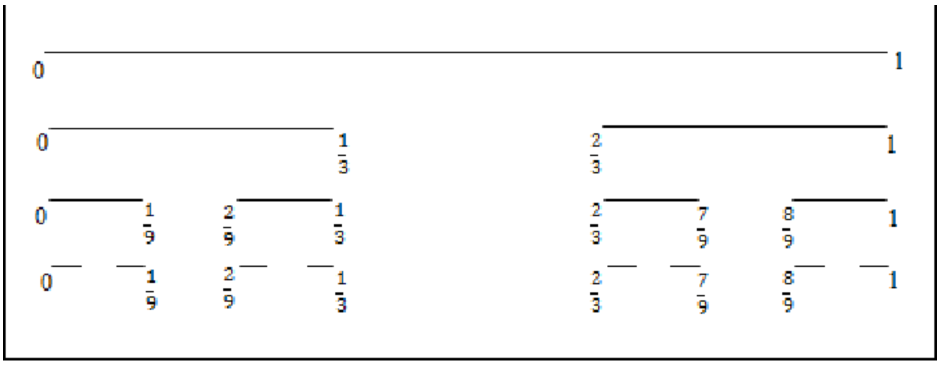

Figure 3: Construction of cantor Set

We start with $[0, 1]$ in $\mathbb{R}$. we divide this segment in three equal parts and in the first step we remove the middle one third of it. So now we are left with two line segments. In the next step we remove the middle one third of both of them. So in n^{th} step we remove the middle one third of each of the line segment left in $(n-1)^{th}$ step. Let the thing left after first step is $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Then $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. In this way after n steps we are left with C_n which is a subset of $[0, 1]$. We define our cantor set to be $C = \cap_{n\in\mathbb{N}} C_n$. So it is just the left part after deleting every middle parts. This is called Cantor's Dust as well. Because the left part is just like dust particles lies in the interval $[0, 1]$.

First we need to show that C is uncountable. We leave this as an exercise to the reader with a small hint. Use the fact that $[0, 1]$ is uncountable and show that there exists a bijection between $[0, 1]$ and C. You can use the binary expansions of the numbers between 0 and 1 and ternary expansions of the numbers in C once you know what they mean. Then try to observe that the ternary expansions of the numbers in C doesn't contain the digit 1, they are just formed by 0 and 2 s. And the binary expansion of a number contains only 0 and 1. So now you can define an obvious bijection between them.

Now we show that $m^*(C) = 0$. Let's calculate first the length of the interval we remove in each step. For step 1 it is $\frac{1}{3}$. Then for step 2 it becomes $\frac{1}{3} \times \frac{2}{3}$. In this way for step n it is $\frac{1}{3} \times (\frac{2}{3})^{n-1}$. So the total length we remove is

$$\Sigma_{n=0}^{\infty} \frac{1}{3} \times (\frac{2}{3})^n = \frac{\frac{1}{3}}{1-\frac{2}{3}} = 1$$

. As the total length of $[0,1]$ is 1. So the length remains after removing these is 0. So $m^*(C) = 0$.
One can assume or try to prove from the definition of series that for $a, r \in \mathbb{R}$ and $0 < r < 1$, we have $a + ar + ar^2 + \cdots = \dfrac{a}{1-r}$.

Now we are going to state the theorem which was the main goal of this section. We find the *Lebesgue criterion* of integrability.

Theorem 59 ***Lebesgue criterion of Integrability :*** *Let Q be a rectangle in $\mathbb{R}^n$. $f : Q \to \mathbb{R}$ be a bounded function. Let D be the set of points of Q where f is not continuous. i.e.*

$$D = \{x \in Q \mid f \ \ is \ \ not \ \ continuous \ \ at \ \ x\}$$

Then $\int_Q$ exists if and only if $m^*(D) = 0$.

As we have seen an important result that set of points of discontinuity of an integrable function has measure zero, let's solve some problems using this criterion.

Exercise 60 *Define the greatest integer function $\lfloor . \rfloor : \mathbb{R} \to \mathbb{Z}$ by $\lfloor x \rfloor =$ the greatest integer less than or equal to x. For example, $\lfloor 3.42 \rfloor = 3$, $\lfloor -5.19 \rfloor = -6$ and $\lfloor 1 \rfloor = 1$. Prove that $\lfloor . \rfloor$ is integrable over $[-\pi, 7.31]$.*

Exercise 61 *Let Q be a rectangle in $\mathbb{R}^n$. $f : Q \to \mathbb{R}$ is integrable over Q. Prove the following*

1. *If f vanishes except on a set of measure 0, then $\int_Q f = 0$. i.e. Let the set $V = \{x \in Q \mid f(x) \neq 0\}$ and $m^*(V) = 0$ then $\int_Q f = 0$.*

2. *If $f(x) \geq 0$ for all $x \in Q$ and $\int_Q f = 0$ then the set $V = \{x \in Q \mid f(x) \neq 0\}$ has measure 0 i.e. f vanishes except on a set of measure 0.*

11 Extended Integrals

11.1 Integration over bounded sets

Till now we have done integrations only over closed and bounded rectangles in $\mathbb{R}^n$. Throughout this section we will figure out how else we can define integrations, specifically over what other sets. Our first goal is to do integrations over any bounded set, which is not necessary closed rectangle.

Let A be a bounded set in $\mathbb{R}$. So $A \subseteq [a,b]$ for some $a, b \in \mathbb{R}$. Similarly let S be a bounded set in $\mathbb{R}^n$. Now $S \subseteq Q$ for some closed rectangle $Q \in \mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be a bounded function. Till now we don't know the meaning of integration of f over S. Here we are going to develop that idea.

Choose a closed rectangle $\Theta \in \mathbb{R}^n$ such that $S \subseteq \Theta$. Define a function $f_S : \Theta \to \mathbb{R}$ by

$$f_S(x) = \begin{cases} f(x) & if \;\; x \in S \\ 0 & if \;\; x \in \Theta \setminus S \end{cases}$$

Note that f_S is a bounded function. Now we can talk about integrability of the function f_S over Θ. Now we are ready to define integration of f over S.

Definition 62 *With the afore mentioned set up f is integrable over S if f_S is integrable over Θ for any choice of Θ. And the value of the integration of f over S is*

$$\int_S f = \int_\Theta f_S$$

Now we see a theorem which is an initiative to define integration over open sets.

Theorem 63 *Let S be a bounded set in $\mathbb{R}^n$. $f : S \to \mathbb{R}$ is a bounded continuous function. Let $A = int(S)$. If f is integrable over S, f is also integrable over A, and $\int_A f = \int_S f$.*

Our next goal is to define a new concept namely rectifiability of a bounded set. Later we will see that some special kind of sets (called compact rectifiable sets) take an important role in the theory of integration.

11.2 Rectifiability

Definition 64 (Rectifiable set) *Let S be a bounded set in $\mathbb{R}^n$. If the constant function $f : S \to \mathbb{R}$ by $f(x) = 1 \; \forall \; x \in S$ is integrable over S, we say that S is rectifiable.*

Interestingly, for a rectifiable set S, the measure/volume of S or value of S is defined by,

$$\mathcal{V}(S) = \int_S 1$$

Similarly we can define volume of any bounded set A in $\mathbb{R}^n$ by the same formula $\mathcal{V}(A) = \int_A 1$ if it exists.

Recall that for a subset A of $\mathbb{R}^n$ we define **Boundary** $Bdy(A) = \overline{A} \setminus int(A)$ where $\overline{A}$ is the closure of A and $int(A)$ is the interior of A. The next theorem establishes a relation between measure and rectifiability.

Theorem 65 *A bounded set S of $\mathbb{R}^n$ is rectifiable $\Leftrightarrow$ $Bdy(S)$ has measure 0.*

It is a highly non-trivial fact that there exists bounded non-rectifiable set i.e. there exists bounded set such that the constant function 1 is not integrable over that. We give a counter example by producing a bounded set whose boundary has measure greater than 0. Set of rational numbers $\mathbb{Q}$ is countable. So denote the set of all rational numbers in $(0, 1)$ by a sequence $q_1, q_2, \cdots$. Fix a real number a such that $0 < a < 1$. Now for each $i \in \mathbb{N}$ we can choose an interval $(a_i, b_i) \subset (0, 1)$ such that $b_i - a_i < \frac{a}{2^i}$ and $q_i \in (a_i, b_i)$. Define $S = \cup_{i \in \mathbb{N}} (a_i, b_i)$. Note that $\overline{S} = [0, 1]$. So $m^*(\overline{S}) = 1$. Now $int(S)$ is an open set.

So $0 < m^*(int(S)) < \Sigma_{i \in \mathbb{N}} \frac{a}{2^i} = a$. Hence, $m^*(Bdy(S)) > 1 - a > 0$. So S is a bounded set which is not rectifiable.

Now we will see some properties of rectifiability. Note that all the sets we will talk about in the following properties are bounded subsets of $\mathbb{R}^n$.

1. If S is rectifiable, $\mathcal{V}(S) \geq 0$.

2. If S_1, S_2 are rectifiable and $S_1 \subseteq S_2$, then $\mathcal{V}(S_1) \leq \mathcal{V}(S_2)$.

3. If S_1 and S_2 are rectifiable, so are $S_1 \cup S_2$ and $S_1 \cap S_2$. And

$$\mathcal{V}(S_1 \cup S_2) = \mathcal{V}(S_1) + \mathcal{V}(S_2) - \mathcal{V}(S_1 \cap S_2)$$

4. Suppose S is rectifiable then $\mathcal{V}(S) = 0$ if and only if S has Lebesgue measure 0 i.e. $m^*(S) = 0$.

5. If S is rectifiable so is $A = int(S)$ and $\mathcal{V}(A) = \mathcal{V}(S)$.

6. If S is rectifiable and $f : S \to \mathbb{R}$ is bounded continuous function, then f is integrable over S.

11.3 Integration over simple regions

Now we proceed one step further and define integrations on a special kind of sets called simple regions.

Definition 66 (Simple region) *Let C be a compact rectifiable set in $\mathbb{R}^{n-1}$. Let ϕ, ψ be continuous functions from C to $\mathbb{R}$ such that $\phi(x) \leq \psi(x)$ for all $x \in C$. The subset S of $\mathbb{R}^n$ defined by*

$$S = \{(x,t) \in \mathbb{R}^n \mid x \in C \wedge \phi(x) \leq t \leq \psi(x)\}$$

is called a simple region in $\mathbb{R}^n$.

Lemma 67 *If S is a simple region in $\mathbb{R}^n$, then S is compact and rectifiable.*

$Lemma$ 67 ensures us that we can define integrations over simple regions because they are bounded being a compact set. Also any bounded continuous function is integrable over them ($property$ 6) because they are rectifiable. Next theorem shows how to calculate integrations over simple regions. We state the theorems and see some applications of it.

Theorem 68 (Fubini's theorem for simple regions) *Let C be a compact rectifiable set in $\mathbb{R}^{n-1}$. Let ϕ, ψ be continuous functions from C to $\mathbb{R}$ such that $\phi(x) \leq \psi(x)$ for all $x \in C$. $S = \{(x,t) \in \mathbb{R}^n \mid x \in C \wedge \phi(x) \leq t \leq \psi(x)\}$ is simple region in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be a bounded continuous function. Then*

$$\int_S f = \int_C \int_{t=\phi(x)}^{t=\psi(x)} f(x,t)\,dt\,dx$$

Exercise 69 *Find the area enclosed between the parabola $y = x^2$ and the straight line $x + y = 2$ in Cartesian plane.*

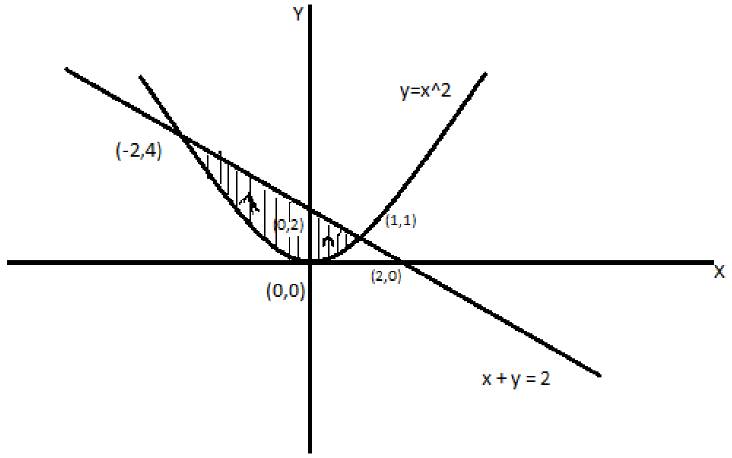

Figure 4: Exercise-69

Hint : Calculate $\int_{x=-2}^{x=1}\int_{y=x^2}^{y=2-x} 1dydx$ and before that try to see how to get this using Fubini's theorem. **(See Figure 4)**

Till now we have seen integrations on closed and bounded rectangles, bounded sets, compact and rectifiable sets, simple regions. Our current goal is to define integrations on open sets.

11.4 Integration over open sets

Let A be an open set in $\mathbb{R}^n$. Let $f : A \to \mathbb{R}$ be a continuous function. If $f(x) \geq 0$ for all $x \in A$, we can define **Extended integral** of f over A, denoted by $\int_A f$.

Definition 70 (Extended integrals) *With the afore mentioned set up $\int_A f$ is the supremum of the set $S = \{\int_D f \;\mid\; D\ is\ a\ compact\ rectifiable\ subset\ of\ A\}$ if it exists. If $\sup S$ doesn't exist extended integral of f over A doesn't exist.*

We were working with non-negative function f here. Now we will do it more generally. Let f is a continuous function from A to $\mathbb{R}$. Define $f_+(x) = max\{f(x), 0\}$ and $f_-(x) = max\{-f(x), 0\}$. Note that both f_+ and f_- are non negative functions from A to $\mathbb{R}$. We have $f = f_+ - f_-$ and $|f| = f_+ + f_-$. If both $\int_A f_+$ and $\int_A f_-$ exists, we say extended integrals $\int_A f$ and $\int_A |f|$ exist where

$$\int_A f = \int_A f_+ - \int_A f_- \quad and \quad \int_A | f | = \int_A f_+ + \int_A f_-$$

Before $definition$ 70 we were always working with bounded functions. Here we dropped that condition. But then for the well-defineness, we need to show $\int_D f$ exists for any compact rectifiable set D and continuous function f from D to $\mathbb{R}$. That will be ensured by the following theorem. Also note that D is bounded because it is compact.

Theorem 71 *Let A be a bounded subset of a metric space X. $f : X \to \mathbb{R}$ is a continuous function. Then $f(A)$ is bounded in $\mathbb{R}$.*

So wherever, we have defined bounded and continuous real valued function on a bounded set it is unnecessary because only the continuity on a bounded set implies the boundedness of the function.

In the above definition, we took the supremum over all the compact rectifiable subsets of an open set. The next Lemma shows that in an open set there exits a compact rectifiable subset. In fact, there are plenty of them.

Lemma 72 *Let A be an open set in $\mathbb{R}^n$. Then there exists sequence of compact rectifiable subsets $C_1, C_2, \cdots$ of A such that $\cup_{i\in\mathbb{N}} C_i = A$ and $C_n \subseteq int(C_{n+1})$ for all $n \in \mathbb{N}$.*

Theorem 73 *Let A be an open set in $\mathbb{R}^n$. Let $C_1, C_2, \cdots$ be a sequence of compact rectifiable subsets of A such that $\cup_{i\in\mathbb{N}} C_i = A$ and $C_n \subseteq int(C_{n+1})$ for all $n \in \mathbb{N}$. Then the extended integral of f over A exists if and only if the sequence $\int_{C_n} |f|$ is bounded. In this case the value of the extended integral is*

$$\int_A f = \lim_{n\to\infty} \int_{C_n} f$$

As of now, we have a notion of ordinary integrals over bounded sets and extended integrals over any open set. The next theorem shows the relation between them in a bounded open set.

