Analysis

Course 221

David Simms

School of Mathematics

Trinity College

Dublin

Typesetting

Eoin Curran.

Contents

Chapter 1

Lebesgue Measure

1.1 Algebra of Subsets

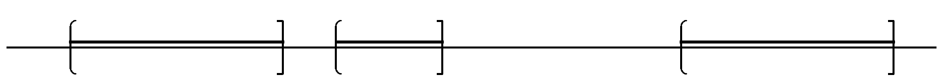
Definition Let X be a set. A collection $\mathcal{A}$ of subsets of X is called an *algebra* of subsets of X if

1. $\emptyset \in \mathcal{A}$
2. $E \in \mathcal{A} \implies E' \in \mathcal{A}$. $\qquad E' = \{x \in X : x \notin E\}$
3. $E_1, \ldots, E_k \in \mathcal{A} \implies E_1 \cup E_2 \cup \cdots \cup E_k \in \mathcal{A}$. i.e., $\mathcal{A}$ is closed under complements and under finite unions. Hence $\mathcal{A}$ is closed under finite intersection. (because $E_1 \cap \cdots \cap E_k = ({E_1}' \cup \cdots \cup {E_k}')'$)

Example Let $\mathcal{I}$ be the collection of all finite unions of intervals in $\mathbf{R}$ of the form:

1. $(a, b] = \{x \in \mathbf{R} : a < x \leq b\}$
2. $(-\infty, b] = \{x \in \mathbf{R} : x \leq b\}$
3. $(a, \infty) = \{x \in \mathbf{R} : a < x\}$
4. $(-\infty, \infty) = \mathbf{R}$

$\mathcal{I}$ is an algebra of subsets of $\mathbf{R}$, called the *Interval Algebra*.



We want to assign a 'length' to each element of the interval algebra $\mathcal{I}$, so we want to allow '∞' as a length. We adjoin to the real no's the 2 symbols ∞ and $-\infty$ to get the *Extended Real Line*:

Definition The *Extended Real Line*:

$$\mathbf{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$$

We extend ordering, addition, multiplication to $[-\infty, \infty]$ by:

1. $-\infty < x < \infty \ \forall x \in \mathbf{R}$
2. Addition
   (a) $\infty + \infty = x + \infty = \infty + x = \infty$
   (b) $(-\infty) + (-\infty) = x + (-\infty) = (-\infty) + x = -\infty$

   (Don't define $\infty + (-\infty)$ or $(-\infty) + \infty$.)
3. Multiplication
   (a) $\infty\infty = \infty = (-\infty)(-\infty)$
   (b) $\infty(-\infty) = (-\infty)\infty = -\infty$
   (c) $x\infty = \infty x = \left\{ \begin{array}{ll} \infty & x > 0 \\ 0 & x = 0 \\ -\infty & x < 0 \end{array} \right\}$
   (d) $x(-\infty) = (-\infty)x = \left\{ \begin{array}{ll} -\infty & x > 0 \\ 0 & x = 0 \\ \infty & x < 0 \end{array} \right\}$

Definition If $A \subset [-\infty, \infty]$ then

1. $supA = \infty$ if $\infty \in A$ or if A has no upper bound in $\mathbf{R}$.
2. $inf A = -\infty$ if $-\infty \in A$ or if A has no lower bound in $\mathbf{R}$.

Definition Let $\mathcal{A}$ be an algebra of subsets of X. A function $m : \mathcal{A} \longrightarrow [0, \infty]$ is called a *measure* on $\mathcal{A}$ if:

1. $m(\emptyset) = 0$
2. if $E = \bigcup_{j=1}^{\infty} E_j$ is countably disjoint with $E_j \in \mathcal{A} \ \forall j$ then $m(E) = \sum_{j=1}^{\infty} m(E_j)$. ($m$ is *Countably Additive*.)

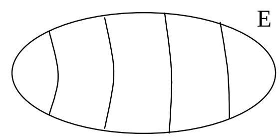


Figure 1.1: E is a disjoint union of sets

Note:

1. *Countable* means that $E_1, E_2, \ldots$ is either a finite sequence or can be labelled by $(1, 2, 3, \ldots)$.
2. *Disjoint* means that $E_i \cap E_j = \emptyset \;\; \forall i \neq j$
3. $m(E) = \sum_{j=1}^{\infty} m(E_j)$ means either:
   (a) $\sum_{j=1}^{\infty} m(E_j)$ is a convergent series of finite no's with the sum of the series as $m(E)$
   (b) $\sum_{j=1}^{\infty} m(E_j)$ is a divergent series with $m(E) = \infty$
   (c) $m(E_j) = \infty$ for some j, and $m(E) = \infty$

1.2 The Interval Algebra

Definition For each $E \in \mathcal{I}$, the interval algebra (See Section 1.1), write

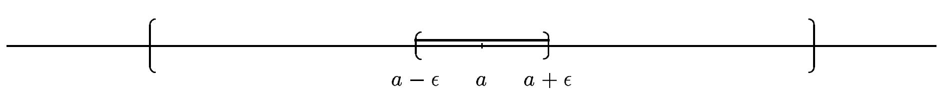
$$E = E_1 \cup \cdots \cup E_k$$

(disjoint) where each E_i is of the form $(a, b], (-\infty, b], (a, \infty), or (-\infty, \infty)$. The *length* of E is $m(E) = m(E_1) + m(E_2) + \cdots + m(E_k)$ where $m(a, b] = b - a$.

$$m(-\infty, b] = m(a, \infty) = m(-\infty, \infty) = \infty$$

Definition a set $V \subset \mathbf{R}$ is called an *Open Subset* of $\mathbf{R}$ if:

$$\text{for each } a \in V \; \exists \epsilon > 0 \;\; \text{s.t.} (a - \epsilon, a + \epsilon) \subset V$$



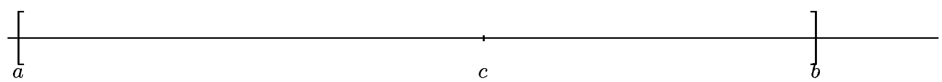
i.e., every point is an *interior point*: it has no end points.

Theorem 1.2.1. (Heine-Borel-Lebesgue) *The closed interval is compact.*

Let $\{V_i\}_{i\in I}$ *be a family of open sets in* $\mathbb{R}$ *which cover the closed interval* $[a,b]$. *Then* $\exists$ *a finite number of them:* $V_{i_1},\dots,V_{i_k}$ *(say) which cover* $[a,b]$ *(i.e.* $[a,b]\subset V_{i_1}\cup\cdots\cup V_{i_k}$*.)*

Proof. Put

$$K=\{x\in[a,b] \text{ s.t. } \exists(\text{ finite set } \{i_1,\dots,i_r\}\subset I \ \text{ s.t. } [a,x]\subset V_{i_1}\cup\cdots\cup V_{i_r})\}$$



$$a\in K \implies K\neq\emptyset$$

Let

$$c=\sup K$$

Then

$$a\le c\le b \quad c\in V_j \text{ (say)}$$

V_j is open, therefore

$$\exists\,\epsilon>0 \text{ s.t. } (c-\epsilon,c+\epsilon)\subset V_j$$

and

$$\exists\, k\in K \text{ s.t. } c-\epsilon<k\le c,$$

$$[a,k]\subset V_{i_1}\cup\cdots\cup V_{i_r} \text{ (say)}$$

Then

$$[a,min(c+\frac{\epsilon}{2},b)]\subset(V_{i_1}\cup\cdots\cup V_{i_r}\cup V_j)$$

therefore

$$min(c+\frac{\epsilon}{2},b)\in K$$

But $c+\frac{\epsilon}{2}\notin K$ (as $c=\sup K$) $\implies b\in K$ as required. □

1.3 The length measure

Theorem 1.3.1. *The length function*

$$m : \mathcal{I} \longrightarrow [0, \infty]$$

is a measure on the interval algebra $\mathcal{I}$

Proof. We have to check that m is countably additive, which reduces to showing that if

$$(a, b] = \bigcup_{j=1}^{\infty} (a_j, b_j]$$

is a countable disjoint union then

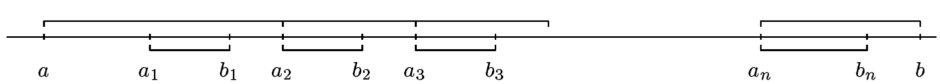
$$\sum_{j=1}^{\infty} (b_j - a_j)$$

is a convergent series with sum $b - a$.

Take the first n intervals and relabel them:

$$(a_1, b_1], \dots, (a_n, b_n]$$

so that:



then

$$\begin{aligned} \textstyle\sum_{j=1}^{n} (b_j - a_j) &= b_1 - a_1 \quad + b_2 - a_2 \quad + b_3 - a_3 \quad + \cdots \quad + b_n - a_n \\ &\le a_2 - a \quad + a_3 - a_2 \quad + a_4 - a_3 \quad + \cdots \quad + b - a_n \\ &= b - a \end{aligned}$$

therefore

$$\sum_{j=1}^{\infty} (b_j - a_j)$$

converges and has sum $\le b - a$

Let $\epsilon > 0$ and put $\epsilon_0 = \frac{\epsilon}{2}$, $\epsilon_1 = \frac{\epsilon}{4}, \dots, \epsilon_j = \frac{\epsilon}{2^{j+1}}, \dots$

so $\epsilon_j > 0$ and $\sum_{j=1}^{\infty} \epsilon_j = \epsilon$

Then

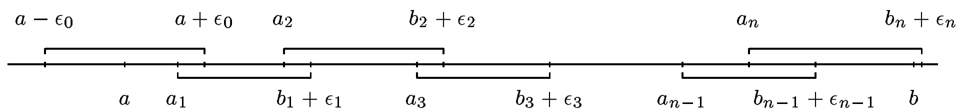
$$(a - \epsilon_0, a + \epsilon_0), (a_1, b_1 + \epsilon_1), (a_2, b_2 + \epsilon_2), \dots$$

is a family of open sets which cover $[a, b]$

By the compactness of $[a, b]$ $\exists$ a finite number of these open sets which cover $[a, b]$, and which by renumbering and discarding some intervals if necessary we can take as:

$$(a - \epsilon_0, a + \epsilon_0), (a_1, b_1 + \epsilon_1), \ldots, (a_n, b_n + \epsilon_n)$$

so that:



$$\begin{array}{lllllll} b - a & = a_1 - a & + a_2 - a_1 & + \cdots & + a_n - a_{n-1} & + b - a_n \\ & < \epsilon_0 & + (b_1 + \epsilon_1 - a_1) & + \cdots & + (b_{n-1} + \epsilon_{n-1} - a_{n-1}) & + (b_n + \epsilon_n - a_n) \\ & < \epsilon & + \sum_{j=1}^{\infty} (b_j - a_j) & & & \end{array}$$

is true $\forall\ \epsilon > 0$. Therefore

$$b - a \leq \sum_{j=1}^{\infty} (b_j - a_j)$$

Therefore

$$\sum_{j=1}^{\infty} (b_j - a_j) = b - a$$

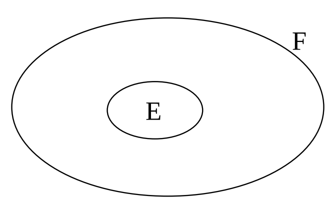
as required. $\square$

Let $m :\longrightarrow [0, \infty]$ be a measure m on an algebra $\mathcal{A}$ of subsets of X.