Theorem 74 *Let A be a bounded open subset of $\mathbb{R}^n$. Let $f : A \to \mathbb{R}$ be a bounded continuous function. Then the extended integral $\int_A f$ exists.*
If the ordinary integral also exists then they are equal.

Now we are ready to state the final theorem of this section. This helps us to define integrations over a larger class of sets in a handy way.

Theorem 75 *Let A be open in $\mathbb{R}^n$. $f : A \to \mathbb{R}$ is continuous. Let $U_1 \subseteq U_2 \subseteq U_3 \subseteq \cdots$ be a sequence of open sets such that $\cup_{i\in\mathbb{N}} U_i = A$ (This sequence always exists because A itself is an open set). Then the extended integral $\int_A f$ exists if and only if the sequence $\int_{U_n} \mid f \mid$ exists and it is bounded. In that case the value of the extended integral is*

$$\int_A f = \lim_{n\to\infty} \int_{U_n} f$$

Now we see very important applications of these techniques of doing integration over open sets.

Exercise 76 *Calculate $\int_1^\infty \frac{1}{x} dx$.*

Hint : We are asked to calculate the extended integral of the function $f(x) = \frac{1}{x}$ in the open set $(1, \infty)$. For all $n \in \mathbb{N}$ define open sets $U_n = (1, n)$. Now $(1, \infty) = \cup_{n\in\mathbb{N}}(1, n)$. Also observe that $\int_{U_n} \frac{1}{x} dx = \int_1^n \frac{1}{x} dx = 1 - \frac{1}{n^2}$. Now $\lim_{n\to\infty} \int_{U_n} \frac{1}{x} dx = \lim_{n\to\infty}(1 - \frac{1}{n^2}) = 1$. So $\int_1^\infty \frac{1}{x} dx = 1$.

Exercise 77 *Calculate* $\int_1^\infty \int_1^\infty \frac{1}{x^2y^2}dxdy$

Hint : Open set is $(1,\infty) \times (1,\infty) \subseteq \mathbb{R}^2$. Define $U_n = (1,n) \times (1,n)$. Observe that $\int_{U_n} \frac{1}{x^2y^2} = 1 - \frac{2}{n} + \frac{1}{n^2}$. Find the limit as n approaches to infinity.

In the theory of integration there is something called **Improper integrals**. They are used to calculate integrals on $\mathbb{R}$ over some unbounded open sets. The limit of the integration becomes ∞ or $-\infty$. Note that, the theory of improper integrals is nothing but the theory of extended integrals we already discussed. They are just restricted to $\mathbb{R}$ in place of $\mathbb{R}^n$ i.e. $n = 1$.

12 Integrals as anti-derivatives

12.1 Fundamental theorem of calculus

Till now we were busy with developing the theory. Now it's time to learn some basic techniques to compute integrals. Before starting this section we will state an absurd statement, which has no such concrete and well-defined meaning in the literature, but for an intuitionist the statement makes sense. In the chapter we will observe to what extent the statement is true.

$$f' = g \implies \int g = f$$

Now its time to formalise the concepts.

Theorem 78 *Let the function* $f : [a,b] \to \mathbb{R}$ *be continuous. Define a function* $F : [a,b] \to \mathbb{R}$ *by* $F(x) = \int_a^x f(t)dt$. *(Note that* t *is just the name of the variable it's just that* $F(x) = \int_{[a,x]} f$*). Then* F *is differentiable on* $[a,b]$ *and* $F'(t) = f(t)$ *for all* $t \in [a,b]$.

The proof of this is straightforward from the definitions. Our next theorem is the most basic one.

Theorem 79 (Fundamental theorem of integral calculus(FTIC)) *Suppose that the function* $F : [a,b] \to \mathbb{R}$ *is differentiable and* F' *is continuous. i.e.* $F \in \mathcal{C}^1[a,b]$. *Then* F' *is integrable over* $[a,b]$ *and*

$$\int_a^b F'(x)dx = F(b) - F(a)$$

This can be proved using the previous theorem along with Mean value theorem.

In one-variable calculus one learns various techniques to find anti-derivatives; i.e., given continuous f, one finds F such that $F' = f$. Once this is done, evaluating $\int_a^b$ is mere plug-in to the **FTIC**. But since not all continuous functions have anti-derivatives that are readily found, or even possible to write in an elementary form (for example, try $f(x) = e^{-x^2}$ or $f(x) = sin(x^2)$), the **FTIC** has its limitations.

Another tool for evaluating one-dimensional integrals is the Change of Variable Theorem. The idea is to transform one integral to another that may be better suited to the **FTIC**.

12.2 Substitutions

Theorem 80 (Change of Variable Theorem) *Let $\phi : [a, b] \to \mathbb{R}$ be differentiable with continuous derivative i.e. of class $\mathcal{C}^1$ and let $f : \phi[a, b] \to \mathbb{R}$ be continuous. Then*

$$\int_a^b (f \circ \phi)\phi' = \int_{\phi(a)}^{\phi(b)} f$$

this theorem is also known as **Forward substitution formula**. This is a very useful technique in terms of computation. For example let's say we have to calculate $\int_1^e \frac{(\log\ x)^2}{x} dx$. We define $\phi : \mathbb{R}^+ \to \mathbb{R}$ by $\phi(x) = log\ x$ and $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Then we use this formula and calculate $\int_0^1 x^2 dx$ which is much easy to calculate than the actual function.

Corollary 81 (Inverse substitution formula) *With the assumptions of theorem 80 and an extra asumption that ϕ is invertible we have*

$$\int_a^b f \circ \phi = \int_{\phi(a)}^{\phi(b)} f\ (\phi^{-1})'$$

Now we will try to differentiate inside the integral sign.

Theorem 82 (Leibniz rule) *Let $\phi, \psi : A \to \mathbb{R}$ be differentiable functions where A is a bounded subset of $\mathbb{R}$ and $\phi(x) \leq \psi(x)$ for all $x \in A$. Define $S = \{(x, t) \mid x \in A \wedge \phi(x) \leq t \leq \psi(x)\}$. $f : S \to \mathbb{R}$ be a differentiable function. We denote $g'(x)$ by $\frac{d}{dx}g$. Then*

$$\frac{d}{dx}\int_{\phi(x)}^{\psi(x)} f(x, t)dt = f(x, \psi(x))\psi'(x) - f(x, \phi(x))\phi'(x) + \int_{\phi(x)}^{\psi)(x)} \frac{\partial}{\partial x} f(x, t)dt$$

Special cases :
If f is a function of t only i.e. $S = \{t \mid \phi(x) \leq t \leq \psi(x)\}$ then

$$\frac{d}{dx}\int_{\phi(x)}^{\psi(x)} f(t)dt = f(\psi(x))\psi'(x) - f(\phi(x))\phi'(x)$$

If ϕ and ψ are constant functions a and b then

$$\frac{d}{dx}\int_a^b f(x, t)dt = \int_a^b \frac{\partial}{\partial x} f(x, t)dt$$

13 Fubini's theorem and change of variable

13.1 Fubini's theorem

With existence theorems for the integral now in hand, this section present tools to compute integrals.
An $n - fold$ iterated integral is n one-dimensional integrals nested inside each other, such as

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_n}^{b_n} f(x_1, x_2, x_3, \cdots, x_n)dx_n dx_{n-1}\cdots dx_2 dx_1$$

So we can think of this as following. In the first step let $x_1, x_2, \cdots, x_{n-1}$ act as constants. First we do the integration by taking only x_n as the variable and over $[a_n, b_n]$. Now the result is a function of the variables $x_1, x_2, \cdots, x_{n-1}$. similarly we think of $x_1, x_2, \cdots, x_{n-2}$ as constants and integrate the function of x_{n-1} over $[a_{n-1}, b_{n-1}]$. And we follow the same for n steps to get the final value of the integral.

Fubini's Theorem says that under suitable conditions, the n dimensional integral is equal to the $n-fold$ iterated integral. The theorem thus provides an essential calculational tool for multi variable integration.

Theorem 83 (Fubini's theorem) *Let $\Theta = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ be a rectangle in $\mathbb{R}^n$. $f : \Theta \to \mathbb{R}$ be a continuous function. Then*

$$\int_\Theta f = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x_1, x_2, x_3, \cdots, x_n) dx_n dx_{n-1} \cdots dx_2 dx_1$$

Proof of this theorem is simple using the concept of definition of integration repetitively. Next we just solve some problems using this.

Exercise 84 *Calculate $\int_{[0,1]\times[0,2]} xy^2$ and $\int_{[0,1]^3} x^2 + y^2 + z^2$*

The best part of Fubini's theorem is that the value of the integral doesn't depend on the order inside the $n-fold$ integral. For example let $R = [a, b] \times [c, d]$. Then the value of $\int_R f$ can be both of the following and they must be equal because $\int_R f$ is unique.

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

Theorem 85 (Differentiation under integral sign) *Consider a function $f : [a, b] \times [c, d] \to \mathbb{R}$. Let f is of class C^1. Then the following exists and the expression holds true.*

$$\frac{d}{dx} \int_c^d f(x, y) dy = \int_c^d \frac{\partial}{\partial x} f(x, y) dy$$

Proof : Define $g : [a, b] \to \mathbb{R}$ by $g(x) = \int_c^d f(x, y) dy$. Take $x \in [a, b]$ arbitrarily. Then we have $g(x) =$

$= \int_c^d f(x, y) dy$

$= \int_c^d \left(\int_a^x \dfrac{\partial}{\partial x} f(x, y) dx + f(a, y) \right) dy \qquad (use\ FTIC\ for\ F(x) = f(x, y))$

$= \int_c^d \int_a^x \dfrac{\partial}{\partial x} f(x, y) dx dy + C \qquad where\ C = \int_c^d f(a, y) dy$

$= \int_a^x \int_c^d \dfrac{\partial}{\partial x} f(x, y) dy dx + C \qquad use\ Fubini's\ theorem$

Now we apply $theorem$ 78 on the function $g(x) - C$ and get $g'(x) = \int_c^d \frac{\partial}{\partial x} f(x, y) dy$. Hence proved. ■

Exercise 86 *Calculate $\int_0^1 \int_{x=y}^{x=1} \frac{sin(x)}{x} dx dy$. (**Hint :** The way the integral is given it is difficult to calculate. We will interchange limits. Try to see the simple region, sketch that. Then do a smart use of Fubini's theorem and see that the integral is same as $\int_0^1 \int_{y=0}^{y=x} \frac{sin(x)}{x} dy dx$*

13.2 Change of variable

Now we will discuss the change of variable formula and see it's applications. Before that recall that if $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable, then it's derivative is a $m \times n$ matrix. Also recall what Jacobian is. (Section 2 and 3). The Change of Variable Theorem says in some generality how to transform an integral from one coordinate system to another.

Theorem 87 (Change of variables) *Let $K \subseteq \mathbb{R}^n$ is a compact, connected set and it's boundary has measure 0. Let A is an open set of $\mathbb{R}^n$ and $K \subseteq A$. $\Phi : A \to \mathbb{R}^n$ is of class $\mathcal{C}^1$ such that Φ is injective on $int(K)$ and $det[\mathcal{J}(\Phi)(x)] \neq 0$ for all $x \in int(K)$. Let $f : \Phi(K) \to \mathbb{R}$ be a continuous function. Then*

$$\int_{\Phi(K)} f = \int_K (f \circ \Phi). \mid det[\mathcal{J}(\Phi)] \mid$$

This is a generalised statement and the proof is technical. But our goal from this theorem is to look at some useful transformation of co-ordinates which we often use in practical purposes. In next three subsections we discuss some usual change of co-ordinates in $\mathbb{R}^2$ and $\mathbb{R}^3$.

13.3 Change of co-ordinates : polar, cylindrical, spherical

Polar co-ordinate

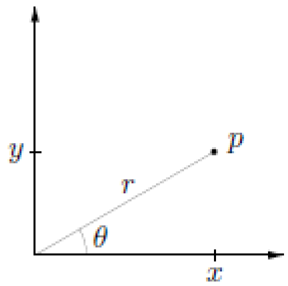

Figure 5: polar co-ordinate

Here we change our usual two dimensional (x, y) system to (r, θ) system. Let the co-ordinate of a point p in usual Cartesian system is (x, y). Note that we can uniquely determine this position by (r, θ) where r is the distance of this point from origin. i.e. $r = \sqrt{x^2 + y^2}$. And then θ determines the angular position of this point in the circle of radius r around origin. θ is the angle between the line joining origin and p and x $axis$.

If θ ranges over any fixed semi-open interval of length 2π on $\mathbb{R}$, we have $\tan\theta = \frac{y}{x}$. We also have $x = r\cos\theta$ and $y = r\sin\theta$. Now we see how we transform variables inside the integrations. We will define the general mapping Φ for polar co-ordinates and then depending on the question we will fix/restrict the domain and use $theorem$ 87.
Define $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ by $\Phi(r, \theta) = (r\cos\theta, r\sin\theta)$. In general, $r \geq 0$ and $0 \leq \theta \leq 2\pi$. It is

easy to check that Φ is of class $\mathcal{C}^1$. Now we see what $\mathcal{J}(\Phi)$ is.

$$\mathcal{J}(\Phi) = \begin{bmatrix} \frac{\partial}{\partial r} r\cos\theta & \frac{\partial}{\partial \theta} r\cos\theta \\ \frac{\partial}{\partial r} r\sin\theta & \frac{\partial}{\partial \theta} r\sin\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

Hence $det[\mathcal{J}(\Phi)] = r(\cos^2\theta + \sin^2\theta) = r$.

Let $A = \{v \in \mathbb{R}^2 \mid a < ||v|| < b\}$. For example we integrate the function $f(x, y) = x^2 + y^2$ over A. So we can observe that $A = \Phi([a, b] \times [0, 2\pi]) = \Phi(K)$. Φ is injective on $(a, b) \times (0, 2\pi) = int(K)$. and $0 \le a < r < b$ for all $r \in int(K)$. So $det[\mathcal{J}(\Phi)] \neq 0$ in $int(K)$. Now we are ready to calculate the integral.

Note that $f(\Phi(r, \theta)) = f(r\cos\theta, r\sin\theta) = r^2$. Now $\int_A f = \int_K (f \circ \Phi) r = \int_a^b \int_0^{2\pi} r^3 d\theta dr$. So we have $\int_A f = \frac{\pi}{2}(b^4 - a^4)$.

Exercise 88 *Calculate* $\int_0^1 \int_0^1 e^{x^2+y^2} dxdy$ *and* $\int_0^\infty e^{-x^2} dx$

Cylindrical co-ordinate

Cylindrical co-ordinate is just a replica of polar co-ordinate in 3 dimension. We change our usual (x, y, z) system to (r, θ, z) system. Where (r, θ) comes from the polar co-ordinate transformation of (x, y) and z remains as it is.

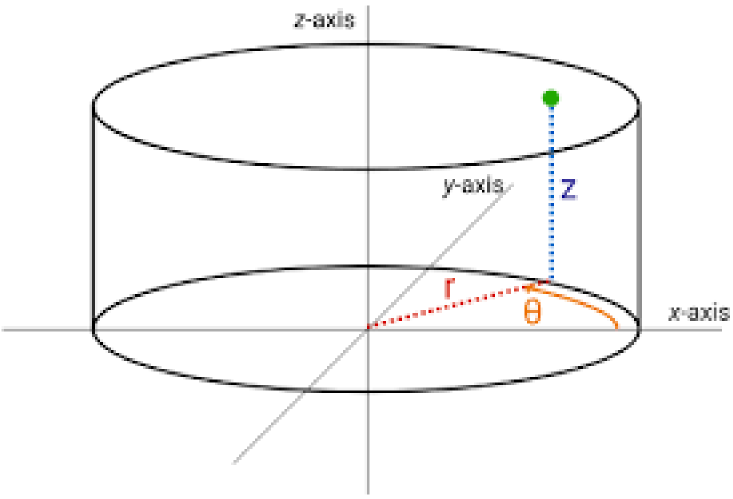

Figure 6: cylindrical co-ordinate

Now we define the general co-ordinate changing mapping. Define $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ by $\Phi(r, \theta, z) = (r\cos\theta, r\sin\theta, z)$. In general, $r \ge 0$, $0 \le \theta \le 2\pi$ and $z \in \mathbb{R}$. It is easy to check that Φ is of class $\mathcal{C}^1$. Now we see what $\mathcal{J}(\Phi)$ is.

$$\mathcal{J}(\Phi) = \begin{bmatrix} \frac{\partial}{\partial r} r\cos\theta & \frac{\partial}{\partial \theta} r\cos\theta & \frac{\partial}{\partial z} r\cos\theta \\ \frac{\partial}{\partial r} r\sin\theta & \frac{\partial}{\partial \theta} r\sin\theta & \frac{\partial}{\partial z} r\sin\theta \\ \frac{\partial}{\partial r} z & \frac{\partial}{\partial \theta} z & \frac{\partial}{\partial z} z \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hence $det[\mathcal{J}(\Phi)] = 1(r(\cos^2\theta + \sin^2\theta)) - 0 + 0 = r$.

Let $C = \{(x, y, z) \mid x^2 + y^2 \leq 1,\ 0 \leq z \leq 2\}$. For example we integrate the function $f(x, y, z) = y^2 z$ over C. So we can observe that $C = \Phi([0, 1] \times [0, 2\pi] \times [0, 2]) = \Phi(K)$. Φ is injective on $(0, 1) \times (0, 2\pi) \times (0, 2) = int(K)$. and $0 < r < 1$ for all $r \in int(K)$. So $det[\mathcal{J}(\Phi)] \neq 0$ in $int(K)$. Now we are ready to calculate the integral.

Note that $f(\Phi(r, \theta, z)) = f(r\cos\theta, r\sin\theta, z) = r^2\sin^2\ \theta\ z$. Now $\int_C f = \int_K (f \circ \Phi) r = \int_0^1 \int_0^{2\pi} \int_0^2 r^2 \sin^2\theta\ z\ r\ dz d\theta dr$. So we have $\int_C f = \frac{\pi}{2}$.

Exercise 89 *Let Ω be the region bounded above by the sphere $x^2+y^2+z^2 = 6$ and bounded below by the paraboloid $x^2+y^2 = z$. Calculate $\int_\Omega z$. (**Hint :** Use Fubini's theorem for simple region with change of co-ordinates. Final expression should be*

$$\int_\Omega z = \int_0^{2\pi} \int_0^{\sqrt{2}} \int_{r^2}^{\sqrt{6-r^2}} zr\ dz dr d\theta$$

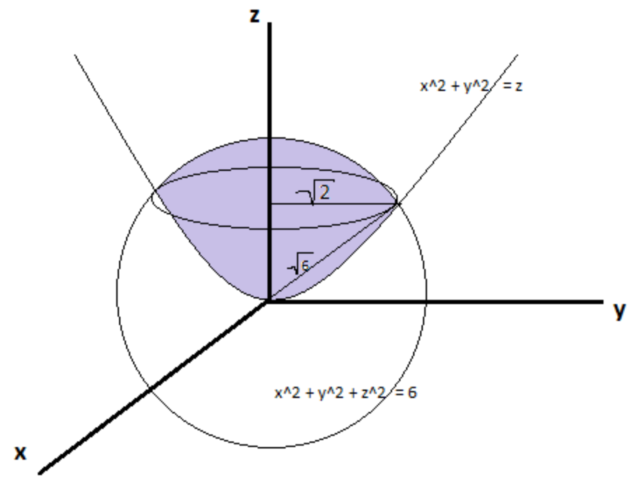

Figure 7: (Exercise-89 ; Shaded portion is Ω)

Spherical co-ordinate

In 3 dimension we change usual (x, y, z) system to (r, θ, ϕ) system through the help of spherical co-ordinate system. Let the co-ordinate of a point p in usual Cartesian system is (x, y, z). Note that we can uniquely determine this position by (r, θ, ϕ) where r is the distance of this point from origin. i.e. $r = \sqrt{x^2 + y^2 + z^2}$. And then ϕ determines the angular position of this point in the circle of radius r around origin on xy plane. ϕ is the angle between the line joining origin and projection of p on xy plane and $x\ axis$. Also θ is the angular position with respect to the z axis i.e. θ is the angle between positive z axis and line joining the origin and p.

If ϕ ranges over any fixed semi-open interval of length 2π on $\mathbb{R}$, and If θ ranges over any fixed semi-open interval of length π on $\mathbb{R}$, we have $z = r\cos\theta$, $x = r\cos\phi\sin\theta$ and $y = r\sin\phi\sin\theta$. Now we see how we transform variables inside the integrations. We

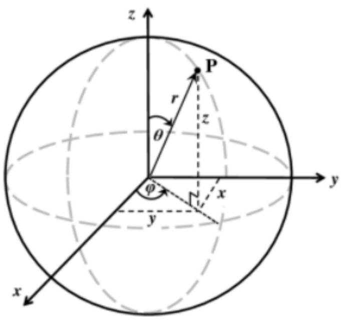

Figure 8: Spherical co-ordinate

will define the general mapping Φ for spherical co-ordinates and then depending on the question we will fix/restrict the domain and use *theorem* 87.

Define $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ by $\Phi(r,\theta,\phi) = (r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\cos\theta)$. In general, $r \geq 0$, $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$ It is easy to check that Φ is of class $\mathcal{C}^1$. Now we see what $\mathcal{J}(\Phi)$ is.

$$\mathcal{J}(\Phi) = \begin{bmatrix} \frac{\partial}{\partial r} r\cos\phi\sin\theta & \frac{\partial}{\partial \theta} r\cos\phi\sin\theta & \frac{\partial}{\partial \phi} r\cos\phi\sin\theta \\ \frac{\partial}{\partial r} r\sin\phi\sin\theta & \frac{\partial}{\partial \theta} r\sin\phi\sin\theta & \frac{\partial}{\partial \phi} r\sin\phi\sin\theta \\ \frac{\partial}{\partial r} r\cos\theta & \frac{\partial}{\partial \theta} r\cos\theta & \frac{\partial}{\partial \phi} r\cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos\phi\sin\theta & r\cos\phi\cos\theta & -r\sin\phi\sin\theta \\ \sin\phi\sin\theta & r\sin\phi\cos\theta & r\cos\phi\sin\theta \\ \cos\theta & -r\sin\theta & 0 \end{bmatrix}$$

Hence $det[\mathcal{J}(\Phi)] = r^2(\sin^3\theta\cos^2\phi + \sin\theta\cos^2\theta\cos^2\phi + \sin^3\theta\sin^2\phi + \sin\theta\cos^2\theta\sin^2\phi) = r^2(\sin^3\theta + \sin\theta\cos^2\theta) = r^2\sin\theta$.

Let $S = \{(x,y,z) \mid x^2+y^2+z^2 = \rho^2\}$. i.e. S is the solid sphere of radius ρ. For example we integrate the function $f(x,y,z) = 1$ over S. i.e. we calculate the volume of S. So we can observe that $S = \Phi([0,\rho]\times[0,\pi]\times[0,2\pi]) = \Phi(K)$. Φ is injective on $(0,\rho)\times(0,\pi)\times(0,2\pi) = int(K)$. and $0 < r^2\sin\theta < \rho^2$ for all $r \in int(K)$. So $det[\mathcal{J}(\Phi)] \neq 0$ in $int(K)$. Now we are ready to calculate the integral.

Note that $f(\Phi(r,\theta,\phi)) = f(r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\cos\theta) = 1$. Now $\int_S f = \int_K (f\circ\Phi)r^2\sin\theta = \int_0^\rho\int_0^\pi\int_0^{2\pi} r^2\sin\theta d\phi d\theta dr$. So we have $\int_S f = \frac{4}{3}\pi\rho^3$.

Exercise 90 *Let Λ be the region bounded by $x^2+y^2+z^2 = 2$ and $z = \sqrt{x^2+y^2}$ in $\mathbb{R}^3$. Calculate $\int_\Lambda z$. (**Hint :** See figure 9 and final expression after change of co-ordinate should be $\int_0^{2\pi}\int_0^{\frac{\pi}{4}}\int_0^{\sqrt{2}} r^3\cos\theta\sin\theta\ dr\ d\theta\ d\phi$*

Exercise 91 (Pappu's theorem) *Let K be a compact set in the (x,z) plane lying to the right of the z axis and with boundary of volume zero. Let S be the solid obtained by rotating K about the z axis in $\mathbb{R}^3$. Also $\overline{x} = \dfrac{\int_K x}{area(K)}$ Then*

$$vol(S) = 2\pi\overline{x}\ area(K)$$

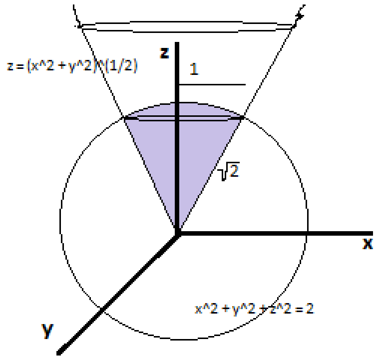

Figure 9: (Ex - 90 ; Shaded portion is Λ)

Hint : *Use change of variable for cylindrical co-ordinate*

Part III
Vector Calculus

14 Scalar and vector field

From this section to the last one we cover a whole new branch namely vector calculus. We define integrations over curves and surfaces. This has a huge application to different fields of physics.

We study the calculus of vector fields. (These are functions that assign vectors to points in space.) Speaking in a physicist's language we define line integrals (which can be used to find the work done by a force field in moving an object along a curve). Then we define surface integrals (which can be used to find the rate of fluid flow across a surface). The connections between these new types of integrals and the single, double, and triple integrals that we have already met are given by the higher-dimensional versions of the Fundamental Theorem of Calculus, Green's Theorem, Stokes' Theorem, and the Divergence Theorem.

Recall that if f is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ we can write $f(x) = (f_1(x), f_2(x), \cdots, f_{m-1}(x), f_m(x))$ for all $x \in \mathbb{R}^n$. Here f_i s can be considered as functions from $\mathbb{R}^n$ to $\mathbb{R}$. These f_i s are just composition of f and projection mappings. (Go back to section 3.3 for detailed discussion). If $e_1, e_2, \cdots, e_m$ is the standard orthonormal basis of $\mathbb{R}^m$, we write $f(x) = f_1(x)e_1 + f_2(x)e_2 + \cdots + f_m(x)e_m$. Now we formally define vector and scalar field.

Definition 92 (Scalar field) $A \subseteq \mathbb{R}^n$. *A scalar field on A is a function from a subset A of $\mathbb{R}^n$ to $\mathbb{R}$.*

Definition 93 (Vector field) $A \subseteq \mathbb{R}^n (n > 1)$. *A vector field on A is a function from a subset A of $\mathbb{R}^n$ to $\mathbb{R}^n$. So if f is a vector field on $\mathbb{R}^n$ we have $f = \Sigma_{i=1}^n f_i e_i$ where $\{e_i\}_1^n$ is the standard orthonormal basis of $\mathbb{R}^n$ and f_i is scalar field for all $1 \leq i \leq n$.*

From now we will use a notation for our standard orthonormal basis of $\mathbb{R}^2$ and $\mathbb{R}^3$. In $\mathbb{R}^2$ we have $\hat{\mathbf{i}} = (1,0), \hat{\mathbf{j}} = (0,1)$. Similarly in $\mathbb{R}^3$ we have $\hat{\mathbf{i}} = (1,0,0), \hat{\mathbf{j}} = (0,1,0), \hat{\mathbf{k}} = (0,0,1)$.

Let F be a vector field on $\mathbb{R}^2$. The best way to picture a vector field is to draw the arrow representing the vector $F(x,y)$ starting at the point (x,y). Of course, it's impossible to do this for all points , but we can gain a reasonable impression of by doing it for a few representative points in as in Figure 10. Since $F(x,y)$ is a two-dimensional vector, we can write it in terms of its component functions and as follows. $F(x,y) = P(x,y)\hat{\mathbf{i}} + Q(x,y)\hat{\mathbf{j}}$ where P and Q are scalar fields. The next figure (Figure-10) shows drawings of three vector fields, $F_1(x,y) = (-y,x), F_2(x,y) = (y, \sin x)$ and $F_3(x,y) = \log(1+y^2)\hat{\mathbf{i}} + \log(1+x^2)\hat{\mathbf{j}}$ respectively.

Now we give some practical examples of vector fields. Newton's Law of Gravitation states that the magnitude of the gravitational force between two objects with masses m and M is $|F| = \frac{mMG}{r^2}$ where r is the distance between the objects and G is the gravitational constant. (This is an example of an inverse square law.) Let's assume that

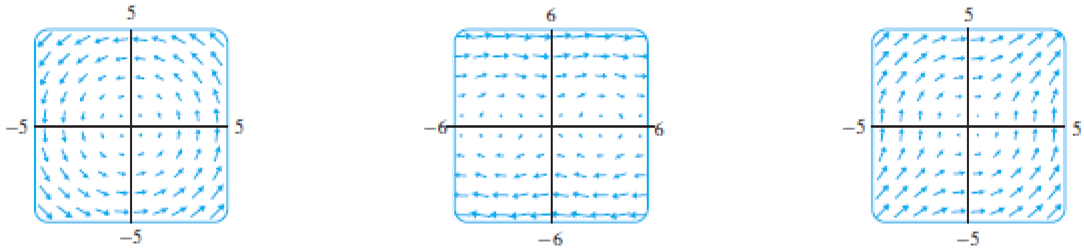

Figure 10: Vector fields on $\mathbb{R}^2$

the object with mass m is located at the origin in $\mathbb{R}^3$. And the position of the object with mass M is (x, y, z). We also have $r = \sqrt{x^2 + y^2 + z^2}$ i.e. $r = \| (x, y, z) \|$. Then the gravitational force exerted on this second object acts toward the origin, and the unit vector in this direction is $-\frac{1}{r}(x, y, z)$. Therefore the gravitational force acting on the object at (x, y, z) is a vector field $\mathfrak{G} : \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\mathfrak{G}(x, y, z) = -\frac{mMG}{\| (x, y, z) \|^3}(x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}})$$

Another example is electric field. Suppose an electric charge Q is located at the origin. According to Coulomb's Law, the electric force $F(v)$ exerted by this charge on a charge q located at a point (x, y, z) with position vector v is $F(v) = \frac{\varepsilon q Q}{\|v\|^3}v$, where ε is a constant (that depends on the units used). For like charges, we have $qQ > 0$ and the force is repulsive; for unlike charges, we have $qQ < 0$ and the force is attractive. Notice the similarity between Formulas for $\mathfrak{G}$ and F. Both vector fields are examples of force fields. Instead of considering the electric force , physicists often consider the force per unit charge i.e.

$$\mathfrak{E}(x, y, z) = \frac{\varepsilon Q}{\| v \|^3}(x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}})$$

. Then is a vector field $\mathfrak{E} : \mathbb{R}^3 \to \mathbb{R}^3$, called the electric field of Q. Now we define some new concepts.