$$E \subset F \implies \left\{ \begin{array}{l} Then m(E) \leq m(F) \\ and \\ m(F \cap E') = m(F) - m(E) \text{ if } m(F) \neq \infty \end{array} \right\}$$

Then

$$F = E \cup (F \cap E')$$



is a disjoint union, and

$$m(F) = m(E) + m(F \cap E')$$

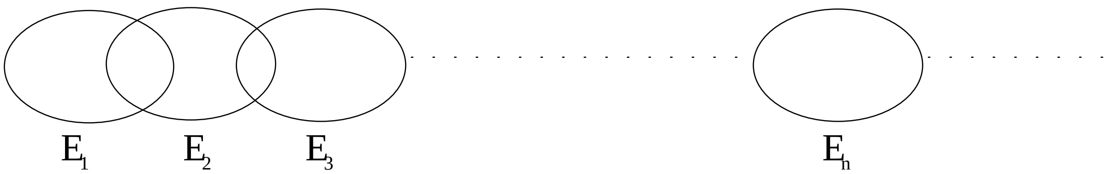
Theorem 1.3.2. (m is subadditive on countable unions)

$$m(\bigcup_{j=1}^{\infty}) \leq \sum_{i=1}^{\infty} m(E_i)$$

Proof. We proceed as follows:

Let

$$F_n = E_1 \cup E_2 \cup \cdots \cup E_n$$



$$\bigcup_{j=1}^{\infty} E_j = E_1 \cup (E_2 \cap F_1{}') \cup (E_3 \cap F_2{}') \cup \cdots$$

$$\begin{array}{lllll} m(\bigcup_{j=1}^{\infty} E_j) & = m(E_1) & +m(E_2 \cap F_1{}') & +m(E_3 \cap F_2{}') & +\cdots \\ & \leq m(E_1) & +m(E_2) & +m(E_3) & +\cdots \end{array}$$

hence the result:

$$m(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} m(E_j)$$

m is *subadditive on countable unions.* □

1.4 The σ-algebra

Our aim is to extend the notion of length to a much wider class of subsets of $\mathbb{R}$. In particular to sets obtainable from $\mathcal{I}$ by a sequence of taking countable unions and taking complements.

Definition an algebra $\mathcal{A}$ of subsets of X is called a *σ-algebra* if for each sequence

$$E_1, E_2, E_3, \ldots$$

of elements of $\mathcal{A}$, their union

$$E_1 \cup E_2 \cup E_3 \cup \cdots$$

is also an element of $\mathcal{A}$.

So $\mathcal{A}$ is closed under countable unions, and hence also under countable intersections.

1.5 The outer measure

Definition We define the *outer measure* $\hat{m}$ associated with m to be the function:

$$\hat{m} : \{\text{all subsets of X}\} \longrightarrow [0, \infty]$$

given by:

$$\hat{m}(E) = \inf \sum_{j=1}^{\infty} m(E_j)$$

where the inf is taken over all sequences E_j of elements of $\mathcal{A}$ such that:

$$E \subset \bigcup_{j=1}^{\infty} E_j$$

Theorem 1.5.1. *$\hat{m}(E) = m(E)\ \ \forall E \in \mathcal{A}$ (i.e. $\hat{m}$ agrees with m on $\mathcal{A}$)*

Proof.

1. if

$$E \in \mathcal{A}$$

and

$$E \subset \bigcup_{j=1}^{\infty} E_j \text{ with } E_j \in \mathcal{A}$$

then:

$$m(E) \leq \sum_{j=1}^{\infty} m(E_j)$$

therefore

$$m(E) \leq \hat{m}(E)$$

2. if

$$E \in \mathcal{A}$$

then

$$E, \emptyset, \emptyset, \ldots$$

is a sequence of elements of $\mathcal{A}$ whose union contains E. Therefore:

$$\hat{m}(E) \leq m(E) + 0 + 0 + \cdots$$

therefore

$$\hat{m}(E) \leq m(E)$$

□

Theorem 1.5.2. $E \subset F \implies \hat{m}(E) \leq \hat{m}(F)$

Proof. if $E \subset F$ then each sequence in $\mathcal{A}$ which covers F will also cover E. Therefore:

$$\hat{m}(E) \leq \hat{m}(F)$$

□

Theorem 1.5.3. $\hat{m}$ is subadditive on countable unions

$$\hat{m}(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} \hat{m}(E_j)$$

for all sequences $E_1, E_2, \ldots$ of subsets of X.

Proof. Let $\epsilon > 0$. Choose:

$$\epsilon_0, \epsilon_1, \epsilon_2, \ldots \ > 0$$

such that

$$\sum_{j=0}^{\infty} \epsilon_j = \epsilon$$

choose $B_{ij} \in \mathcal{A}$ s.t.

$$E_i \subset B_{i1} \cup B_{i2} \cup \cdots \cup B_{ij} \cup \cdots$$

and s.t.

$$\sum_{j=1}^{\infty} m(B_{ij}) \leq \hat{m}(E_i) + \epsilon_i$$

Then

$$\bigcup_{i=1}^{\infty} E_i \subset \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} B_{ij}$$

and

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} m(B_{ij}) \leq \sum_{i=1}^{\infty} \hat{m}(E_i) + \epsilon$$

therefore

$$\hat{m}(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} \hat{m}(E_j)$$

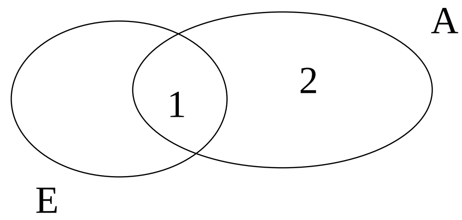
as required. □

Definition We call a subset $E \subset X$ *measurable w.r.t m* if:

$$\hat{m}(A) = \hat{m}(A \cap E) + \hat{m}(A \cap E')$$

for all $A \subset X$

(E splits every set A into two pieces whose outer measures add up.)



1.6 Extension of measure to σ-algebra, using outer measure

We can now prove the central:

Theorem 1.6.1. *Let m be a measure on an algebra $\mathcal{A}$ of subsets of X, $\hat{m}$ the associated outer measure, and M the collection of all subsets of X which are measurable with respect to m.*

Then M is a σ-algebra containing $\mathcal{A}$ and $\hat{m}$ is a (countably additive) measure on M.

Corollary 1.6.2. *the measure m on the algebra $\mathcal{A}$ can be extended to a measure (also denoted by m) on the σ-algebra M by defining:*

$$m(E) = \hat{m}(E) \ \ \forall E \in M$$

in particular:

Corollary 1.6.3. *the length measure m on the interval algebra $\mathcal{I}$ can be extended to a measure (also denoted by m) on the σ-algebra M of measurable sets w.r.t. m. We call the elements of M the* Lebesgue Measureable *sets, and the extended measure the (one-dimensional)* Lebesgue Measure.

$$\mathcal{I} \quad \subset \quad \{\textit{Lebesgue Measurable Sets}\} \quad \subset \quad \textit{all subsets of } \mathbb{R}$$

length measure m	*Lebesgue measure* (countably additive)	*outer measure $\hat{m}$* (not countably additive $\Longrightarrow$ not a measure)
$\searrow$	$\downarrow$	$\swarrow$
	$[0, \infty]$	

Proof. 1. $\mathcal{A} \subset M$

Let $E \in \mathcal{A}$, let $A \subset X$

Need to show:

$$\hat{m}(A) = \hat{m}(A \cap E) + \hat{m}(A \cap E')$$

For $\epsilon > 0$. Let

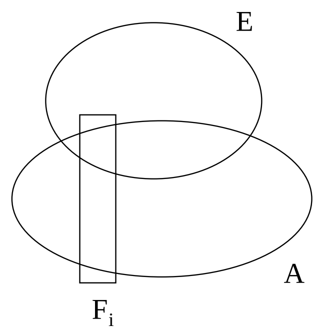
$$F_1, F_2, \ldots$$

be a sequence in $\mathcal{A}$ s.t.

$$A \subset \bigcup_{j=1}^{\infty} F_j$$

and s.t.

$$\sum_{j=1}^{\infty} m(F_j) \leq \hat{m}(A) + \epsilon$$



Then

$$\begin{aligned}
\hat{m}(A) &\leq \hat{m}(A \cap E) + \hat{m}(A \cap E') && \text{since } \hat{m} \text{ is subadditive} \\
&\leq \hat{m}\left(\textstyle\bigcup_{j=1}^{\infty} F_j \cap E\right) + \hat{m}\left(\textstyle\bigcup_{j=1}^{\infty} F_j \cap E'\right) \\
&\leq \textstyle\sum_{j=1}^{\infty} \hat{m}(F_j \cap E) + \textstyle\sum_{j=1}^{\infty} \hat{m}(F_j \cap E') \\
&= \textstyle\sum_{j=1}^{\infty} m(F_j \cap E) + \textstyle\sum_{j=1}^{\infty} m(F_j \cap E') && \text{since } F_j \cap E, F_j \cap E' \in \mathcal{A} \text{ and } \hat{m} = m \text{ on } \mathcal{A} \\
&= \textstyle\sum_{j=1}^{\infty} m(F_j) && \text{since } m \text{ is additive} \\
&\leq \hat{m}(A) + \epsilon
\end{aligned}$$

therefore,

$$\hat{m}(A) \leq \hat{m}(A \cap E) + \hat{m}(A \cap E') \leq \hat{m}(A) + \epsilon$$

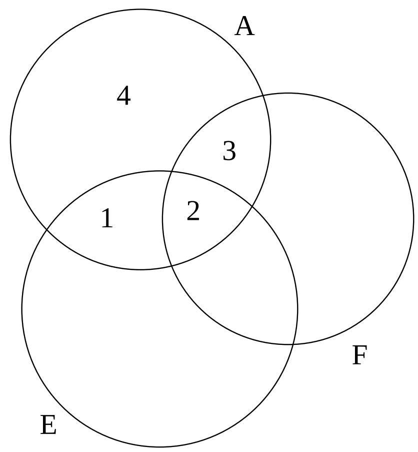
$\forall \epsilon > 0$, and therefore

$$\hat{m}(A) = \hat{m}(A \cap E) + \hat{m}(A \cap E')$$

as required.

2. M is an algebra

   let $E, F \in M$; Let $A \subset X$



Then

$$\begin{aligned}
\hat{m}(A) &= \hat{m}(1+2) && +\hat{m}(3+4) && && \text{since } E \in M \\
&= \hat{m}(1+2) && +\hat{m}(3) && +\hat{m}(4) && \text{since } F \in M \\
&= \hat{m}(1+2+3) && +\hat{m}(4) && && \text{since } E \in M
\end{aligned}$$

therefore

$$E \cup F \in M$$

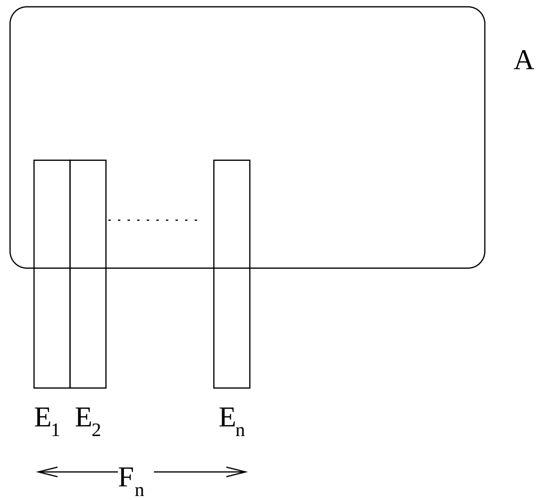
and M closed under finite unions.

Also, M closed under complements by symmetry in the definition.

3. M is a σ-algebra

   Let $E = \bigcup_{i=1}^{\infty} E_i$ be a countable disjoint union with $E_i \in M$. We need to show that $E \in M$. Put

$$F_n = E_1 \cup E_2 \cup \cdots \cup E_n$$



$F_n \in M$ since M is an algebra. Let $A \subset X$. Then

$$\begin{aligned} \hat{m}(A) &= \hat{m}(A \cap F_n) + \hat{m}(A \cap F_n{}') && \text{since } F_n \in M \\ &= \textstyle\sum_{k=1}^{n} \hat{m}(A \cap E_k) + \hat{m}(A \cap F_n{}') && \text{since } E_1, \ldots, E_n \in M, \text{ and are disjoint} \\ &\geq \textstyle\sum_{k=1}^{n} \hat{m}(A \cap E_k) + \hat{m}(A \cap E') && \text{since } E' \subset F_n{}' \end{aligned}$$

is true $\forall n$. Therefore:

$$\left.\begin{aligned} \hat{m}(A) &\geq \textstyle\sum_{k=1}^{\infty} \hat{m}(A \cap E_k) + \hat{m}(A \cap E') \\ &\geq \hat{m}(\textstyle\bigcup_{k=1}^{\infty} A \cap E_k) + \hat{m}(A \cap E') \quad \text{since } \hat{m} \text{ is countably subadditive} \end{aligned}\right\} (*)$$

$$\begin{aligned} &= \hat{m}(A \cap E) + \hat{m}(A \cap E') \\ &\geq \hat{m}(A) \qquad \text{since } \hat{m} \text{ is subadditive} \end{aligned}$$

All the above are equalities:

$$\hat{m}(A) = \hat{m}(A \cap E) + \hat{m}(A \cap E')$$

therefore

$$E \in M$$

and M is closed under countable disjoint unions. But any countable union:

$$E_1 \cup E_2 \cup E_3 \cup \cdots$$

can be written as a countable disjoint union

$$E_1 \cup (E_2 \cap F_1{}') \cup (E_3 \cap F_2{}') \cup \cdots$$

where

$$F_n = E_1 \cup \cdots \cup E_n$$

therefore M is closed under countable unions as required.

4. $\hat{m}$ is countably additive on M

   Put $A = E$ in (*) to get

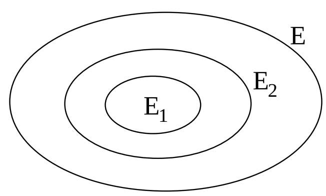
$$\sum_{k=1}^{\infty} \hat{m}(E_k) = \hat{m}\left(\bigcup_{k=1}^{\infty} E_k\right)$$

   as required.

□

1.7 Increasing Unions, Decreasing Intersections

Definition We use the notation $E_j \uparrow E$ to denote that

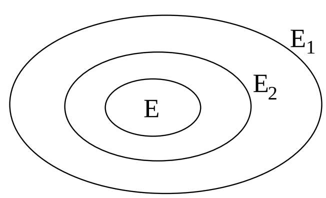


$$E_1 \subset E_2 \subset \cdots \subset E_j \subset \cdots$$

is an increasing sequence of sets such that

$$\bigcup_{j=1}^{\infty} E_j = E$$

Definition We use the notation $E_j \downarrow E$ to denote that



$$E_1 \supset E_2 \supset \cdots \supset E_j \supset \cdots$$

is a decreasing sequence of sets such that

$$\bigcap_{j=1}^{\infty} E_j = E$$

Theorem 1.7.1. $E_j \uparrow E \implies \lim_{j\to\infty} m(E_j) = m(E)$

Proof. 1. if $m(E_j) = \infty$ for some j then the result holds

2. if $m(E_j)$ is finite $\forall j$ then

$$E = E_1 \cup (E_2 \cap E_1{}') \cup (E_3 \cap E_2{}') \cup \cdots$$

is a countable disjoint union and

$$\begin{aligned} m(E) &= m(E_1) + [m(E_2) - m(E_1)] + [m(E_3) - m(E_2)] + \cdots \\ &= \lim_{n\to\infty}\Big\{ m(E_1) + [m(E_2) - m(E_1)] + \cdots + [m(E_n) - m(E_{n-1})] \Big\} \\ &= \lim_{n\to\infty} m(E_n) \end{aligned}$$

as required.

□

Theorem 1.7.2.

$$\left.\begin{array}{l} E_j \downarrow E \\ m(E_1) \neq \infty \end{array}\right\} \implies \lim_{j\to\infty} m(E_j) = m(E)$$

Proof.

$$(E_1 \cap E_j{}') \uparrow (E_1 \cap E')$$

therefore:

$$\lim_{j\to\infty} m(E_1 \cap E_j{}') = m(E_1 \cap E')$$

and

$$\lim_{j\to\infty} [m(E_1) - m(E_j)] = m(E_1) - m(E)$$

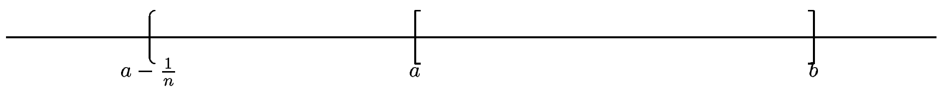
so

$$\lim_{j\to\infty} m(E_j) = m(E)$$

as required. □

Example

$$(a - \frac{1}{n}, b] \downarrow [a, b]$$



for Lebesgue measure m:

$$\begin{aligned} m[a,b] &= \lim_{n\to\infty} m(a - \tfrac{1}{n}, b] \\ &= \lim_{n\to\infty} (b - a + \tfrac{1}{n}) \\ &= b - a \end{aligned}$$

as expected.

1.8 Properties of Lebesgue Measure

We now show that the Lebesgue measure is the only way of extending the 'length' measure on the interval algebra $\mathcal{I}$ to a measure on the Lebesgue measurable sets.

We need:

Definition a measure m on an algebra $\mathcal{A}$ of subsets of X is called *σ-finite* if $\exists$ a sequence X_i in $\mathcal{A}$ such that

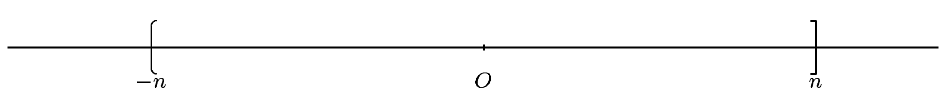
$$X_i \uparrow X$$

and $m(X_i)$ is finite for all i.

Example the Lebesgue measure m is σ-finite because:

$$(-n, n] \uparrow \mathbb{R}$$

and $m(-n, n] = 2n$ is finite



Theorem 1.8.1. (Uniqueness of Extension) *Let m be a σ-finite measure on an algebra $\mathcal{A}$ of subsets of X.*

Let M be the collection of measurable sets w.r.t. m.

Let l be any measure on M which agrees with m on $\mathcal{A}$.

Then $l(E) = \hat{m}(E)\ \forall E \in M$

Proof. Let $E \in M$. Then for each sequence $\{A_i\}$, $A_i \in \mathcal{A}$ covering E:

$$E \subset \bigcup_{j=1}^{\infty} A_i$$

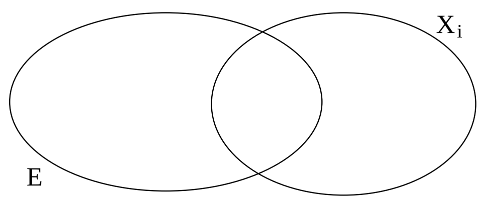
we have

$$l(E) \quad \leq \textstyle\sum_{i=1}^{\infty} l(A_i) \quad = \sum_{i=1}^{\infty} m(A_i) \quad \text{since } l = m \text{ on } \mathcal{A}$$

Therefore,

$$(*) \quad l(E) \quad \leq \hat{m}(E) \qquad \text{by definition of the outer measure } \hat{m}$$

Now let $X_i \uparrow X$ and $m(X_i)$ finite, $X_i \in \mathcal{A}$. Consider



$$l(X_i \cap E) + l(X_i \cap E') = l(X_i) = m(X_i) = \hat{m}(X_i \cap E) + \hat{m}(X_i \cap E)$$

By (*) it follows that

$$l(X_i \cap E) = \hat{m}(X_i \cap E)$$

Take $\lim_{i\to\infty}$ to get $l(E) = \hat{m}$ as required. □

As a consequence we have:

Lemma 1.8.2. *Let m be the Lebesgue measure on $\mathbb{R}$. Then for each measurable set E and each $c \in \mathbb{R}$ we have:*

1. $m(E + c) = m(E)$. $\quad m$ *is* translation invariant
2. $m(cE) = |c|m(E)$

Proof. Let M be the collection of measurable sets

1. define a measure m_c on M by

$$m_c(E) = m(E + c)$$

then

$$m_c(a, b] = m(a + c, b + c] = (b + c) - (a + c) = b - a = m(a, b]$$

therefore m_c agrees with m on the interval algebra $\mathcal{I}$. Therefore m_c agrees with m on M.