Definition 94 (Gradient field) *Let f be a scalar field on $\mathbb{R}^n$ such that all its first order partial derivatives exist. Then gradient field of f is a vector field on $\mathbb{R}^n$ defined as $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ by* $\nabla f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) & \frac{\partial}{\partial x_2} f(x) & \cdots & \cdots & \frac{\partial}{\partial x_n} f(x) \end{bmatrix}$

Definition 95 (Conservative vector field and potential) *A vector field f on $\mathbb{R}^n$ i.e. $f : \mathbb{R}^n \to \mathbb{R}^n$ is called a conservative vector field on $\mathbb{R}^n$ if $f = \nabla g$ for some scalar field $g : \mathbb{R}^n \to \mathbb{R}$. In this situation g is called a potential function for f.*

Not all vector fields are conservative, but such fields do arise frequently in physics. For example, observe that gravitational field $\mathfrak{G}$ is a conservative vector field. Because $\mathfrak{G} = \nabla f_{\mathfrak{G}}$ where the scalar field $f_{\mathfrak{G}}$ is defined by

$$f_{\mathfrak{G}}(x, y, z) = \frac{mMG}{\| (x, y, z) \|}$$

This $f_{\mathfrak{G}}$ is called gravitational potential.

Exercise 96 *Check whether electric field $\mathfrak{E}$ is a conservative vector field. If so find electric potential.*

15 Line integrals

As of now we haven't defined curves or surfaces. There is a field of mathematics called "Differential geometry" where we need to define curves and surfaces properly. Then we get a concept of parametrized curves and parametrized surfaces. But to do integrations over curves or surfaces in a "Multivariable calculus" course we directly jump into the simplified definition of parametrized curves and surfaces.

15.1 Curves on $\mathbb{R}^2$

In this section we define an integral that is similar to a single integral except that instead of integrating over an interval $[a, b]$, we integrate over a curve C. Such integrals are called line integrals, although "curve integrals" would be better terminology. They were invented in the early 19th century to solve problems involving fluid flow, forces, electricity, and magnetism.

Definition 97 (Curve/Parametrized curve) *Let I be an interval in $\mathbb{R}$. (It can be closed, open or semi-open). A parametrized curve α on a subset U of $\mathbb{R}^2$ is a smooth [1] function $\alpha : I \to U$. The image C of this function α i.e. $C = \alpha(I)$ is called a curve on U.*

Let $I = [a, b]$ and $\alpha : I \to \mathbb{R}^2$ be smooth. Note that $\alpha(t) = (\alpha_1(t), \alpha_2(t))$ for all $t \in I$. Also α_1 and α_2 are smooth functions from I to $\mathbb{R}$ because they are just compositions of projections and smooth function α.

So if we say C is a curve on $\mathbb{R}^2$ there is a parametrized equation of C i.e.

$$C = (x, y) \ where \ x = x(t) \ and \ y = y(t); \ a \leq t \leq b$$

Don't get confused between the same notation x and y for variables and smooth function. We just mean that there exists $\alpha : [a, b] \to \mathbb{R}^2$ is a smooth function such that $\alpha(t) = (x(t), y(t))$ for al $a \leq t \leq b$. And $C = Im(\alpha)$.

So whenever we will talk about a parametrized curve we will think of the smooth function α but geometrically the curve is just $Im(\alpha)$. We can call both of them curves depending on the context. But technically speaking $C = Im(\alpha)$ is the curve and α is called it's parametrization. So we call α as parametrized curve. The function α is abstract, it doesn't have a physical interpretation so this is just parametrization but we can see $Im(\alpha)$ as a physical object in $\mathbb{R}^2$ so that is the curve.

Velocity, speed and length

Let $\alpha : I \to \mathbb{R}^2$ be a parametrized curve such that $\alpha(t) = (x(t), y(t))$ for all $t \in I$. Velocity of α at t_0 is just the derivative of α at t_0. So for any $t \in I$ it is defined by

$$\dot{\alpha}(t) = D\alpha(t) = \begin{bmatrix} \frac{d}{dt}x(t) \\ \frac{d}{dt}y(t) \end{bmatrix}$$

[1] In general in the definition of parametrized curve α we just require α to be continuous, but in the context of integration almost everywhere we need our curve to be smooth. So we fix the smoothness of α as a definition throughout.

Note that as x and y are functions from $I \to \mathbb{R}$ we have $\frac{d}{dt}x(t) = \frac{\partial}{\partial t}x(t)$.
The speed of α at t_0 is the norm of velocity of it at t_0. So for any $t \in I$ it is defined by

$$\| \dot{\alpha}(t) \| = \sqrt{(\frac{d}{dt}x(t))^2 + (\frac{d}{dt}y(t))^2}$$

Now we are going to define the most crucial tool for line integrals. i.e. the length of a parametrized curve. It is nothing but the integral of the speed of it over I i.e. $l(\alpha) = \int_I \| \dot{\alpha}(t) \| dt$.

Definition 98 (Length of a curve/parametrized curve) *Let $\alpha : [a, b] \to \mathbb{R}^2$ be a parametrized curve such that $\alpha(t) = (x(t), y(t))$ for all $a \leq t \leq b$. Hence $C = Im(\alpha)$ is the curve on $\mathbb{R}^2$. Then the length of C is same as the length of the parametrization α given by*

$$l(C) = l(\alpha) = \int_a^b \sqrt{(\frac{d}{dt}x(t))^2 + (\frac{d}{dt}y(t))^2 dt}$$

Now if a curve C on $\mathbb{R}^2$ is given to us we know there exists a parametrization α such that $Im(\alpha) = C$. But now the question we need to ask is whether such parametrization is unique and the answer is no. There can be some other parametrization β of C such that $Im(\beta) = C$. So now we state that the length of a curve C doesn't depend on the choice of its parametrization α. Readers can take this statement for granted or try to prove. The proof is not so trivial but easy.

15.2 Orientation of curves

Once again this "orientation" is a technical term one can rigorously learn after properly define curves and surfaces. But as we have only defined it through parametrization we will only have a geometrical intuition of orientation.

Let $\alpha : [a, b] \to \mathbb{R}^2$ is parametrization of curve C. We need to understand that we can traverse through C in two ways, one is from $\alpha(a)$ to $\alpha(b)$ with increasing values of t. And another one is the reverse of it i.e. from $\alpha(b)$ to $\alpha(a)$ with decreasing values of t. Depending on these directions we can put arrows on our curve. Note that there can be exactly two choices of arrows in a particular point of C depending on the way we traverse. In case of the first way of traversing we say C has a positive orientation. The next way is called negative orientation of C.

A oriented curve C of a parametrized curve α on $\mathbb{R}^2$ is a curve together with a choice of orientation of it. The following example will explain it better. Consider the boundary of the unit circle i.e. $C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ with a parametrization $\alpha : [0, 2\pi] \to \mathbb{R}^2$ by $\alpha(\theta) = (\cos\theta, \sin\theta)$. Figure 11 shows the two different orientations of C. In the first image we traverse through the curve from $\alpha(2\pi)$ to $\alpha(0)$ i.e. with decreasing value of t. Hence this has a negative orientation. For the reverse traversing the next one has positive orientation. So any curve can have exactly two oriented curves obtained from it.

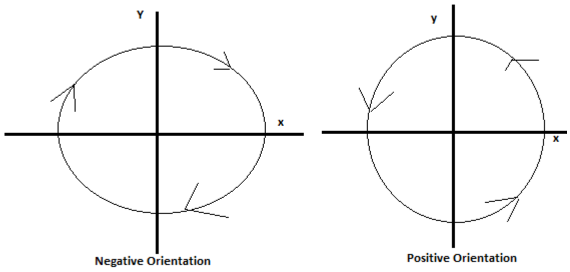

Figure 11: Orientation of unit circle in $\mathbb{R}^2$

15.3 Line integration on $\mathbb{R}^2$

Now we are ready to define our main concept i.e line integrals. C is a curve on $\mathbb{R}^2$. Let $\alpha : [a,b] \to \mathbb{R}^2$ be the parametrization of C such that $\alpha(t) = (x(t), y(t))$ for all t. $f : C \to \mathbb{R}$ is a bounded function with it's set of point of discontinuity having measure zero. Then the **line integral of** f **along** C equals

$$\int_C f(x,y)dC = \int_a^b f(x(t),y(t))\sqrt{(\frac{d}{dt}x(t))^2 + (\frac{d}{dt}y(t))^2}dt$$

We can also think of this in terms of Riemann sum. Recall the special case of Riemann sum discussed at the end of section 9. Let $\Delta x = \frac{b-a}{n}$ and hence $a, a+\Delta x, a+2\Delta x, \cdots, a+n\Delta x$ is a partition of $[a,b]$ Let $t_i = a + i\Delta x$. Let $\overline{t_i}$ be the midpoint of the interval $[a+(i-1)\Delta x, a+i\Delta x]$. So informally speaking, we are trying to find a partition of C by applying α to the t_i s. But the concept of length of partition will be a bit different here. Define $\Delta c_i = \int_{t_{i-1}}^{t_i} \parallel \dot{\alpha}(t) \parallel dt$. In this set up the **line integral of** f **along** C equals

$$\int_C f(x,y)dC = \lim_{n\to\infty} \Sigma_{i=1}^n f(x(\overline{t_i}), y(\overline{t_i}))\Delta c_i$$

Again we state a non-trivial fact that the line integration of f along C doesn't depend on the choice of parametrization if the orientation of them are same. i.e. if orientation of C is same (either positive for both or negative for both) for two parametrized curves α and β then the following are equal.

$$\int_a^b f(\alpha(t)) \parallel \dot{\alpha}(t) \parallel dt = \int_C f dC = \int_c^d f(\beta(t)) \parallel \dot{\beta}(t) \parallel dt$$

where $[a,b]$ and $[c,d]$ are domains of α and β respectively. If orientations of C are different for two different curves then also the absolute value of the integral is same but one of them is positive and another one is negative.

Now we will see examples. Let C be the upper half of the unit circle $x^2 + y^2 = 1$. Calculate $\int_C (2 + x^2 y)dC$. First we find any parametrization of C. Define a parametrization $\alpha : [0,\pi] \to \mathbb{R}^2$ by $\alpha(t) = (\cos t, \sin t)$. Now $\int_C (2 + x^2 y)dC$
$= \int_0^\pi (2 + \cos^2 t \sin t)\sqrt{(\frac{d}{dt}\cos t)^2 + (\frac{d}{dt}\sin t)^2}dt = \int_0^\pi (2 + \cos^2 t \sin t)\sqrt{\sin^2 t + \cos^2 t}dt = \int_0^\pi (2 + \cos^2 t \sin t)dt = \left[2t - \frac{\cos^3 t}{3}\right]_0^\pi = 2\pi + \frac{2}{3}$.

There are physical interpretations of line integrals. For example, suppose we are given the semi circular wire C as the above example. Let f denote the mass density $\rho(x, y)$ of the wire. Then the above integral calculates the mass m of the wire. And if (X, Y) is the centre of mass of the wire then $X = \frac{1}{m}\int_C x\rho(x, y)dC$ and $Y = \frac{1}{m}\int_C y\rho(x, y)dC$.

Now we define **line integrals of f along C with respect to a single variable x or y**. With the set up of the previous definition now we define two more things e.g.

$$\int_C f(x, y)dx = \int_a^b f(x(t), y(t))x'(t)dt \quad ; \quad \int_C f(x, y)dy = \int_a^b f(x(t), y(t))y'(t)dt$$

Exercise 99 *Let C be a curve on $\mathbb{R}^2$ and P and Q be two continuous functions from C to $\mathbb{R}^2$. Then $\int_C P + Q = \int_C P + \int_C Q$.*

15.4 Curves on $\mathbb{R}^3$

Previously on $\mathbb{R}^2$ we had a parametrization $\alpha(t) = (x(t), y(t))$. Now when we define parametrization of a curve on $\mathbb{R}^3$ this becomes a smooth function from some interval I of $\mathbb{R}$ to $\mathbb{R}^3$ i.e. $\alpha(t) = (x(t), y(t), z(t))$. Similarly we can define velocity, length and speed of it as earlier.

For example let C be the straight line joining $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$. Then a parametrization of C is the following. For $0 \leq t \leq 1$ define $x(t) = tu_1 + (1-t)v_1$, $y(t) = tu_2 + (1-t)v_2$ and $z(t) = tu_3 + (1-t)v_3$.

15.5 Line integrations on $\mathbb{R}^3$

C is a curve on $\mathbb{R}^3$. Let $\alpha : [a, b] \to \mathbb{R}^3$ be the parametrization of C such that $\alpha(t) = (x(t), y(t), z(t))$ for all t. $f : C \to \mathbb{R}$ is a bounded function with it's set of point of discontinuity having measure zero. Then the **line integral of f along C** equals

$$\int_C f(x, y)dC = \int_a^b f(x(t), y(t), z(t))\sqrt{(\frac{d}{dt}x(t))^2 + (\frac{d}{dt}y(t))^2 + (\frac{d}{dt}z(t))^2}dt$$

Similarly we can define the Riemann sum criterion with a little modification of that of $\mathbb{R}^2$.

15.6 Curves and line integral on higher dimensions

Having a solid background and concept from $\mathbb{R}^2$ and $\mathbb{R}^3$ we can develop on it. The parametrized curve on $\mathbb{R}^n$ is a smooth function α from some interval I of $\mathbb{R}$ to $\mathbb{R}^n$ i.e. $\alpha(t) = (\alpha_1(t), \alpha_2(t), \cdots, \alpha_n(t))$ for all $t \in I$. Now we can define line integration on $\mathbb{R}^n$ in similar manner.

15.7 Line integrals of vector fields

From the beginning (section 9) till now we have only integrate real valued function. But now for the first time we will go one step further and integrate vector fields on $\mathbb{R}^n$

i.e. we will integrate functions of n components in general. The concept of integrating vector fields was inspired by a physical activity i.e. calculating work done by a given force field in moving a particle from one point to another.

Though we will define for any n, in practical purposes we always have two cases $n = 2$ or 3. Let C be a curve on $\mathbb{R}^n$ with it's parametrization $r(t); a \leq t \leq b$. So $r(t) = ((r_1(t)), (r_2(t)), \cdots, (r_n(t)))$. Let F be a continuous vector field on C i.e. $F : C \to \mathbb{R}^n$ continuous. Define $T(t) = \dfrac{\dot{r}(t)}{\| \dot{r}(t) \|}$. Then the integral of F along C is defined by

$$\int_C F \cdot dr = \int_a^b F(r(t)) \cdot \dot{r}(t)\; dt = \int_C F \cdot T\; dC$$

One can informally think of $T(t)$ as an unit tangent vector to C at $r(t)$. For concrete definition we have to wait till section 17.3.

Exercise 100 *Find the work by the force field* $F(x, y) = x^2\hat{\boldsymbol{i}} - xy\hat{\boldsymbol{j}}$ *in moving a particle along the quarter circle* $r(t) = \cos t\hat{\boldsymbol{i}} + \sin t\hat{\boldsymbol{j}};\ 0 \leq t \leq \frac{\pi}{2}$.

Solution : $F(r(t)) = \cos^2 t\hat{\mathbf{i}} - \cos t \sin t\hat{\mathbf{j}}$ and $\dot{r}(t) = -\sin t\hat{\mathbf{i}} + \cos t\hat{\mathbf{j}}$. Therefore the work done by $F = \int_C F \cdot dr = \int_0^{\frac{\pi}{2}} F(r(t)) \cdot \dot{r}(t)\; dt = \int_0^{\frac{\pi}{2}} \left(\cos^2 t\hat{\mathbf{i}} - \cos t \sin t\hat{\mathbf{j}}\right)\left(-\sin t\hat{\mathbf{i}} + \cos t\hat{\mathbf{j}}\right) = \int_0^{\frac{\pi}{2}} -2\cos^2 t \sin t dt = \dfrac{2\cos^3 t}{3}\Big]_0^{\frac{\pi}{2}} = -\frac{2}{3}$.