2. define a measure m_c on M by

$$m_c(E) = m(cE)$$

Then

$$\begin{aligned} m_c(a, b] &= m(c(a, b]) \\ &= \left\{ \begin{array}{ll} m(ca, cb] & c > 0 \\ m\{0\} & c = 0 \\ m[cb, ca) & c < 0 \end{array} \right\} \\ &= \left\{ \begin{array}{l} cb - ca \\ 0 \\ ca - cb \end{array} \right\} \\ &= |c|(b - a) \\ &= |c|m(a, b] \end{aligned}$$

Therefore m_c agrees with $|c|m$ on the interval algebra. Therefore m_c agrees with $|c|m$ on M.

□

1.9 Borel Sets

Definition if $\mathcal{V}$ is any collection of subsets of X, we denote by $G(\mathcal{V})$ the intersection of all the σ-algebras of subsets of X which contain $\mathcal{V}$. We have:

1. $G(\mathcal{V})$ is a σ-algebra containing $\mathcal{V}$
2. if $\mathcal{W}$ is any σ-algebra which contains $\mathcal{V}$ then

$$\mathcal{V} \subset G(\mathcal{V}) \subset \mathcal{W}$$

Thus $G(\mathcal{V})$ is the smallest σ-algebra of subsets of X which contains $\mathcal{V}$. $G(\mathcal{V})$ is called the σ-algebra *generated by* $\mathcal{V}$.

Definition The σ-algebra generated by the open sets of $\mathbb{R}$ is called the algebra of *Borel Sets* of $\mathbb{R}$.

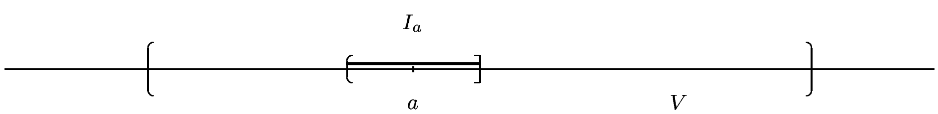
Theorem 1.9.1. *The σ-algebra generated by the interval algebra $\mathcal{I}$ is the algebra of Borel sets of* $\mathbb{R}$.

Proof. Let $\mathcal{V}$ be the collection of open sets of $\mathbb{R}$

1. $(a, b] = \bigcap_{n=1}^{\infty}(a, b + \frac{1}{n})$ is the intersection of a countable family of open sets. Therefore $(a, b] \in G(\mathcal{V})$.

   Therefore
   $$\mathcal{I} \subset G(\mathcal{V})$$
   and
   $$G(\mathcal{I}) \subset G(\mathcal{V})$$

2. let V be an open set in $\mathbb{R}$. For each $a \in V$ choose an interval $I_a \in \mathcal{I}$ with rational endpoints s.t. $a \in I_a \subset V$.



   $$V = \bigcup_{a \in V} I_a$$

   so V is the union of a countable family of elements of $\mathcal{I}$. Therefore

   $$V \in G(\mathcal{I})$$

   implies
   $$\mathcal{V} \subset G(\mathcal{I}) \implies G(\mathcal{V}) \subset G(\mathcal{I})$$

So, combining these two results,

$$G(\mathcal{I}) = G(\mathcal{V}) = \text{algebra of Borel sets}$$

□

We have:

$$\mathcal{I} \subset \text{Borel Sets in } \mathbb{R} \subset \text{Lebesgue measurable Sets in } \mathbb{R}$$

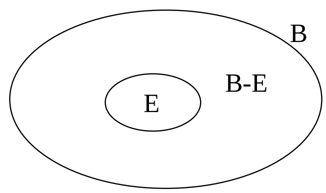
The following theorem shows that any Lebesgue measurable set in $\mathbb{R}$ can be obtained from a Borel set by removing a set of measure zero.

Theorem 1.9.2. *Let E be a Lebesgue measurable subset of $\mathbb{R}$. Then there is a Borel set B containing E such that*

$$B - E = B \cap E'$$

has measure zero, and hence

$$m(B) = m(E)$$



Proof. 1. Suppose $m(E)$ is finite. Let k be an integer > 0. Choose a sequence

$$I_1, I_2, I_3, \ldots$$

in $\mathcal{I}$ such that

$$E \subset \bigcup_{i=1}^{\infty} I_i$$

and s.t.

$$m(E) \le \sum_{i=1}^{\infty} m(I_i) \le m(E) + \frac{1}{k}$$

Put $B_k = \bigcup_{i=1}^{\infty} I_i$ then $E \subset B_k$, B_k is Borel, and

$$m(E) \le m(B_k) \le m(E) + \frac{1}{k}$$

Put $B = \bigcap_{k=1}^{\infty} B_k$. Then $E \subset B$, B is Borel, and

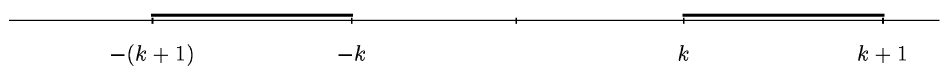
$$m(E) \le m(B) \le m(E) + \frac{1}{k}$$

for all k. So

$$m(E) = m(B) \implies m(B - E) = 0$$

since $m(E)$ is finite.

2. Suppose $m(E) = \infty$.



Put

$$E_k = \{x \in E : k \le |x| < k+1\}$$

Then

$$E = \bigcup_{k=0}^{\infty} E_k$$

is a countable disjoint union and $m(E_k)$ is finite.

For each integer k choose a Borel set B_k s.t.

$$E_k \subset B_k$$

and

$$m(B_k - E_k) = 0$$

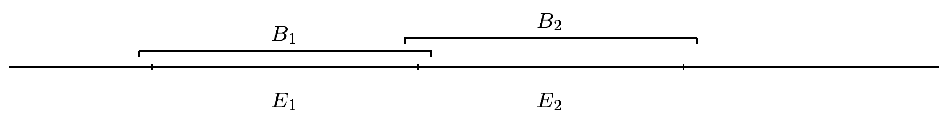
Put $B = \bigcup_{k=1}^{\infty} B_k$. Then $E \subset B$, B is Borel, and

$$B - E \subset \bigcup_{k=1}^{\infty} (B_k - E_k)$$

therefore

$$m(B - E) \le \sum_{k=1}^{\infty} m(B_k - E_k) = \sum_{k=1}^{\infty} 0 = 0$$

as required.



□

Chapter 2

Integration

2.1 Measure Space, Measurable sets

Definition We fix a set X, a σ-algebra M of subsets of X, and a measure m on M. The triple (X, M, m) is then called a *measure space*, the elements of M are called the *measurable sets* of the measure space.

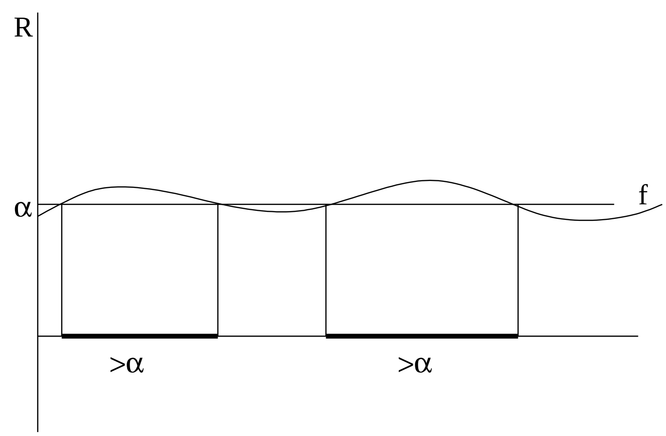
Definition We call a function

$$f : X \longrightarrow [-\infty, \infty]$$

measurable if

$$f^{-1}(\alpha, \infty] = \{x \in X : f(x) > \alpha\}$$

is a measurable $\forall \alpha \in \mathbb{R}$.



Example The collection

$$\{E \subset \mathbb{R} : f^{-1}(E) \text{ is measurable}\}$$

is a σ-algebra containing the sets

$$\{(\alpha, \beta] : \alpha, \beta \in \mathbb{R}\}$$

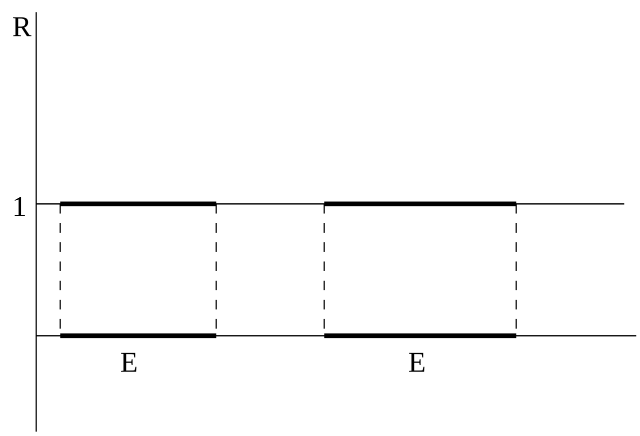
therefore contains the Borel sets. Therefore f^{-1} is measurable for each Borel set $B \subset \mathbb{R}$.

2.2 Characteristic Function

Definition If $E \subset X$, we denote by χ_E the function on X:

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

χ_E is called the *characteristic function* of E.



Definition A real valued function

$$\phi : X \longrightarrow \mathbb{R}$$

is called *simple* if it takes only a finite number of distinct values

$$a_1, a_2, \ldots, a_n$$

(say). Each simple function ϕ can be written in a unique way as:

$$\phi = a_1\chi_{E_1} + \cdots + a_n\chi_{E_n}$$

where $a_1, \ldots, a_n$ are distinct and

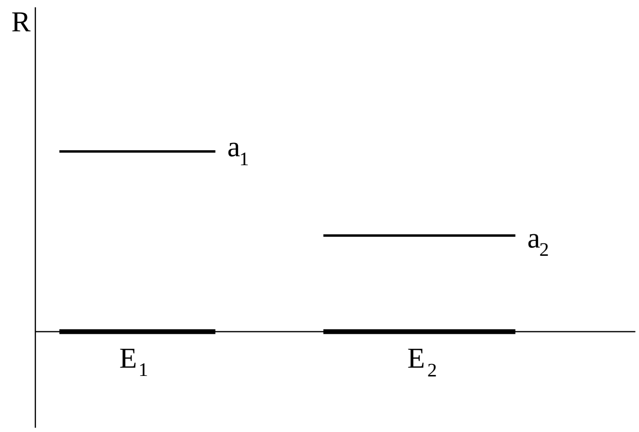
$$X = E_1 \cup \cdots \cup E_n$$

is a disjoint union.

2.3 The Integral

Definition If ϕ is a non-negative measurable simple function with $\phi = a_1\chi_{E_1} + \cdots + a_n\chi_{E_n}$; $X = E_1 \cup \cdots \cup E_n$ disjoint union we define the *integral of* ϕ w.r.t. the measure m to be

$$\int \phi \, dm = a_1 m(E_1) + \cdots + a_n m(E_n) \in [0, \infty]$$



Recall that $0.\infty = 0$, $a.\infty = \infty$ if $a > 0$.

Definition if E is a measurable subset of X and ϕ is a non-negative measurable simple function on X, we define the *integral of* ϕ *over* E w.r.t. the measure m to be:

$$\int_E \phi \, dm = \int \phi \chi_E \, dm$$

We note that:

1. $\int (\phi + \psi) \, dm = \int \phi \, dm + \int \psi \, dm$
2. $\int c\phi \, dm = c \int \phi \, dm \;\; \forall c \geq 0$

To see 1. we put:

$$\phi = \sum a_i \chi_{E_i}$$

$$\psi = \sum b_j \chi_{F_j}$$

Let $\{c_k\}$ be the set of distinct values of $\{a_i + b_j\}$. Then

$$\phi + \psi = \sum c_k \chi_{G_k}$$

where $G_k = \bigcup E_i \cap F_j$, the union taken over $\{i, j : a_i + b_j = c_k\}$.

$$\begin{aligned}
\int(\phi+\psi) &= \sum c_k m(G_k) \\
&= \sum_k c_k \sum_{\{i,j: a_i+b_j=c_k\}} m(E_i\cap F_j) \\
&= \sum_{i,j}(a_i+b_j)m(E_i\cap F_j) \\
&= \sum a_i m(E_i\cap F_j) + \sum b_j m(E_i\cap F_j) \\
&= \sum_i a_i m(E_i) + \sum_j b_j m(F_j) \\
&= \int \phi\, dm + \int \psi\, dm
\end{aligned}$$

which proves 1.

A very useful property of the integral is:

Theorem 2.3.1. *Fix a simple non-negative measurable function ϕ on X. For each $E \in M$ put*

$$\lambda(E) = \int_E \phi\, dm$$

Then λ is a measure on M.

Proof. Let

$$\phi = a_1\chi_{E_1} + \cdots a_n\chi_{E_n}$$

for $a_i \geq 0$. Then

$$\begin{aligned}
\lambda E &= \int_E \phi\, dm \\
&= \int \phi\chi_E\, dm \\
&= \int (a_1\chi_{E_1\cap E} + \cdots + a_n\chi_{E_n\cap E})\, dm \\
&= a_1 m(E_1\cap E) + \cdots \qquad +a_n m(E_n\cap E) \\
&= a_1 m_{E_1}(E) + \cdots \qquad +a_n m_{E_n}(E)
\end{aligned}$$

therefore

$$\lambda = a_1 m_{E_1} + \cdots + a_n m_{E_n}$$

is a linear combination of measures with non-negative coefficients. Therefore λ is a measure. □

Corollary 2.3.2. *1. If $E = \bigcup_{n=1}^{\infty} E_n$ is a countable disjoint union then*

$$\int_E \phi \, dm = \sum_{n=1}^{\infty} \int_{E_n} \phi \, dm$$

2. If $E = \bigcup_{n=1}^{\infty} E_n$ is a countable increasing union then

$$\int_E \phi \, dm = \lim_{n\to\infty} \int_{E_n} \phi \, dm$$

3. If $E = \bigcap_{n=1}^{\infty} E_n$ is a countable decreasing intersection and $\int_{E_1} \phi \, dm < \infty$ then

$$\int_E \phi \, dm = \lim_{n\to\infty} \int_{E_n} \phi \, dm$$

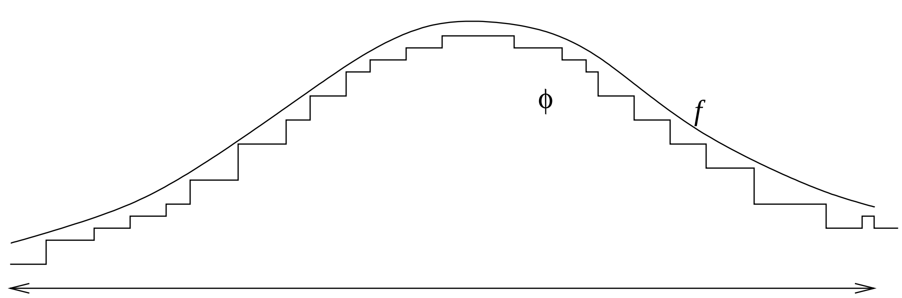
We can now define the integral of any non-negative measurable function.

Definition Let $f : X \longrightarrow [0, \infty]$ be a measurable function. Then we define the *integral of f w.r.t. the measure m* to be:

$$\int f \, dm = \sup_{\phi} \int \phi \, dm$$

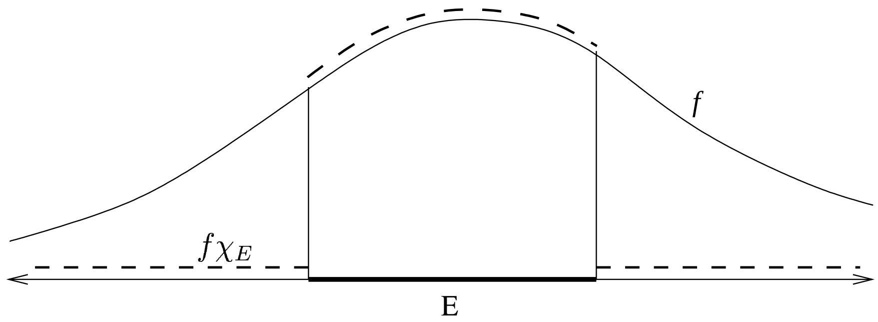
where the sup is taken over all simple measurable functions ϕ such that:

$$0 \leq \phi \leq f$$



Definition If E is a measurable subset of X then we define the *integral of f over E* w.r.t. m to be:

$$\int_E f \, dm = \int f \chi_E \, dm$$



We then have:

1. $f \leq g \implies \int f\,dm \leq \int g\,dm$
2. $E \subset F \implies \int_E f\,dm \leq \int_F f\,dm$

2.4 Monotone Convergence Theorem

We can now prove our first important theorem on integration.

Theorem 2.4.1. (Monotone Convergence Theorem, MCT) *Let* $\{f_n\}$ *be a monotone increasing sequence of non-negative measurable functions. Then*

$$\int \lim f_n\,dm = \lim \int f_n\,dm$$

Proof. We write $f_n \uparrow f$ to denote that $\{f_n\}$ is an increasing sequence of functions with $\lim f_n = f$. We need to prove that:

$$\lim \int f_n\,dm = \int f\,dm$$

1. $$f_n \leq f_{n+1} \leq f$$

   for all n. Therefore:

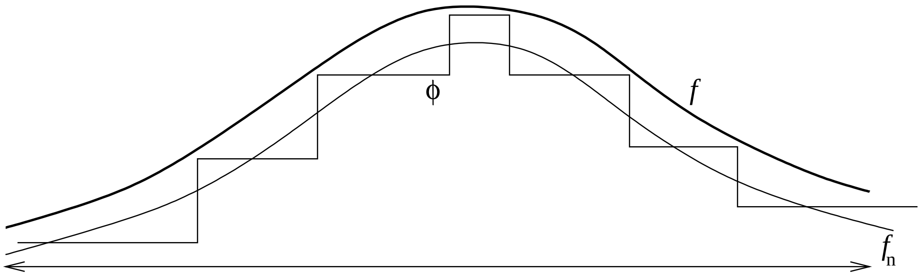
   $$\int f_n\,dm \leq \int f_{n+1}\,dm \leq \int f\,dm$$

   for all n. Therefore:

   $$\lim \int f_n\,dm \leq \int f\,dm$$

2. Let ϕ be a simple measurable function s.t.:

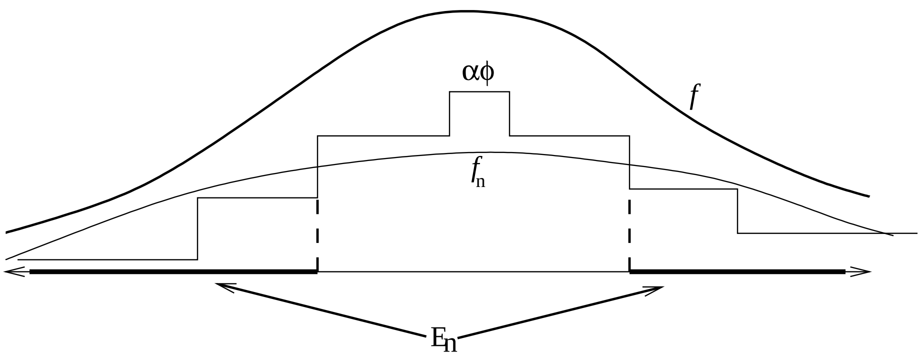
$$0 \leq \phi \leq f$$



Let $0 < \alpha < 1$. For each integer $n > 0$ put

$$E_n = \{x \in X : f_n(x) \geq \alpha\phi(x)\}$$

so $E_n \uparrow X$. Now:



$$\int_{E_n} \alpha\phi \, dm \leq \int_{E_n} f_n \, dm \leq \int f_n \, dm$$

Let $n \to \infty$:

$$\alpha \int \phi \, dm \leq \lim \int f_n \, dm$$

is true $\forall 0 < \alpha < 1$. Therefore:

$$\int \phi \, dm \leq \lim \int f_n \, dm$$

so:

$$\int f \, dm \leq \lim \int f_n \, dm$$

hence

$$\int f \, dm = \lim \int f_n \, dm$$

□

Definition In *probability theory* we have a measure space:

$$\begin{pmatrix} \underset{\substack{\text{sample}\\ \text{space}}}{X} & , & \underset{\text{events}}{M} & , & P \end{pmatrix}$$

with $P(X) = 1$. X is the *sure event.*

A measurable function:

$$f : X \longrightarrow \mathbb{R}$$

is called a *random variable.*

$$P\{x : f(x) \in B\}$$

is the probability that the random variable f takes value in the Borel set $B \subset \mathbb{R}$. If

$$f = a_1\chi_{E_1} + \cdots + a_n\chi_{E_n}$$

with $a_1, \ldots, a_n; E_1, \ldots, E_n$ disjoint, and $E_1 \cup \cdots \cup E_n = X$, then the probability that f takes value a_i is

$$P\{x \in X : f(x) = a_i\} = P(E_i)$$

and

$$\int f\, dP = a_1 P(E_1) + \cdots + a_n P(E_n)$$

is the *average value* or the *expectation* of f.

We define the *expectation* of any random variable f to be:

$$\mathbb{E}(f) = \int f\, dP$$

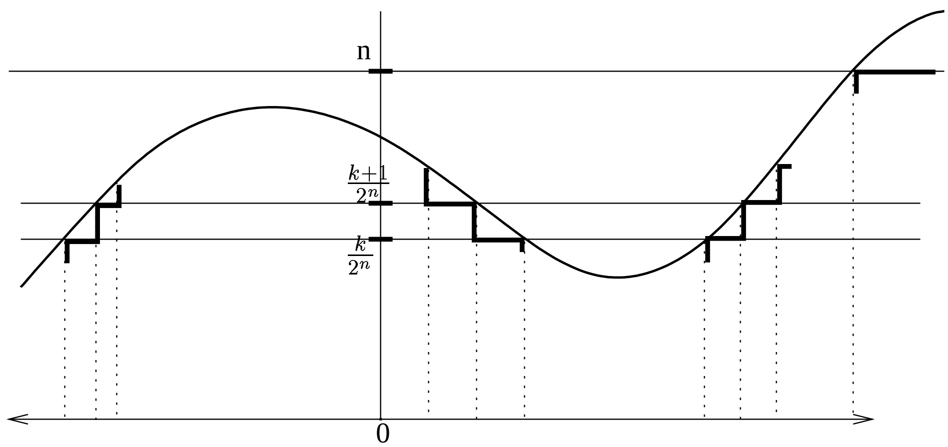
2.5 Existence of Monotone Increasing Simple Functions converging to f

In order to apply the MCT effectively we need:

Theorem 2.5.1. *Let f be a non-negative measurable function $f : X \longrightarrow [0, \infty]$. Then there exists a monotone increasing sequence ϕ_n of simple measurable functions converging to f*

Proof. Put each integer $n > 0$:

$$\phi_n(x) = \begin{cases} \frac{k}{2^n} & \frac{k}{2^n} \le f(x) < \frac{k+1}{2^n} \quad k = 0, 1, 2, \ldots, n2^n - 1 \\ n & f(x) \ge n \end{cases}$$



Then

1. $0 \le \phi_n(x) \le \phi_{n+1}(x)$
2. each ϕ_n is simple and measurable
3. $\lim_{n\to\infty} \phi_n(x) = f(x)$

□

Theorem 2.5.2. *Let f, g be non-negative measurable functions mapping X to $[0, \infty]$ and $c \ge 0$. Then*

1. $\int cf\, dm = c \int f\, dm$
2. $\int (f + g)\, dm = \int f\, dm + \int g\, dm$

Proof. Let ϕ_n, ψ_n be monotonic increasing sequences of non-negative simple functions with $f = \lim \phi_n$, $g = \lim \psi_n$. Then

1. $\int cf\, dm \overset{MCT}{=} \lim \int c\phi_n\, dm = c \lim \int \phi_n\, dm \overset{MCT}{=} = c \int f\, dm$
2. $\int (f + g)\, dm \overset{MCT}{=} \lim \int (\phi_n + \psi_n) = \lim \int \phi_n\, dm + \lim \int \psi_n\, dm \overset{MCT}{=} \int f\, dm + \int g\, dm$

□

This enables us to deal with series:

Theorem 2.5.3. *Let f_n be a sequence of non-negative measurable functions $X \longrightarrow [0, \infty]$. Then:*

$$\int (\sum_{n=1}^{\infty} f_n)\, dm = \sum_{n=1}^{\infty} \int f_n\, dm$$

Proof. put $s_n = f_1 + \cdots + f_n$ sum to n terms. s_n is monotone increasing:

$$\int (\lim_{n\to\infty} s_n)\, dm \stackrel{MCT}{=} \lim_{n\to\infty} \int s_n\, dm = \lim_{n\to\infty} \int \sum_{r=1}^{n} f_r\, dm = \lim_{n\to\infty} \sum_{r=1}^{n} \int f_r\, dm$$

Therefore:

$$\int (\sum_{r=1}^{\infty} f_r)\, dm = \sum_{r=1}^{\infty} \int f_r\, dm$$

□

Theorem 2.5.4. *Let $f : X \longrightarrow [0, \infty]$ be a non-negative measurable and put*

$$\lambda(E) = \int_E f\, dm$$

for each $E \in M$. Then λ is a measure in M.

Proof. Let $E = \bigcup_{k=1}^{\infty} E_k$ be a countable disjoint union. Then

$$\begin{aligned} \lambda(E) &= \textstyle\int_E f\, dm \\ &= \textstyle\int f\chi_E\, dm \\ &= \textstyle\int (\sum_{k=1}^{\infty} f\chi_{E_k})\, dm \\ &\stackrel{MCT}{=} \textstyle\sum_{k=1}^{\infty} \int f\chi_{E_k}\, dm \\ &= \textstyle\sum_{k=1}^{\infty} \int_{E_k} f\, dm \\ &= \textstyle\sum_{k=1}^{\infty} \lambda(E_k) \end{aligned}$$

therefore λ is countably additive, as required. □

Corollary 2.5.5.

1. *if $E = \bigcup_{k=1}^{\infty} E_k$ countable disjoint union then*

$$\int_E f\, dm = \sum_{k=1}^{\infty} \int_{E_k} f\, dm$$

2. *if $E_k \uparrow E$ then $\lim_{k\to\infty} \int_{E_k} f\, dm = \int_E f\, dm$*

3. *If $E_k \downarrow E$ and $\int_{E_1} f\, dm < \infty$, then $\lim_{k\to\infty} \int_{E_k} f\, dm = \int_E f\, dm$*

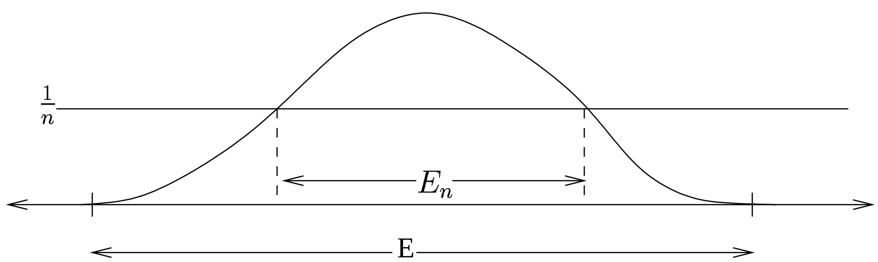
2.6 'Almost Everywhere'

Definition Let $f, g : X \longrightarrow [0, \infty]$. Then we say that $f = g$ *almost everywhere (a.e.)* or $f(x) = g(x)$ *almost all* $x \in X$ *(a.a.x)* if $\{x \in X : f(x) \neq g(x)\}$ has measure zero.

Theorem 2.6.1. *Let* $f : X \longrightarrow [0, \infty]$ *be non-negative measurable. Then* $\int f\,dm = 0 \Leftrightarrow f = 0$ *a.e.*

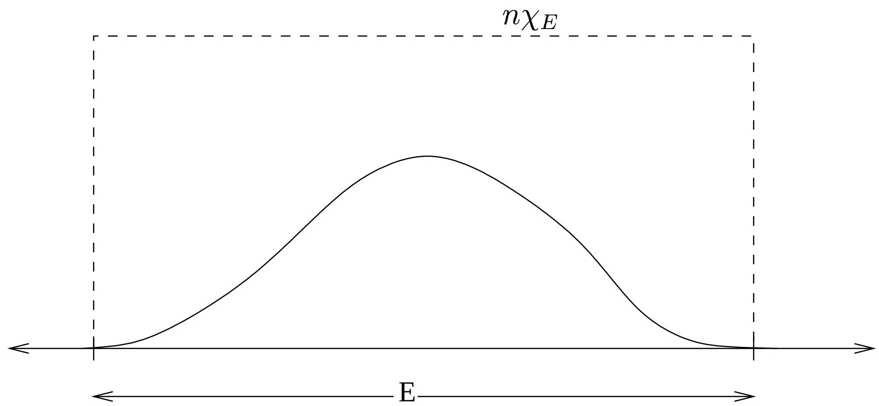
Proof. Put $E = \{x \in X : f(x) \neq 0\}$

1. Put
$$E_n = \{x \in X : f(x) > \frac{1}{n}\}$$
for each integer $n > 0$, so
$$f > \frac{1}{n}\chi_{E_n}$$



Suppose that $\int f\,dm = 0$. Then
$$0 = \int f\,dm \geq \frac{1}{n}m(E_n)$$
so
$$m(E_n) = 0$$
for all n. But $E_n \uparrow E$, so
$$m(E) = \lim m(E_n) = \lim 0 = 0$$
therefore
$$f = 0 \text{ a.e.}$$