Exercise 101 *Let C and $-C$ denote the curve C on $\mathbb{R}^n$ with positive and negative orientation respectively. r is the parametrization of C. F is a vector field on C. Show that* $\int_{-C} F \cdot dr = -\int_C F \cdot dr$

Exercise 102 *Let C be the curve given by* $x(t) = t, y(t) = t^2, z(t) = t^3;\ 0 \leq t \leq 1$. *Calculate* $\int_C xye^{yz}\; dy$

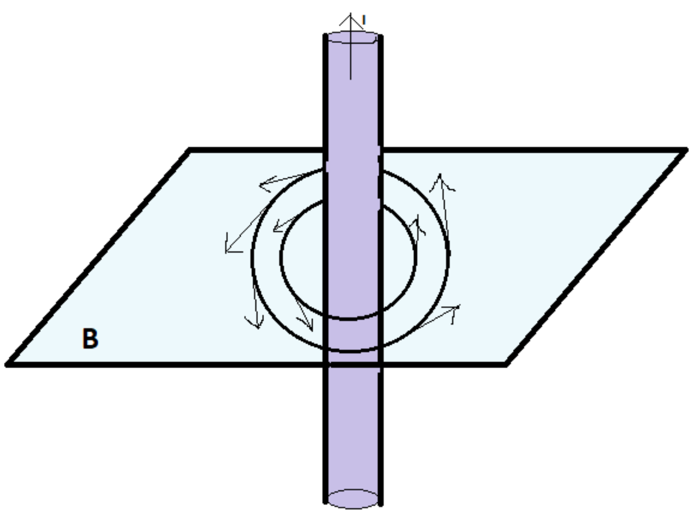

Figure 12: Exercise - 104

Exercise 103 *Show that a constant force field does zero work on a particle that moves uniformly (constant speed) once around the circle $x^2 + y^2 = 1$. Is this also true if the force field is given by $F(x, y) = \pi(x, y)$?*

Exercise 104 *Experiments show that a steady current I in a long wire produces a magnetic field B that is tangent to any circle that lies in the plane perpendicular to the wire and whose center is the axis of the wire (as in the figure). Ampère's Law relates the electric current to its magnetic effects and states that $\int_C B \cdot dr = \mu_0 I$ where I is the net current that passes through any surface bounded by a closed curve C, and μ_0 is a constant called the permeability of free space. By taking C to be a circle with radius, show that the magnitude $\mathcal{B}$ of the magnetic field B at a distance r from the center of the wire is $\mathcal{B} = \frac{\mu_0 I}{2\pi r}$*

15.8 The fundamental theorem for line integration

Recall theorem 79 (section 12) i.e. if F is a function of class $\mathcal{C}^1$ on $[a, b]$ then

$$\int_a^b F'(x)dx = F(b) - F(a)$$

We also called this **the Net Change Theorem**. The integral of a rate of change is the net change. If we think of the gradient vector ∇f of a function of two or three variables as a sort of derivative of f, then the following theorem can be regarded as a version of the Fundamental Theorem for line integrals.

Theorem 105 *Let C be a curve on $\mathbb{R}^n$ (Mostly we consider $n = 2$ or 3) with it's parametrization $r(t); a \leq t \leq b$. Let f be a differentiable function from C to $\mathbb{R}$ such that its gradient vector field ∇f is continuous i.e. f is of class $\mathcal{C}^1$. Then*

$$\int_C \nabla f \cdot dr = f(r(b)) - f(r(a))$$

The proof of this theorem directly follows from the definition of line integration of vector field and the fundamental theorem of calculus.

Theorem 105 says that we can evaluate the line integral of a conservative vector field (the gradient vector field of the potential function f) simply by knowing the value of it at the endpoints of C. In fact, Theorem 105 says that the line integral of ∇f is the net change in f.

15.9 Independence of path

Let C be a curve on $\mathbb{R}^n$ such that it has parametrization $\alpha : I \to \mathbb{R}^n$. Let a and b be the endpoints of I with $a < b$. Then $\alpha(a)$ and $\alpha(b)$ are the initial and end points respectively of the positive oriented C and $\alpha(b)$ and $\alpha(a)$ are the initial and end points respectively of the negative oriented $-C$.

Exercise 106 *Prove or disprove. Let F be a vector field on $\mathbb{R}^n$ and C_1 and C_2 be two different curves on $\mathbb{R}^n$ with same initial and end points. Then $\int_{C_1} F \cdot dr_1 = \int_{C_2} F \cdot dr_2$ where r_1 and r_2 are parametrizations of C_1 and C_2 respectively.*

Hint : This exercises asks whether the value of line integration of a vector field between two points depends on the curve joining them. The answer is yes. i.e. there exists C_1 and C_2 such that $\int_{C_1} F \cdot dr_1 \neq \int_{C_2} F \cdot dr_2$. Let F be a vector field on $\mathbb{R}^2$ defined by $F(x,y) = y^2\hat{\mathbf{i}} + x\hat{\mathbf{j}}$. Take C_1 as the line segment joining $(-5,-3)$ to $(0,2)$ and C_2 as the parabola $x = 4 - y^2$ from $(-5,-3)$ to $(0,2)$. Find suitable parametrization r_1 and r_2 of them and calculate them to show that $\int_{C_1} F \cdot dr_1 = -\frac{5}{6}$ and $\int_{C_2} F \cdot dr_2 = \frac{245}{6}$.

So we have seen that in general the line integral of a vector field is path dependent. But the next theorem ($theorem$ 107) shows that if the vector field is conservative i.e. it's a gradient of some scalar field then the line integral of a vector field is path independent. Before that we define a notion of closed curve.

A curve on $\mathbb{R}^n$ is closed if it has same initial and end points. i.e. Let C be a curve on $\mathbb{R}^n$ with parametrization $r(t);\ a \leq t \leq b$. Then C is said to be a closed curve if $r(a) = r(b)$.

Theorem 107 *Let f be a continuous vector field on an open connected set S of $\mathbb{R}^n$. Then the following are equivalent.*

1. *f is gradient of some scalar field on S i.e. $f = \nabla g$ for some $g : S \to \mathbb{R}$.*

2. *The line integral of f is independent of the path. i.e. If $a, b \in S$ and C_1 and C_2 are two curves on S with initial point a and end point b, then $\int_{C_1} f = \int_{C_2} f$.*

3. *The line integral of f is zero along every closed path on S i.e. if C is a closed curve on S then $\int_C f dC = 0$.*

The above theorem gives a criterion of determining whether a vector field is conservative (gradient field of some scalar field) in terms of line integration. But now we will see some other criterion for a vector field to be conservative.

Theorem 108 *Let f be a vector field of class $\mathcal{C}^1$ on an open connected set S of $\mathbb{R}^n$. Then f is conservative if and only if $\dfrac{\partial f_i}{\partial x_j}(x) = \dfrac{\partial f_j}{\partial x_i}(x)$ for all $1 \leq i, j \leq n$ and for all $x \in S$.*
The above statement is same as saying f is gradient of some scalar field on S if and only if the matrix $\mathcal{J}f(x)$ is symmetric for all $x \in S$.

Exercise 109 *Define a vector field F on $\mathbb{R}^2$ by $F(x,y) = (ye^x + \sin y)\hat{\boldsymbol{i}} + (e^x + x\cos y)\hat{\boldsymbol{j}}$. Check whether F is conservative. If so, find it's potential scalar field.*

Exercise 110 *Let C denote the curve the earth follows to revolve around the sun. Suppose the earth was a km apart from the sun on $March\ 20, 2020$ and it was b km apart from the sun on $August\ 20, 2020$. Let G be the gravitational constant and m_E and m_S are the masses of the earth and the sun respectively. Find the work done by the gravitational force due to movement of the earth around the sun in between two given dates.*

We will use a new notation now. If C is a closed curve we denote $\int_C f$ by $\oint_C f$.

16 Green's theorem

Let C be a curve with parametrization $r(t); a \leq t \leq b$. We already know that C closed if $r(a) = r(b)$. We say C is simple if α is one to one in $[a, b]$. We say a closed curve C is simple if α is one to one in $[a, b)$.

On $\mathbb{R}^2$ Green's Theorem gives the relationship between a line integral around a simple closed curve C and a double integral over the plane region R bounded by C. (See Figure 13.) We assume that R consists of all points inside C as well as all points on C. In stating Green's Theorem we use the convention that the positive orientation of a simple closed curve refers to a single counter-clockwise traversal of C. Thus if C is given by the vector function $r(t); a \leq t \leq b$, then the region is always on the left as the point traverses. (See Figure 13.)

Let f be a vector field on $\mathbb{R}^2$ i.e. $f(v) = (f_1(v), f_2(v)$ for scalar fields f_1 and f_2 on $\mathbb{R}^2$. Then observe that $\int_C f \cdot dr = \int_a^b f(r(t)) \cdot \dot{r}(t)\ dt = \int_a^b f(x(t), y(t)) \cdot (\dot{x}(t), \dot{y}(t))\ dt = \int_a^b \big[f_1(x(t), y(t))\dot{x}(t) + f_2(x(t), y(t))\dot{y}(t)\big] dt = \int_C f_1\ dx + f_2\ dy$. (Check section 15.2 to see the meaning of $\int_C g(x, y)\ dx$ and $\int_C g(x, y)\ dy$)

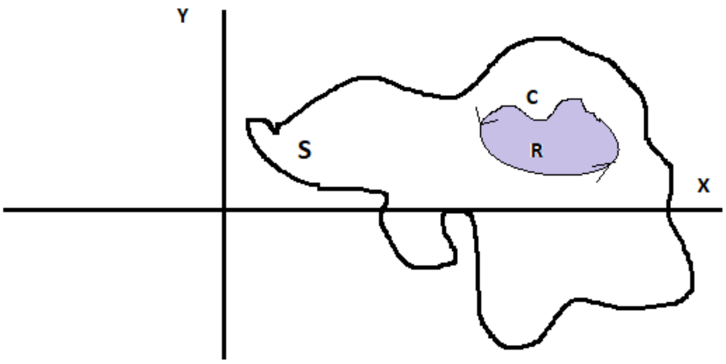

Figure 13: Green's theorem

Theorem 111 (Green's theorem) *Let S be an open set in $\mathbb{R}^2$. P and Q are two scalar fields on S of class $\mathcal{C}^1$. C is a simple closed curve on S. Let R be the plane region enclosed by C and C itself i.e. $R = C \cup I(C))$ where $I(C)$ is the region inside the curve C. Then*

$$\int\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\ dx\ dy = \oint_C P\ dx + Q\ dy$$

Exercise 112 *Evaluate the following*

1. $\oint_C (3y - e^{\sin x})dx + (7x + \sqrt{y^4 + 1})dy$ *where C is the circle $x^2 + y^2 = 9$.*

2. $\oint_C y^2\ dx + 3xy\ dy$ *where C is the boundary of the semi-circular region D in the upper half plane between the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$.*

Exercise 113 *Use Green's theorem to prove the change of variable formula for double integrals.*

17 Curl and Divergence

17.1 Curl

Let $F(x,y,z) = P\hat{\mathbf{i}} + Q\hat{\mathbf{j}} + R\hat{\mathbf{k}}$ is a vector field on $\mathbb{R}^3$ such that P, Q, R are scalar fields and all of their first order partial derivatives exist. Then

$$curl\ F = (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})\hat{\mathbf{i}} + (\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x})\hat{\mathbf{j}} + (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\hat{\mathbf{k}}$$

Define an operator $\nabla = \frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}$. Though this 'operator' is a technical term, one can think of this as a symbol with the following meanings associated with it.If f is a scalar field on $\mathbb{R}^3$ then

$$\nabla f = \frac{\partial f}{\partial x}\hat{\mathbf{i}} + \frac{\partial f}{\partial y}\hat{\mathbf{j}} + \frac{\partial f}{\partial z}\hat{\mathbf{k}}$$

This we already knew from the concepts of Jacobian. But our next goal is to define $curl$ in terms of this ∇ operator. Let $u = a\hat{\mathbf{i}} + b\hat{\mathbf{j}} + c\hat{\mathbf{k}}$ and $v = d\hat{\mathbf{i}} + e\hat{\mathbf{j}} + f\hat{\mathbf{k}}$ be two vectors of $\mathbb{R}^3$. We define their cross product as follows.

$$u \times v = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a & b & c \\ d & e & f \end{vmatrix} = \big(bf - ec\big)\hat{\mathbf{i}} + \big(dc - af\big)\hat{\mathbf{j}} + \big(ae - bd\big)\hat{\mathbf{k}}$$

So for the previously defined F we have

$$\nabla \times F = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})\hat{\mathbf{i}} + (\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x})\hat{\mathbf{j}} + (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\hat{\mathbf{k}} = curl\ F$$

For example, if $F(x,y,z) = xz\hat{\mathbf{i}} + xyz\hat{\mathbf{j}} - y^2\hat{\mathbf{k}}$ then we can calculate and observe that $curl\ F = -y(2+x)\hat{\mathbf{i}} + x\hat{\mathbf{j}} + yz\hat{\mathbf{k}}$. The next theorem says that curl of a gradient field is zero.

Theorem 114 *Let f be scalar field on $\mathbb{R}^3$ and f is of class C^2 then*

$$curl\ (\nabla f) = 0$$

Proof : We have $curl\ (\nabla f) = \nabla \times (\nabla f) = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} = (\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y})\hat{\mathbf{i}} + (\frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z})\hat{\mathbf{j}} + (\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x})\hat{\mathbf{k}} = 0\hat{\mathbf{i}} + 0\hat{\mathbf{j}} + 0\hat{\mathbf{k}} = 0$. This 0 comes from Claircut's theorem.

The converse of this theorem is not true in general. But the next theorem shows that the converse is true if the domain of the vector field is whole $\mathbb{R}^3$.

Theorem 115 *Let $F : \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field such that all its component functions are of class $\mathcal{C}^1$. i.e. if $F(x) = (F_1(x), F_2(x), F_3(x))$ then F_1, F_2, F_3 are of class $\mathcal{C}^1$.*
Let $curl\ F = 0$. Then F is a conservative vector field i.e. $F = \nabla f$ for some scalar field f on $\mathbb{R}^3$.

17.2 Divergence

Let $F(x, y, z) = P\hat{\mathbf{i}} + Q\hat{\mathbf{j}} + R\hat{\mathbf{k}}$ is a vector field on $\mathbb{R}^3$ such that P, Q, R are scalar fields and $\frac{\partial P}{\partial x}, \frac{\partial Q}{\partial y}, \frac{\partial R}{\partial z}$ exist. Then

$$div\ F = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

So stating in terms of ∇ operator we have

$$\nabla \cdot F = \left(\frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}\right) \cdot \left(P\hat{\mathbf{i}} + Q\hat{\mathbf{j}} + R\hat{\mathbf{k}}\right) = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = div\ F$$

The next theorem shows that divergence of curl is zero.