2. Suppose $f = 0$ a.e., so $m(E) = 0$. Now,
$$0 \leq f \leq \lim n\chi_E$$



therefore

$$\begin{aligned} 0 &\leq \int f\,dm \\ &\leq \int \lim n\chi_E\,dm \\ &\stackrel{MCT}{=} \lim \int n\chi_E\,dm \\ &= \lim nm(E) \\ &= \lim 0 = 0 \end{aligned}$$

so

$$\int f\,dm = 0$$

□

Corollary 2.6.2. *Let $f : X \longrightarrow [0,\infty]$ be non-negative and measurable and let E have measure zero. Then*

$$\int_E f\,dm = 0$$

Proof.

$$f\chi_E = 0 \text{ a.e.}$$

therefore

$$\int_E f\,dm = \int f\chi_E\,dm = 0$$

□

Corollary 2.6.3. *Let $f, g : X \longrightarrow [0,\infty]$ be non-negative and measurable and $f = g$ a.e. then $\int f\,dm = \int g\,dm$*

Proof. Let $f = g$ on E and $m(E') = 0$ then

$$\int f\,dm = \int_E f\,dm + \int_{E'} f\,dm = \int_E g\,dm + \int_{E'} g\,dm = \int g\,dm$$

□

Thus changing f on a set of measure zero makes no difference to $\int f\,dm$. Also

Theorem 2.6.4. *If f_n are non-negative and $f_n \uparrow f$ a.e. then*

$$\lim \int f_n\,dm = \int f\,dm$$

Proof. suppose $f_n \geq 0$ and $f_n \uparrow f$ on E with $m(E') = 0$. Then

$$\begin{aligned} \int f\,dm &= \int_E f\,dm + \int_{E'} f\,dm \\ &\overset{MCT}{=} \lim \int_E f_n\,dm + 0 \\ &= \lim \left[\int_E f_n\,dm + \int_{E'} f_n\,dm\right] \\ &= \lim \int f_n\,dm \end{aligned}$$

□

So far we have dealt with functions

$$X \longrightarrow [0, \infty]$$

which are non-negative, but have allowed the value ∞.

Now we look at functions

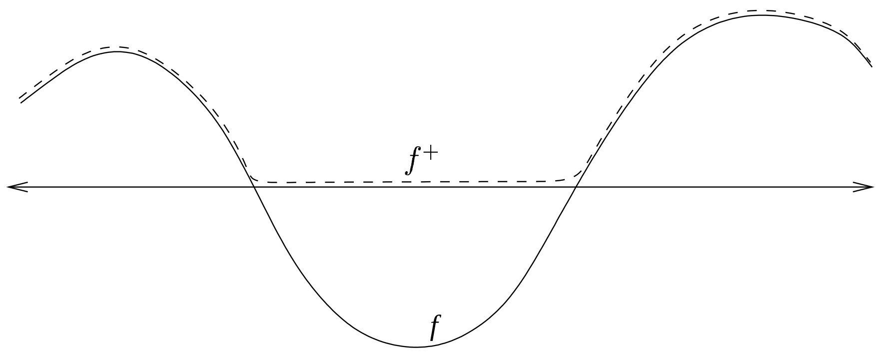
$$X \longrightarrow \mathbb{R}$$

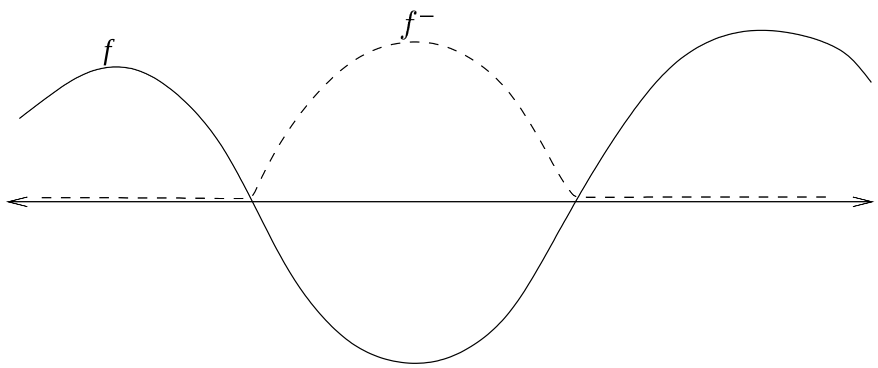
which may be negative and we do *not* allow ∞ as a value.

Definition Let $f : X \longrightarrow \mathbb{R}$ be measurable. Put

$$f^+(x) = \left\{ \begin{array}{ll} f(x) & f(x) \geq 0 \\ 0 & f(x) \leq 0 \end{array} \right\}$$

$$f^-(x) = \left\{ \begin{array}{ll} -f(x) & f(x) \leq 0 \\ 0 & f(x) \geq 0 \end{array} \right\}$$





Thus $f = f^+ - f^-$ and both f^+, f^- are non-negative. We say that f is *integrable w.r.t. m* if

$$\int f^+ \, dm < \infty \text{ and } \int f^- \, dm < \infty$$

and we write

$$\int f \, dm = \int f^+ \, dm - \int f^- \, dm$$

and call it the *integral of f* (w.r.t. measure m).

If E is measurable we write

$$\int_E f \, dm = \int f \chi_E \, dm = \int_E f^+ \, dm - \int_E f^- \, dm$$

Theorem 2.6.5. *Let $f = f_1 - f_2$ where f_1, f_2 are non-negative and measurable and*

$$\int f_1 \, dm < \infty \text{ and } \int f_2 \, dm < \infty$$

Then f is integrable and

$$\int f \, dm = \int f_1 \, dm - \int f_2 \, dm$$

Proof. 1. $f = f_1 - f_2$, therefore $f^+ \leq f_1; f^- \leq f_2$. So

$$\int f^+ \, dm \leq \int f_1 \, dm < \infty; \quad \int f^- \, dm \leq \int f_2 \, dm < \infty$$

therefore f is integrable.

2. $f = f_1 - f_2 = f^+ - f^-$. So

$$f_1 + f^- = f^+ + f_2$$

therefore

$$\int f_1 \, dm + \int f^- \, dm = \int f^+ \, dm + \int f_2 \, dm$$

so

$$\int f_1 \, dm - \int f_2 \, dm = \int f^+ \, dm - \int f^- \, dm = \int f \, dm$$

as required.

□

Theorem 2.6.6. *Let f, g be integrable and $f = g$ a.e. Then*

$$\int f \, dm = \int g \, dm$$

Proof.

$$f^+ = g^+ \text{ a.e.} \implies \int f^+ \, dm = \int g^+ \, dm$$

$$f^- = g^- \text{ a.e.} \implies \int f^- \, dm = \int g^- \, dm$$

so

$$\int f \, dm = \int f^+ \, dm - \int f^- \, dm = \int g^+ \, dm - \int g^- \, dm = \int g \, dm$$

□

So when integrating we can ignore sets of measure zero.

Theorem 2.6.7. *Let $f : X \longrightarrow \mathbb{R}$ be measurable. Then*

1. *f is integrable $\Leftrightarrow |f|$ is integrable*
2. *if f is integrable then*

$$\left| \int f \, dm \right| \leq \int |f| \, dm$$

Proof. 1.

$$\begin{aligned} f \text{ integrable } &\Leftrightarrow \int f^+ \, dm < \infty \text{ and } \int f^- \, dm < \infty \\ &\Leftrightarrow \int (f^+ + f^-) \, dm < \infty \\ &\Leftrightarrow \int |f| \, dm < \infty \\ &\Leftrightarrow |f| \text{ integrable} \end{aligned}$$

2.

$$\begin{aligned} \left| \int f \, dm \right| &= \left| \int f^+ \, dm - \int f^- \, dm \right| \\ &\leq \int f^+ \, dm + \int f^- \, dm \\ &= \int |f| \, dm \end{aligned}$$

□

Definition A complex valued function $f : X \longrightarrow \mathbb{C}$ with

$$f(x) = f_1(x) + i f_2(x)$$

(say) (f_1, f_2 real) is called *integrable* if f_1 and f_2 are integrable and we define:

$$\int f \, dm = \int f_1 \, dm + i \int f_2 \, dm$$

Thus

$$\Re \int f \, dm = \int (\Re f) \, dm$$

$$\Im \int f \, dm = \int (\Im f) \, dm$$

We have $|f_1| \leq |f|$, $|f_2| \leq |f|$, and therefore

$|f|$ integrable $\Leftrightarrow$ $|f_1|$ and $|f_2|$ integrable $\Leftrightarrow$ f_1 and f_2 integrable $\Leftrightarrow$ f integrable.

We also have:

f, g integrable and $c \in \mathbb{C} \implies f + g$ and cf integrable, and

$$\int (f + g) \, dm = \int f \, dm + \int g \, dm$$

and

$$\int (cf) \, dm = c \int f \, dm$$

Thus, the set $\mathcal{L}(X, \mathbb{R}, m)$ of all integrable real valued functions on X is a real vector space, and the set $\mathcal{L}(X, \mathbb{C}, m)$ of all integrable complex valued functions on X is a complex vector space. And on each space:

$$f \longrightarrow \int f\,dm$$

is a linear form.

If $f : X \longrightarrow \mathbb{C}$ is integrable then

$$\int f\,dm = \left|\int f\,dm\right| e^{i\theta} \text{ (say) } \theta \text{ real}$$

therefore

$$\begin{aligned}\overset{\text{(real)}}{\left|\textstyle\int f\,dm\right|} &= e^{-i\theta}\textstyle\int f\,dm \\ &= \textstyle\int \overset{\text{(real)}}{e^{-i\theta} f}\,dm \\ &= \textstyle\int \Re[e^{-i\theta} f]\,dm \\ &\leq \textstyle\int \left|e^{-i\theta} f\right|\,dm \\ &= \textstyle\int |f|\,dm\end{aligned}$$

therefore

$$\left|\int f\,dm\right| \leq \int |f|\,dm$$

2.7 Integral Notation

Definition When dealing with 1-dimensional Lebesgue measure we write

$$\int_{[a,b]} f\,dm = \int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

if $a \leq b$. It follows that

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx$$

and

$$\int_a^b dx = b - a$$

$\forall a, b, c$.