Theorem 116 *Let $F(x, y, z) = P\hat{\boldsymbol{i}} + Q\hat{\boldsymbol{j}} + R\hat{\boldsymbol{k}}$ is a vector field on $\mathbb{R}^3$ such that P, Q, R are scalar fields of class $\mathcal{C}^2$. Then*

$$div\ curl\ F = 0$$

Proof : $div\ curl\ F = \nabla \cdot (\nabla \times F)$
$= \left(\frac{\partial}{\partial x}\hat{\mathbf{i}} + \frac{\partial}{\partial y}\hat{\mathbf{j}} + \frac{\partial}{\partial z}\hat{\mathbf{k}}\right) \cdot \left(\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\hat{\mathbf{k}} \right)$
$= \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)$
$= \frac{\partial^2 P}{\partial y \partial z} - \frac{\partial^2 P}{\partial z \partial y} + \frac{\partial^2 Q}{\partial z \partial x} - \frac{\partial^2 Q}{\partial x \partial z} + \frac{\partial^2 R}{\partial x \partial y} - \frac{\partial^2 R}{\partial y \partial x} = 0$. This 0 comes from Claircut's theorem.

Laplace operator

As we already defined the ∇ operator now we will define something more. Let f be a scalar field on $\mathbb{R}^3$ with it's gradient vector field ∇f. Then we have

$$div(\nabla f) = \nabla \cdot (\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

So Laplacian / Laplace operator ∇^2 for a scalar field f is defined as follows

$$\nabla^2(f) = \nabla \cdot \nabla(f)$$

17.3 Vector forms of Green's theorem

Let $F = P\hat{\mathbf{i}} + Q\hat{\mathbf{j}}$ be a vector field on $\mathbb{R}^2$. Assume all the hypothesis of the Green's theorem (theorem 111) and recall that we have

$$\int\int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\, dy = \oint_C P\, dx + Q\, dy = \oint_C F \cdot dr$$

Let's change the set up a bit and regard F as a vector field on $\mathbb{R}^3$ by $F = P\hat{\mathbf{i}} + Q\hat{\mathbf{j}} + 0\hat{\mathbf{k}}$. Then we have

$$curl\ F = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & 0 \end{vmatrix} = (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\hat{\mathbf{k}}$$

Therefore $curl\ F \cdot \hat{\mathbf{k}} = (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})\hat{\mathbf{k}} \cdot \hat{\mathbf{k}} = (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})$. So now we are ready to re-write the Green's theorem in vector form. And that is as follows

$$\oint_C F \cdot dr = \oint_C F \cdot T\ dC = \int\int_R (curl\ F) \cdot \hat{\mathbf{k}}\, dA$$

Note that this dA is just a simplified notation for $dxdy$ which represents area integral i.e. integration on $\mathbb{R}^2$. Similarly in $\mathbb{R}^3$ we use dV in place of $dxdydz$ and it represents volume integral. Find the meaning of the notation $\oint_C F \cdot T\ dC$ where we defined line integrations of vector field.

The above equation expresses the line integral of the tangential component of F (See figure 14) along C as the double integral of the vertical component of $curl\ F$ over the region R enclosed by C. We now derive a similar formula involving the normal component of F.

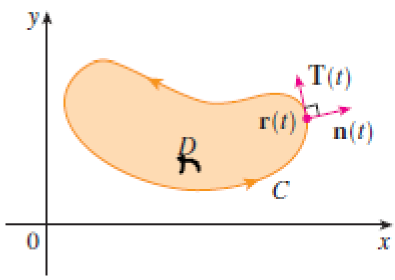

Figure 14: Tangent and normal to C at $r(t)$

Let C has the parametrization $r(t) = (x(t), y(t)); a \leq t \leq b$. Then we define unit tangent vector $T(t)$ to C at the point $r(t)$ as

$$T(t) = \frac{\dot{r}(t)}{\|\ \dot{r}(t)\ \|} = \frac{x'(t)}{\|\ \dot{r}(t)\ \|}\hat{\mathbf{i}} + \frac{y'(t)}{\|\ \dot{r}(t)\ \|}\hat{\mathbf{j}}$$

Our next goal is to define a unit normal vector $n(t)$ to C at the point $r(t)$. Note that $n(t)$ must be perpendicular to $T(t)$ ie. $n(t) \cdot T(t) = 0$. So we have

$$n(t) = \frac{y'(t)}{\|\ \dot{r}(t)\ \|}\hat{\mathbf{i}} - \frac{x'(t)}{\|\ \dot{r}(t)\ \|}\hat{\mathbf{j}}$$

We can also take $-n(t)$ as normal but that will be inward to the curve whereas $n(t)$ is outward and considered widely as convention. (See figure 14)

First we define a function $F \cdot n$ from C to $\mathbb{R}$ by $(F \cdot n)(r(t)) = F(r(t)) \cdot n(t)$. Now from the definition of line integral we have $\oint_C F \cdot n \; dC = \int_a^b (F \cdot n)(r(t)) \parallel \dot{r}(t) \parallel dt = \int_a^b \big[\frac{P(x(t), y(t))\; y'(t)}{\parallel \dot{r}(t) \parallel} - \frac{Q(x(t), y(t))\; x'(t)}{\parallel \dot{r}(t) \parallel}\big] \parallel \dot{r}(t) \parallel dt = \int_a^b \big(P(x(t), y(t))\; y'(t) - Q(x(t), y(t))\; x'(t)\big) dt = \int_C P\; dy - Q\; dx = \int\int_R \big(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\big)\; dA$. (The last equality follows from Green's theorem.

So now we have second vector form of Green's theorem i.e.

$$\oint_C F \cdot n\; dC = \int\int_R div(F)\; dA$$

Exercise 117 *f and g are scalar fields of class $\mathcal{C}^2$. Use Green's theorem and the notations used there to prove the following identities.*

1. $\int\int_R f\nabla^2 g\; dA = \oint_C f(\nabla g) \cdot n\; dC - \int\int_R \nabla f \cdot \nabla g\; dA$
2. $\int\int_R \big(f\nabla^2 g - g\nabla^2 f\big) dA = \oint_C \big(f\nabla g - g\nabla f\big) \cdot n\; dC$

Exercise 118 *This exercise demonstrates a connection between the curl vector and rotations. Let B be a rigid body rotating about the z axis. The rotation can be described by the vector $w = -\omega\hat{\mathbf{k}}$, where ω is the angular speed of B, i.e, the tangential speed of any point P in B divided by the distance d from the axis of rotation. Let $r = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ be the position vector of P.*

1. *By considering the angle θ in figure 15, show that the velocity field of B is given by $v = -w \times r$*
2. *Show that $v = -\omega y\hat{\mathbf{i}} + \omega x\hat{\mathbf{j}}$*
3. *Show that $v = 2w$*

Exercise 119 *Maxwell's equations relating the electric field E and magnetic field B as they vary with time in a region containing no charge and no current can be stated as follows:*

- $div\; E = 0$
- $curl\; E = -\frac{1}{c}\frac{\partial H}{\partial t}$
- $div\; H = 0$
- $curl\; H = \frac{1}{c}\frac{\partial E}{\partial t}$

where c is the speed of the light. Use these equations to prove the following

1. $\nabla \times (\nabla \times E) = -\frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}$
2. $\nabla \times (\nabla \times H) = -\frac{1}{c^2}\frac{\partial^2 H}{\partial t^2}$
3. $\nabla^2 E = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}$
4. $\nabla^2 H = \frac{1}{c^2}\frac{\partial^2 H}{\partial t^2}$

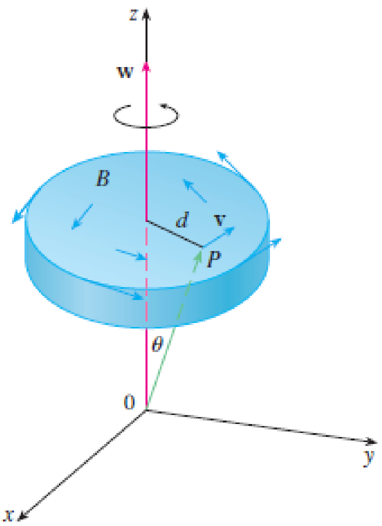

Figure 15: Exercise-118

18 Surface integrals

As we stated at the beginning of Line integrals (section 15) here also we will not give the definition of surfaces. We will only talk about parametrized surfaces like we only talked about parametrized curves while doing line integrals.

18.1 Surfaces on $\mathbb{R}^3$

In the same way that we describe a space curve C by a vector function $r(t)$ of a single parameter , we can describe a surface S by a vector function $r(u, v)$ of two parameters u and v.

Definition 120 (Parametrized surface) *Let Q be a rectangle of $\mathbb{R}^2$. A parametrized surface r on subset A of $\mathbb{R}^3$ is a function $r : Q \to A$. And the image of this function i.e. $r(Q) = S$ is called a surface on A.*

In the above definition if r is continuous we call S a continuous surface.

So if S is a surface on $\mathbb{R}^3$ then there exists a parametrization $r : Q \to \mathbb{R}^3$ where Q is a rectangle of $\mathbb{R}^2$ and $r(u, v) = (x(u, v), y(u, v), z(u, v))$ for all $(u, v) \in Q$. i.e. x, y, z are component functions of r.

Now we will see some examples of parametrizations of surfaces. Note that, given two points in $\mathbb{R}^2$ there exists a unique straight line joining them. Similarly, given three points in $\mathbb{R}^3$ we have an unique triangle. Let $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3), c = (c_1, c_2, c_3)$ be three points. We find the parametrized equation of the triangle having these as vertices. Let $Q = [0, 1] \times [0, 1]$. So we define $r : Q \to \mathbb{R}^3$ by $r(u, v) = (x(u, v), y(u, v), z(u, v))$ where $x(u, v) = uva_1+(1-u)vb_1+(1-v)c_1, y(u, v) = uva_2+(1-u)vb_2+(1-v)c_2, z(u, v) = uva_3 + (1 - u)vb_3 + (1 - v)c_3$.

Suppose we have some equation of surface given, where the third variable is a function of other two. In such cases we have a inherited parametrization. For example consider the paraboloid $z = x^2 + 2y^2$. We have the parametrization $x(u, v) = u, y(u, v) = v, z(u, v) = u^2 + 2v^2; (u, v) \in \mathbb{R}^2$.

Exercise 121 *Find a parametrization of the sphere* $x^2 + y^2 + z^2 = 1$. *(Hint : Spherical co-ordinate)*

Surface of revolution

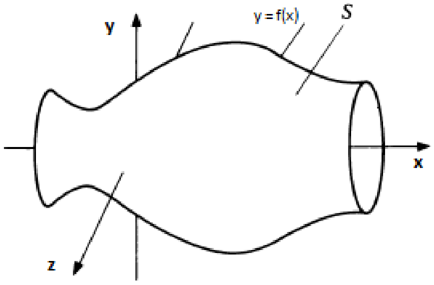

Figure 16: Surface of revolution

Here we start with a positive function $y = f(x)$ in xy plane and then rotate the graph of the function around x axis. So it forms a surface on $\mathbb{R}^3$. We find a parametrization of this surface S. Let $f : [a, b] \to \mathbb{R}$. Then the parametrization of S is the following. $r : [a, b] \times [0, 2\pi] \to \mathbb{R}^3$ by

$$x(u, v) = u, y(u, v) = f(u) \cos v, z(u, v) = f(u) \sin v$$

because for a fixed $u, (y, z)$ lies on the circle of radius $f(u)$ on yz plane. This S is called the surface of revolution of $y = f(x)$ because it is obtained by rotating $y = f(x)$. Observe that surface of revolution of a straight line segment is the curved surface of a cylinder.

Tangent plane

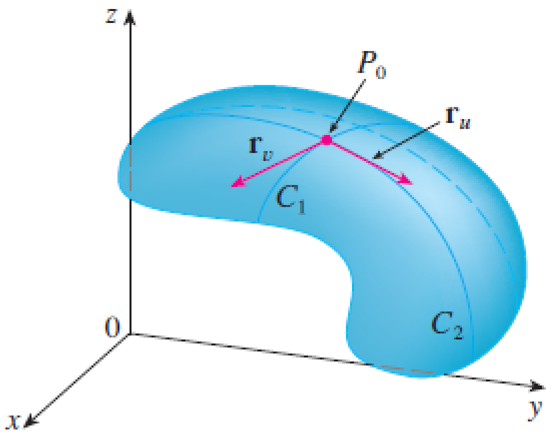

Figure 17: Tangent plane

If we have a curve on $\mathbb{R}^2$ and a point p on it then we can have only two tangents to C at p in two opposite directions. But in case of a surface we can have infinitely many tangents at a point. And all of them belong to a plane called tangent plane. Here we are given a surface S, its parametrization r with component functions x, y, z and a point $P_0 = (a, b, c)$ on it. We find the parametrization of the tangent plane at $P_0 = r(u_0, v_0)$. We fix two curves C_1 and C_2 on S passing through P_0 such that their parametrization domains are consistent with components of domain of r. More precisely, given r, define parametrization of C_1 by $r_1(v) = r(u_0, v)$ and that of C_2 by $r_2(u) = r(u, v_0)$. Then for C_1 we have a tangent

$$r_v(P_0) = \frac{\partial x}{\partial v}(u_0, v_0)\hat{\mathbf{i}} + \frac{\partial y}{\partial v}(u_0, v_0)\hat{\mathbf{j}} + \frac{\partial z}{\partial v}(u_0, v_0)\hat{\mathbf{k}}$$

Similarly for C_2 we have

$$r_u(P_0) = \frac{\partial x}{\partial u}(u_0, v_0)\hat{\mathbf{i}} + \frac{\partial y}{\partial u}(u_0, v_0)\hat{\mathbf{j}} + \frac{\partial z}{\partial u}(u_0, v_0)\hat{\mathbf{k}}$$

So our final goal is to find the parametrized equation of the plane with lines $r_u(P_0)$ and $r_v(P_0)$ on it. Let S_{P_0} denote the tangent plane of S at P_0. Then we define the parametrization $\alpha : \mathbb{R}^2 \to S_{P_0}$ by $\alpha(s, t) = P_0 + sr_u(P_0) + tr_v(P_0)$.

If $r_u(p) \times r_v(p) \neq 0$ for all $p \in S$ then S is called **Smooth surface.**

Normal to a surface

Follow all the previous notations. Then normal to S at $p \in S$ is defined by $r_u(p) \times r_v(p)$. The normal vector field on S is defined as $n : S \to \mathbb{R}^3$ by $n(p) = \dfrac{r_u(p) \times r_v(p)}{\| r_u(p) \times r_v(p) \|}$.

18.2 Orientation of surfaces

For a surface S we can have two choices of normal at each point $p \in S$. One is $n(p)$ and another one is $-n(p)$. Both of them are vector field on S. An oriented/orientable surface S is a surface with a fixed choice of normal vector field. S is called positively oriented if the normal vector field is n, and negatively oriented if the normal vector field is $-n$.

In case of curve the orientation was representing the direction of traversing. Here the orientation of a surface says whether it is the outward surface or the inward surface. (Figure 18)

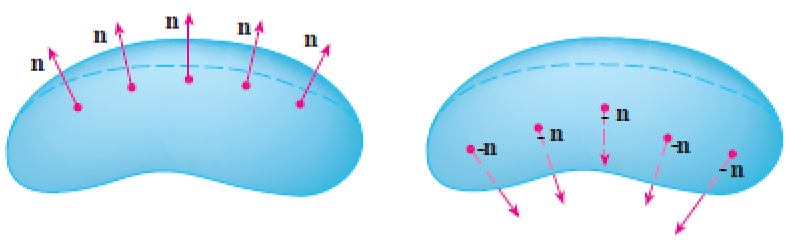

Figure 18: Orientation of surface

Once we define an orientable surface the question we can naturally ask is whether all surfaces are orientable. We give a quick cheap answer to this as "no". But showing the existence of such non-orientable surfaces is highly non-trivial. So we don't directly give the parametrized equation of the example. We illustrate how such surfaces can be formed and then talk about its parametrization.

Mobius strip

Take a rectangular strip (As shown in figure 19). Twist it exactly once and paste its opposite ends as shown in the figure. The object M formed is called Mobius strip.

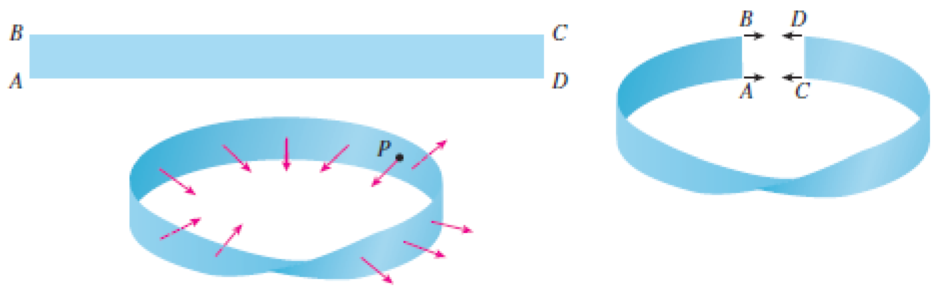

Figure 19: Mobius strip

We have already seen that orientable surfaces has two directions, inward and outward. For example, take a sphere. It has outer and inner orientations of the surface. But in Mobius strip there are no such two directions associated. Take a pencil. Start drawing a line on the Mobius strip. Keep moving on and one will end up reaching the initial position where he started after covering the whole of it. So it has no such inward and outward surface. But in case of sphere if we start drawing a line on the outer surface we will never reach inner surface.

Observe that the parametrization of M is $r : [0, 2\pi] \times [-1, 1] \to \mathbb{R}^3$ by $x(u, v) = (1 + \frac{v}{2} \cos \frac{u}{2}) \cos u, y(u, v) = (1 + \frac{v}{2} \cos \frac{u}{2}) \sin u, z(u, v) = \frac{v}{2} \sin \frac{u}{2}$.

18.3 Surface area

Let S be a surface on $\mathbb{R}^3$. It has a parametrization $r : Q \to \mathbb{R}^3$ i.e. $r(u, v) = x(u, v)\hat{\mathbf{i}} + y(u, v)\hat{\mathbf{j}} + z(u, v)\hat{\mathbf{k}}$ for all $(u, v) \in Q$. Now r_u and r_v are two vector fields on $\mathbb{R}^3$. So for all $p \in S, r_u(p) \times r_v(p) \in \mathbb{R}^3$. But $\| r_u(p) \times r_v(p) \| \in \mathbb{R}$. So the function f defined from S to $\mathbb{R}$ by $p \to \| r_u(p) \times r_v(p) \|$ is a real valued function. Then $f \circ r : Q \to \mathbb{R}$ is a real valued function on Q. i.e. $f \circ r(u, v) = \| r_u(r(u, v)) \times r_v(r(u, v)) \|$ We can talk about its integrability over Q.

Now we are ready to define the surface area of S. So area of $S = A(S) = \int_Q f \circ r$. We just write it below using a more convenient notation.

$$A(S) = \int \int_Q \| r_u \times r_v \| \, dA$$

where $r_u = \frac{\partial x}{\partial u}\hat{\mathbf{i}} + \frac{\partial y}{\partial u}\hat{\mathbf{j}} + \frac{\partial z}{\partial u}\hat{\mathbf{k}}$ and $r_v = \frac{\partial x}{\partial v}\hat{\mathbf{i}} + \frac{\partial y}{\partial v}\hat{\mathbf{j}} + \frac{\partial z}{\partial v}\hat{\mathbf{k}}$.

Graph of a function

Graph of a function is a surface where one variable can be expressed as a function of other two. Generally, the equation of the surface is given as $z = f(x, y)$ for $f : D \to \mathbb{R}$. So we have a natural parametrization i.e. $x(u, v) = u, y(u, v) = v, z(u, v) = f(u, v)$ for all $(u, v) \in D$. We denote this surface as $G(f)$.

Now we find the surface area of such a surface $G(f)$. First we see $r = u\hat{\mathbf{i}} + v\hat{\mathbf{j}} + f(u, v)\hat{\mathbf{k}}$. Hence $r_u = \hat{\mathbf{i}} + \frac{\partial f}{\partial u}\hat{\mathbf{k}}$ and $r_v = \hat{\mathbf{j}} + \frac{\partial f}{\partial v}\hat{\mathbf{k}}$. Now we calculate

$$r_u \times r_v = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 1 & 0 & \frac{\partial f}{\partial u} \\ 0 & 1 & \frac{\partial f}{\partial v} \end{vmatrix} = -\frac{\partial f}{\partial u}\hat{\mathbf{i}} - \frac{\partial f}{\partial v}\hat{\mathbf{j}} + \hat{\mathbf{k}}$$

. Thus we have $\| r_u \times r_v \| = \sqrt{1 + (\frac{\partial f}{\partial u})^2 + (\frac{\partial f}{\partial v})^2}$. Hence the surface area is

$$A(G(f)) = \int\int_D \sqrt{1 + (\frac{\partial f}{\partial u})^2 + (\frac{\partial f}{\partial v})^2} dA$$

Exercise 122 *Prove that the surface area of a sphere of radius a is $4\pi a^2$.*

Exercise 123 *Find the area of the part of the sphere $x^2 + y^2 + z^2 = 4z$ that lies inside the paraboloid $z = x^2 + y^2$.*

18.4 Surface integrals on $\mathbb{R}^3$

Now we are ready to define our main concept i.e surface integrals. S is a surface on $\mathbb{R}^3$. Q is a rectangle in $\mathbb{R}^2$. Let $r : Q \to \mathbb{R}^3$ be the parametrization of S such that $r(u, v) = (x(u, v), y(u, v), z(u, v))$ for all $(u, v) \in Q$. $f : S \to \mathbb{R}$ is a bounded function with it's set of point of discontinuity having measure zero. Then the **surface integral of f over** S equals

$$\int\int_S f(x, y, z)dS = \int\int_Q f(r(u, v)) \| r_u \times r_v \| dA$$

We can also think of this in terms of Riemann sum. Recall the special case of Riemann sum discussed at the end of section 9. Let $Q = [a, b] \times [c, d]$. $\Delta u = \frac{b-a}{m}$ and hence $a, a + \Delta u, a + 2\Delta u, \cdots, a + m\Delta u$ is a partition of $[a, b]$. Similarly $\Delta v = \frac{d-c}{n}$ and hence $c, c + \Delta v, c + 2\Delta v, \cdots, c + n\Delta v$ is a partition of $[c, d]$. Let $u_i = a + i\Delta u$. Let $\overline{u_i}$ be the midpoint of the interval $[a + (i - 1)\Delta u, a + i\Delta u]$. Similarly Let $v_j = c + j\Delta v$. Let $\overline{v_j}$ be the midpoint of the interval $[c + (j - 1)\Delta v, c + j\Delta v]$. So informally speaking, we are trying to find a partition of S by applying r to the (u_i, v_j) s. But the concept of area of partition will be a bit different here. Define $\Delta S_{ij} = \int_{u_{i-1}}^{u_i} \int_{v_{j-1}}^{v_j} \| r_u \times r_v \| dA$. In this set up the **surface integral of f along** S equals

$$\int\int_S f(x, y, z)dS = \lim_{m,n\to\infty} \Sigma_{i=1}^m \Sigma_{j=1}^n f(r(\overline{u_i}, \overline{v_j}))\Delta S_{ij}$$

For example if we take the surface $z = \phi(x, y)$ for $\phi : D \to \mathbb{R}$ and f is a bounded continuous real valued function on it, then

$$\int\int_{G(\phi)} f(x,y,z)dG(\phi) = \int\int_D f(u,v,\phi(u,v))\sqrt{1 + (\frac{\partial \phi}{\partial u})^2 + (\frac{\partial \phi}{\partial v})^2}dA$$

It is not necessary that we have to use the notation $dG(\phi)$ when we integrate over the surface $G(\phi)$. We can use a general notation dS for any surface integrals.

Exercise 124 *Let S be the unit sphere $x^2 + y^2 + z^2 = 1$. Compute the surface integral $\int\int_S x^2 dS$.*

Hint : Find a suitable parametrization of S by

$$x(u,v) = \sin u \cos v, y(u,v) = \sin u \sin v, z(u,v) = \cos u;\ 0 \le u \le \pi,\ 0 \le v \le 2\pi$$

. So $r(u,v) = \sin u \cos v\hat{\mathbf{i}} + \sin u \sin v\hat{\mathbf{j}} + \cos u\hat{\mathbf{k}}$. Now we compute $r_u \times r_v$.

$$r_u \times r_v = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{vmatrix} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \cos u \cos v & \cos u \sin v & -\sin u \\ -\sin u \sin v & \sin u \cos v & 0 \end{vmatrix}$$

$= \sin^2 u \cos v\hat{\mathbf{i}} + \sin^2 u \sin v\hat{\mathbf{j}} + \sin u \cos u\hat{\mathbf{k}}$. Hence $\| r_u \times r_v \| =$ $\sqrt{\sin^4 u \cos^2 v + \sin^4 u \sin^2 v + \sin^2 u \cos^2 u} = \sqrt{\sin^4 u + \sin^2 u \cos^2 u} = \sqrt{\sin^2 u} = \sin u$. Now $\int\int_S x^2 dS = \int_0^{2\pi}\int_0^{\pi} x^2 \ \| r_u \times r_v \| \ dudv = \int_0^{2\pi}\int_0^{\pi} (\sin u \cos v)^2 \ \sin u \ dudv =$ $\int_0^{2\pi}\int_0^{\pi} \cos^2 v \sin^3 u \ dudv$. Now use Fubini's theorem and compute to obtain $\frac{4\pi}{3}$.

Exercise 125 *Evaluate $\int\int_S y \, dS$ where S is the surface $z = x + y^2$, $0 \le x \le 1$, $0 \le y \le 2$.*

18.5 Surface integrals of vector fields

Let F be a vector field on a surface S on $\mathbb{R}^3$. Q is a rectangle in $\mathbb{R}^2$ and $r : Q \to \mathbb{R}^3$ is parametrization of S by $r(u,v) = (x(u,v), y(u,v), z(u,v))$ for all $(u,v) \in Q$. We define $r_u, r_v : Q \to \mathbb{R}^3$ by

$$r_u(u,v) = \frac{\partial x}{\partial u}(u,v)\hat{\mathbf{i}} + \frac{\partial y}{\partial u}(u,v)\hat{\mathbf{j}} + \frac{\partial z}{\partial u}(u,v)\hat{\mathbf{k}}$$

and

$$r_v(u,v) = \frac{\partial x}{\partial v}(u,v)\hat{\mathbf{i}} + \frac{\partial y}{\partial v}(u,v)\hat{\mathbf{j}} + \frac{\partial z}{\partial v}(u,v)\hat{\mathbf{k}}$$

So now define $n : S \to \mathbb{R}^3$ by

$$n(r(u,v)) = r_u(u,v) \times r_v(u,v)$$

Finally define $F \cdot n : S \to \mathbb{R}$ by

$$(F \cdot n)(r(u,v)) = F(r(u,v)) \cdot n(r(u,v))$$

Definition 126 (Flux) *Let F be a continuous vector field on a surface S on $\mathbb{R}^3$. Then the flux of F across S is*

$$\int\int_S F \cdot dS = \int\int_S F \cdot n \, dS = \int\int_Q (F \circ r) \cdot (r_u \times r_v) \, dA$$

The **Surface integral of** F **over** S is the flux of F across S.

Exercise 127 *Find the flux of the vector field $F(x, y, z) = z\hat{\mathbf{i}} + y\hat{\mathbf{j}} + x\hat{\mathbf{k}}$ across the unit sphere $x^2 + y^2 + z^2 = 1$. (**Hint :** Same as exercise 124)*

The motivation of surface integral came from a physical problem i.e. given a fluid flow in space (which is a vector field on $\mathbb{R}^3$) finding its flux. Although this concept of surface integrals arises in other physical situations as well. For instance, if $\mathcal{E}$ is the electric field then the surface integral $\int\int_S \mathcal{E} \cdot dS$ is called the electric flux of E through the surface S. One of the important laws of electro-statistics is **Gauss's Law**, which says that the net charge enclosed by a closed surface S is

$$Q = \varepsilon \int\int_S \mathcal{E} \cdot dS$$

where ε is a constant (called the permittivity of free space) that depends on the units used.

Another application of surface integrals occurs in the study of heat flow. Suppose the temperature at a point (x, y, z) in a body is $t(x, y, z)$. Then the heat flow is defined as the vector field $F = -K\nabla t$ where K is an experimentally determined constant called the conductivity of the substance. The rate of heat flow across the surface S in the body is then given by the surface integral

$$\int\int_S F \cdot dS = -K \int\int_S \nabla t \cdot dS$$

Exercise 128 *Evaluate $\int\int_S y^2 \, dS$ where S is the part of the sphere $x^2 + y^2 + z^2 = 4$ that lies inside the cylinder $x^2 + y^2 = 1$ and above xy plane.*

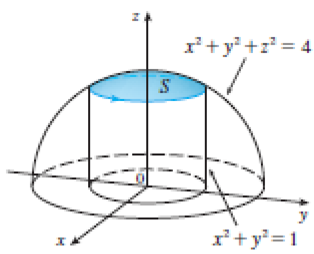

Figure 20: Exercise-128

Exercise 129 *Find the mass of a thin funnel in the shape of a cone $z = \sqrt{x^2 + y^2}$, $1 \leq z \leq 4$, if its density function is $\rho(x, y, z) = 10 - z$.*

Exercise 130 *Use Gauss's Law to find the charge enclosed by the cube with vertices* $(\pm 1, \pm 1, \pm 1)$ *if the electric field is given by* $\mathcal{E}(x, y, z) = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$.

Exercise 131 *The temperature at a point in a ball with conductivity* K *is inversely proportional to the distance of the point from the center of the ball. Find the rate of heat flow across a sphere of radius* a *with center at the center of the ball.*

Exercise 132 *Let* F *be an inverse square field on* $\mathbb{R}^3$ *i.e.* $F(r) = \dfrac{cr}{\| r \|^3}$ *for some constant* c*, where* $r = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$. *Show that the flux of* F *across a sphere* S *with centre at origin is independent of the radius of* S.

19 Stokes' theorem

Stokes' Theorem can be regarded as a higher-dimensional version of Green's Theorem. Whereas Green's Theorem relates a double integral over a plane region R to a line integral around its plane boundary curve C , Stokes' Theorem relates a surface integral over a surface S to a line integral around the boundary curve C of (which is a curve on $\mathbb{R}^3$). Figure 21 shows an oriented surface S with unit normal vector n. The orientation of S induces the positive orientation of the boundary curve C shown in the figure. This means that if one walks in the positive direction around with her head pointing in the direction of n, then the surface S will always be on her left.

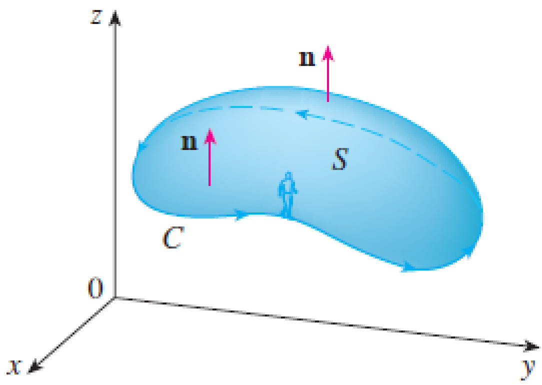

Figure 21: Strokes' theorem

Theorem 133 (Stokes' theorem) *Let* S *be an oriented smooth surface that is bounded by a simple, closed, boundary curve* C *with positive orientation. i.e.* $C = Bdy(S) = \overline{S} \backslash int(S)$. *Let* F *be a vector field whose components have continuous partial derivatives on an open region in* $\mathbb{R}^3$ *that contains* S. *Then*

$$\int_C F \cdot dr = \int \int_S curl\ F \cdot dS$$

Exercise 134 *Evaluate* $\int_C F \cdot dr$ *where* $F(x, y, z) = -y^2\hat{\mathbf{i}} + x\hat{\mathbf{j}} + z^2\hat{\mathbf{k}}$ *and* C *is the curve of intersection of the plane* $y + z = 2$ *and the cylinder* $x^2 + y^2 = 1$. *Also* C *has a counter clock wise orientation when viewed from above.*

From section 15.7 we see $\int_C F \cdot dr = \int_C F \cdot T \ \ dC$ and from definition 126 we see $\int\int_S curl \ \ F \cdot dS = \int\int_S (curl \ \ F) \cdot n \ \ dS$. So Stokes' Theorem says that the line integral of the tangential component of F around the boundary curve of S is equal to the surface integral over of the normal component of the curl of F. The positively oriented boundary curve C of the oriented surface S is often written as ∂S. In that case Stokes' Theorem can be expressed as

$$\int\int_S curl \ F \cdot dS = \int_{\partial S} F \cdot dr \cdots\cdots *$$

There is an analogy among Stokes Theorem, Green's Theorem, and the Fundamental Theorem of Calculus. As before, there is an integral involving derivatives on the left side of $*$ (recall from section 17.3 that $curl \ F$ is a sort of derivative of F) and the right side involves the values of F only on the boundary of S.

In fact, in the special case where the surface is flat and lies in the -plane with upward orientation, the unit normal is $\hat{\mathbf{k}}$, the surface integral becomes a double integral, and Stokes' Theorem becomes

$$\int_C F \cdot dr = \int\int_S curl \ F \cdot dS = \int\int_S (curl \ F) \cdot \hat{\mathbf{k}} \, dA$$

which is precisely the vector form of Green's theorem. So we can say that Green's theorem is a special case of Stokes' theorem.

Exercise 135 *Let C be a simple closed curve that lies in the plane $x+y+z=1$. Show that the line integral $\int_C z \ dx - 2x \ dy + 3y \ dz$ depends only on the area of the region bounded by C and not on the shape of C or its location in the plane.*

Exercise 136 *If S is a sphere and F satisfies the hypothesis of Stokes' theorem, show that $\int\int_S curl \ F \cdot dS = 0$*

20 The divergence theorem

Recall the definition of Simple region in $\mathbb{R}^n$(definition 66). We rewrite the definition of simple region in $\mathbb{R}^3$ for convenience. Let $D \subseteq \mathbb{R}^2$. f and g are two continuous functions from D to $\mathbb{R}$ such that $f(x,y) \leq g(x,y)$ for all $(x,y) \in D$. So a **solid simple region** E in $\mathbb{R}^3$ is defined by $E = \{(x,y,z) \in \mathbb{R}^3 \mid (x,y) \in D \wedge f(x,y) \leq z \leq g(x,y)\}$. So E is the region between surfaces $G(f)$ and $G(g)$ on $\mathbb{R}^3$. Where $G(f)$ is the graph of f.

In section 17.3 we have seen second vector form of Green's theorem

$$\oint_C F \cdot n \, dC = \int\int_R div(F(x,y,z)) \, dA$$

where R is the region in $\mathbb{R}^2$ enclosed by curve C. Now we just extend the same idea for vector field on $\mathbb{R}^3$ i.e. F is a function of x, y and z. So we can naturally guess an extension i.e.

$$\int_S F \cdot n \, dS = \int\int\int_E div(F(x,y,z)) \, dV$$

where E is a region enclosed by surface S on $\mathbb{R}^3$. Till now this is just an assumption but under appropriate hypothesis it may have true implications. One can observe its similarity to Green's Theorem and Stokes' Theorem. In those cases it relates **the integral of a sort of derivative of a function over a region** to the **integral of the original function over the boundary of the region**. Here also we have the same analogy.

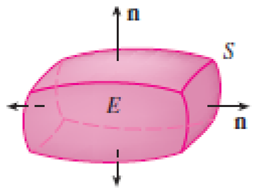

Figure 22: The divergence theorem

Theorem 137 (The divergence theorem) *Let E be a solid simple region in $\mathbb{R}^3$ such that its boundary is a is a surface S on $\mathbb{R}^3$ with a positive outward orientation. Let F be a vector field whose components have continuous partial derivatives on an open region in $\mathbb{R}^3$ that contains E. Then*

$$\int\int_S F \cdot dS = \int\int\int_E div\ F\ dV$$

Exercise 138 *Find the flux of the vector field $F(x,y,z) = z\hat{\boldsymbol{i}} + y\hat{\boldsymbol{j}} + x\hat{\boldsymbol{k}}$ over the unit sphere $x^2 + y^2 + z^2 = 1$.*

Exercise 139 *Evaluate $\int\int_S F \cdot dS$ where $F(x,y,z) = xy\hat{\boldsymbol{i}} + (y^2 + e^{xz^2})\hat{\boldsymbol{j}} + sin(xy)\hat{\boldsymbol{k}}$ and S is the surface of the region E bounded by the parabolic cylinder $z = 1 - x^2$ and the planes $z = 0$, $y = 0$ and $y + z = 2$.*

Exercise 140 *Consider the electric field $\mathfrak{E}$ on $\mathbb{R}^3$ defined by $\mathfrak{E}(v) = \dfrac{\varepsilon Q}{\| v \|^3} v$ where the electric charge Q is located at the origin and the position vector of another unit charge is $v = (x,y,z)$. Show that the electric flux of E through any closed surface S_2 that encloses the origin is*

$$\int\int_{S_2} E \cdot dS_2 = 4\pi\varepsilon Q$$

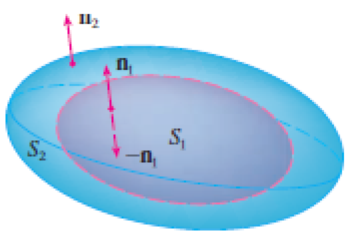

Figure 23: Exercisde - 140

Hint : We will definitely use theorem 137 but we have to find suitable E and S in a tricky way. See figure 23. Here S_1 is any smallest sphere completely inside S_2. We take the inner region of S_1 and S_2 as E. Then the appropriate orientation for it to be consistent with the statement of the divergence theorem should be n_2 for S_2 and $-n_1$ for S_1. Now we use the theorem.

Exercise 141 *A solid occupies a region E with surface S and is immersed in a liquid with constant density ρ. We set up a coordinate system so that the xy plane coincides with the surface of the liquid, and positive values of z are measured downward into the liquid. Then the pressure at depth z is , $p = \rho g z$ where g is the acceleration due to gravity. The total buoyant force on the solid due to the pressure distribution is given by the surface integral*

$$F = -\int\int_S -(pn)\, dS$$

*where n is the outer unit normal. Show that $F = -W\hat{\mathbf{k}}$, where W is the weight of the liquid displaced by the solid. (Note that F is directed upward because z is directed downward.) The result **Archimedes' Principle:** The buoyant force on an object equals the weight of the displaced liquid.*

Part IV
Problems

Readers are encouraged to solve the following problems to get a deeper insight into the subject.

Problem 1 (The wave equation)

1. Let c be a constant, tacitly understood to denote the speed of light. Show that if $u = F(x - ct) + G(x + ct)$ (where F and G are arbitrary $\mathcal{C}^2$ functions of one variable) then $c^2 \dfrac{\partial^2 u}{\partial x^2} = \dfrac{\partial^2 u}{\partial t^2}$

2. Now let $0 < v < c$ (both v and c are constant), and define new space and time variables in terms of old ones by a **Lorentz transformation**,

$$y = \gamma(x - vt), \;\; u = \gamma(t - (\frac{v}{c^2})x) \;\; where \;\; \gamma = (1 - \frac{v^2}{c^2})^{-\frac{1}{2}}.$$

   Show that

$$y^2 - (cu)^2 = x^2 - (ct)^2.$$

   Suppose that a quantity w, viewed as a function of x and t satisfies the wave equation, $c^2 \dfrac{\partial^2 w}{\partial x^2} = \dfrac{\partial^2 w}{\partial t^2}$. What corresponding equation does w satisfy viewed as a function of y and u?

Problem 2 (The monkey saddle)

The graph of the function

$$m(x, y) = 6xy^2 - 2x^3 - 3y^4$$

is called a **monkey saddle**. Find the three critical points of m and classify each as a maximum, minimum or saddle. (The max/min test will work on two. Study $m(x, 0)$ and $m(0, y)$ to classify the third.) Explain the name monkey saddle. (Computer graphing software may help).

Problem 3

Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} x + 2x^2 sin(\frac{1}{x}) + 7e^x & if \;\; x \neq 0 \\ 0 & if \;\; x = 0 \end{cases}$$

1. Show that f is differentiable at $x = 0$ and that $f'(0) \neq 0$. (Since this is a one-dimensional problem you may verify the old definition of derivative rather than the new one.)

2. Despite the result from the previous part of the problem, show that f is not locally invertible at $x = 0$. Why doesn't this contradict the Inverse Function Theorem?

Problem 4 (Diagonal matrix of differentiable operator)

A is open in $\mathbb{R}^n$. For what differentiable mappings $f : A \to \mathbb{R}^m$ is $Df(a)$ a diagonal matrix for all $a \in A$? (A diagonal matrix is a matrix whose $(i,j)^{th}$ entries for all $i \neq j$ are 0)

Problem 5

$f, g : \mathbb{R}^n \to \mathbb{R}^m$ are differentiable everywhere. Prove the following

1. If U is open in $\mathbb{R}^m$, $f^{-1}(U)$ is open in $\mathbb{R}^n$.
2. If F is closed in $\mathbb{R}^m$, $f^{-1}(F)$ is closed in $\mathbb{R}^n$.
3. If K is compact subset of $\mathbb{R}^n$, $f(K)$ is compact in $\mathbb{R}^m$.
4. $A \subseteq \mathbb{R}^n$ such that $f(a) = g(a)$ for all $a \in A$. Show that $f(a) = g(a)$ for all $a \in \bar{A}$ (Closure of A)
5. If T is connected subset of $\mathbb{R}^n$, $f(T)$ is connected in $\mathbb{R}^m$.

Problem 6

Suppose f is an integrable real valued function on $\mathbb{R}^n$. For each $\alpha > 0$ define $E_\alpha = \{x \mid |f(x)| > \alpha\}$. Show that

$$\int_{\mathbb{R}^n} \mid f(x) \mid dx = \int_0^\infty m^*(E_\alpha)\, d\alpha$$

Problem 7 (Measurable sets)

A subset E of $\mathbb{R}^n$ is measurable if for all $\varepsilon > 0$ there exists an open set O_ε with $E \subset O_\varepsilon$ such that $m^*(O_\varepsilon \setminus E) < \varepsilon$. Show that

1. Every open set in $\mathbb{R}^n$ is measurable.
2. Every closed set in $\mathbb{R}^n$ is measurable.
3. Countable union of measurable sets is measurable.
4. If $A \subseteq \mathbb{R}^n$ is measurable then $\mathbb{R}^n \setminus A$ is measurable.
5. Countable intersection of measurable sets is measurable.
6. Cantor's set is measurable.
7. There exists a set $A \subseteq \mathbb{R}^n$ such that A is not measurable.

Problem 8

Let $S = [0,1] \times [0,1] \subset \mathbb{R}^2$. Evaluate

$$I_0 = \int_S \frac{x^2 - y^2}{(x^2+y^2)^2}$$

Also calculate

$$I_1 = \int_0^1 \Big(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2} dx \Big) dy \quad and \quad I_2 = \int_0^1 \Big(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2} dy \Big) dx$$

Deduce that $I_1 \neq I_2$. Explain why Fubini's theorem is not being satisfied here.

Problem 9

Let F, G be vector fields on $\mathbb{R}^3$ and f be a scalar field on $\mathbb{R}^3$. Then $fF, F \cdot G, F \times G$ is defined by

$$fF(x,y,z) = f(x,y,z)F(x,y,z)$$
$$(F \cdot G)(x,y,z) = F(x,y,z) \cdot G(x,y,z)$$
$$(F \times G)(x,y,z) = F(x,y,z) \times G(x,y,z)$$

Now prove the following identities

1. $div(fF) = f\ div\ F + F \cdot \nabla f$
2. $curl(fG) = f\ curl\ G + (\nabla f) \times G$
3. $div(F \times G) = G \cdot curl\ F - F \cdot curl\ G$
4. $curl(curl\ G) = \nabla(div\ G) - \nabla^2 G$
5. $\nabla(F \cdot G) = (F \cdot \nabla)G + (G \cdot \nabla)F + F \times curl\ G + G \times curl\ F$

Problem 10 ($n-$surface)

U is open subset of $\mathbb{R}^{n+1}$. $f : U \to \mathbb{R}$ is smooth. Then $f^{-1}(c)$ is called the level set of f at height c.

$S \subseteq \mathbb{R}^{n+1}$ is called a $n-$ surface if

- S is a level set of some smooth function f from some open subset of $\mathbb{R}^{n+1}$ to $\mathbb{R}$ and
- $(\nabla f)(p) \neq 0$ for all $p \in S$.

$1-$surfaces are called curves. $2-$surfaces are called surfaces.

See the examples of some level set below (Figure 24) and answer the following.

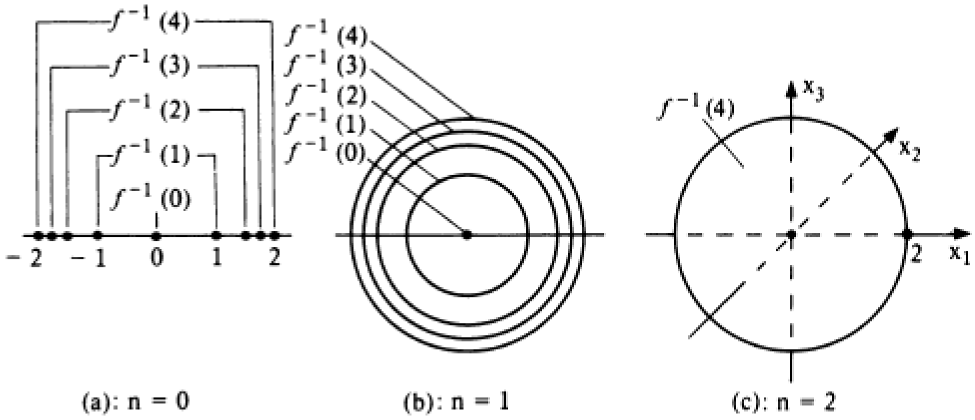

Figure 24: Level sets for the functions $f(x_1, x_2, \cdots, x_{n+1} = x_1^2 + x_2^2 + \cdots + x_{n+1}^2$

1. Give an example of a parametrized curve on $\mathbb{R}^2$ (i.e. image of some parametrization) which is not a curve (i.e. $1-$surface)

2. Give an example of a curve (i.e. $1-$surface)which is not a parametrized curve on $\mathbb{R}^2$(i.e. image of some parametrization)

3. Give an example of a parametrized surface on $\mathbb{R}^3$ (i.e. image of some parametrization) which is not a surface (i.e. $2-$surface)

4. Give an example of a surface (i.e. $2-$surface)which is not a parametrized surface on $\mathbb{R}^3$(i.e. image of some parametrization)