Notice that x is a *dummy symbol* and that

$$\int_a^b f(x)\,dx = \int_a^b f(y)\,dy = \int_a^b f(t)\,dt = \cdots$$

just as

$$\sum_{i=1}^n a_i b_i = \sum_{j=1}^n a_j b_j = \cdots$$

2.8 Fundamental Theorem of Calculus

Theorem 2.8.1. Fundamental Theorem of Calculus *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be continuous and put*

$$F(t) = \int_a^t f(x)\,dx$$

then

$$F'(t) = f(t)$$

Proof. Let $t \in \mathbb{R}$, let $\epsilon > 0$. Then $\exists \delta > 0$ s.t.

$$|f(t+h) - f(t)| \le \epsilon \;\; \forall |h| \le \delta$$

by continuity of f. Therefore

$$\begin{aligned} &\left|\tfrac{F(t+h)-F(t)}{h} - f(t)\right| \\ &= \left|\tfrac{1}{h}\textstyle\int_t^{t+h} f(x)\,dx - \tfrac{1}{h}\int_t^{t+h} f(t)\,dx\right| \\ &= \tfrac{1}{|h|}\left|\textstyle\int_t^{t+h}[f(x) - f(t)]\,dx\right| \\ &\le \tfrac{1}{|h|}\epsilon|h| = \epsilon \end{aligned}$$

for all $|h| \le \delta; h \ne 0$, and hence the result.

Corollary 2.8.2. *If $G : \mathbb{R} \longrightarrow \mathbb{R}$ is C^1 (i.e. has a continuous derivative) then $\int_a^b G'(x)\,dx = G(b) - G(a)$.*

Proof. put $F(t) = \int_a^t G'(x)\,dx$. Then

$$F'(t) = G'(t)$$

for all t. Therefore

$$F(t) = G(t) + c$$

for all t, where c is constant. Therefore

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

as required. □

□

Theorem 2.8.3. (Change of Variable)

Let

$$[t_3, t_4] \xrightarrow{g} [t_1, t_2] \xrightarrow{f} \mathbb{R}$$

with f continuous and $g \in C^1$; $g(t_3) = t_1, g(t_4) = t_2$. Then

$$\int_{t_1}^{t_2} f(x)\,dx = \int_{t_3}^{t_4} f(g(y))g'(y)\,dy$$

Proof. Put

$$G(t) = \int_{t_1}^{t} f(x)\,dx$$

then $G'(t) = f(t)$, and therefore:

$$\begin{aligned}
\textstyle\int_{t_3}^{t_4} f(g(y))g'(y)\,dy \quad &= \textstyle\int_{t_3}^{t_4} G'(g(y)g'(y)\,dy \\
&= \textstyle\int_{t_3}^{t_4} [\frac{d}{dy} G(g(y))]\,dy \\
&= G(g(t_4)) - G(g(t_3)) \\
&= G(t_2) - G(t_1) \\
&= \textstyle\int_{t_1}^{t_2} f(x)\,dx
\end{aligned}$$

as required. □

To deal with sequences which are not monotone we need the concepts of $\liminf$ and $\limsup$.

Definition Let

$$\{a_n\} = a_1, a_2, a_3, \ldots$$

be a sequence in $[-\infty, \infty]$. Put

$$b_n = \inf\{a_n, a_{n+1}, a_{n+2}, \ldots\}$$

$$c_n = \sup\{a_n, a_{n+1}, a_{n+2}, \ldots\}$$

Then

$$b_1 \leq b_2 \cdots \leq b_n \leq b_{n+1} \leq \cdots \leq c_{n+1} \leq c_n \leq \cdots \leq c_2 \leq c_1$$

Define:

$$\liminf a_n = \lim b_n$$

$$\limsup a_n = \lim c_n$$

Then

$$\liminf a_n \leq \limsup a_n$$

and a_n converges iff $\liminf a_n = \limsup a_n (= \lim a_n)$.

2.9 Fatou's Lemma

Theorem 2.9.1. (Fatou's Lemma) *Let f_n be a sequence of non-negative measurable functions:*

$$f_n : X \longrightarrow [0, \infty]$$

Then

$$\int \liminf f_n \, dm \leq \liminf \int f_n \, dm$$

Proof. Put

$$g_r = \inf\{f_r, f_{r+1}, \ldots\}$$

so that

$$\cdots \leq g_r \leq g_{r+1} \leq \cdots$$

Now:

$$\begin{aligned}
& f_n \geq g_r && \forall n \geq r \\
\Longrightarrow \quad & \textstyle\int f_n \geq \int g_r && \forall n \geq r \\
\Longrightarrow \quad & \liminf \textstyle\int f_n \geq \int g_r && \forall r \\
\Longrightarrow \quad & \liminf \textstyle\int f_n \geq \lim \int g_r \overset{MCT}{=} \int \lim g_r = \int \liminf f_n
\end{aligned}$$

as required. ∎

2.10 Dominated Convergence Theorem

Theorem 2.10.1. Lebesgue's Dominated Convergence Theorem, DCT *Let f_n be integrable and $f_n \to f$. Let*

$$|f_n| \leq g$$

for all n where g is integrable. Then f is integrable and $\int f = \lim \int f_n$

Proof. $|f| = \lim |f_n| \leq g$. Therefore $|f|$ is integrable, and therefore f is integrable. Now

$$g \pm f_n \geq 0$$

therefore

$$\int \liminf[g \pm f_n] \leq_{FATOU} \liminf \int [g \pm f_n]$$

so

$$\int [g \pm f] \leq \int g + \liminf \pm \int f_n$$

so

$$\pm \int f \leq \liminf \pm \int f_n$$

Which gives:

$$\oplus : \quad \int f \leq \liminf \int f_n$$

$$\ominus : \quad -\int f \leq -\limsup \int f_n$$

therefore

$$\int f \leq \liminf \int f_n \leq \limsup \int f_n \leq \int f$$

which is equivalent to:

$$\lim \int f_n = \int f$$

as required. □

For series this leads to:

Theorem 2.10.2. Dominated convergence theorem for series *Let f_n be a sequence of integrable functions such that*

$$\sum_{n=1}^{\infty} \int |f_n| \, dm < \infty$$

Then $\sum_{n=1}^{\infty} f_n$ is (equal a.e. to) an integrable function and

$$\int (\sum_{n=1}^{\infty} f_n) \, dm = \sum_{n=1}^{\infty} \int f_n \, dm$$

Proof. Put

$$s_n = f_1 + f_2 + \cdots + f_n$$

sum to n terms. Then

$$|s_n| \le |f_1| + \cdots + |f_n| \le \sum_{r=1}^{\infty} |f_r| = g$$

(say). Then

$$\begin{aligned} \int g\,dm &= \int \textstyle\sum_{r=1}^{\infty} |f_r|\,dm \\ &\stackrel{MCT}{=} \textstyle\sum_{r=1}^{\infty} \int |f_r|\,dm < \infty \end{aligned}$$

Therefore g is integrable (a.e. equal to an integrable fn). Therefore $\lim s_n$ is integrable and

$$\int \lim s_n\,dm \stackrel{DCT}{=} \lim \int s_n\,dm$$

therefore

$$\int \sum_{r=1}^{\infty} f_r\,dm = \sum_{r=1}^{\infty} \int f_r\,dm$$

as required. □

2.11 Differentiation under the integral sign

Another useful application.

Theorem 2.11.1. Differentiation under the integral sign *Let $f(x,t)$ be an integrable function of $x \in X$ for each $a \le t \le b$ and differentiable w.r.t t.*

Suppose

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \le g(x)$$

for all $a \le t \le b$, where g is integrable. Then

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

Proof. Put

$$F(t) = \int f(x,t)\,dx$$

Let $a \le t \le b$. Choose a sequence t_n in $[a,b]$ s.t. $\lim t_n = t$ and $t_n \ne t$. Then, by the Mean Value Theorem:

$$\begin{aligned} |f(x,t_n) - f(x,t)| &= |t_n - t| \left| \tfrac{\partial f}{\partial t}(x, c(n,x,t)) \right| \\ &\le |t_n - t| g(x) \end{aligned}$$

Therefore

$$\left| \frac{f(x,t_n) - f(x,t)}{t_n - t} \right| \le g(x)$$

so

$$\begin{aligned} \lim \tfrac{F(t_n)-F(t)}{t_n-t} &= \lim \int \tfrac{f(x,t_n)-f(x,t)}{t_n-t}\, dx \\ &\overset{DCT}{=} \int \lim \tfrac{f(x,t_n)-f(x,t)}{t_n-t}\, dx \\ &= \int \tfrac{\partial f}{\partial t}(x,t)\, dx \end{aligned}$$

and therefore

$$\frac{dF}{dt} = \int \frac{\partial f}{\partial t}(x,t)\, dx$$

as required. □

Chapter 3

Multiple Integration

3.1 Product Measure

We have established the Lebesgue measure and Lebesgue integral on $\mathbb{R}$. To consider integration on

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$$

we use the concept of a *product measure*.

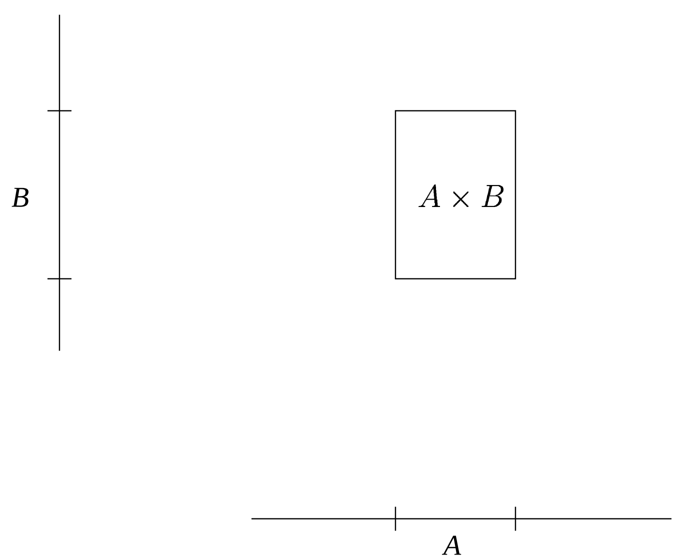
We proceed as follows:

Definition Let l be a measure on a σ-algebra $\mathcal{L}$ of subsets of X. Let m be a measure on a σ-algebra $\mathcal{M}$ of subsets of Y.

Call

$$\{A \times B : A \in \mathcal{L}, B \in \mathcal{M}\}$$

the *set of rectangles* in $X \times Y$



Lemma 3.1.1. *Let $A \times B = \bigcup_{i=1}^{\infty} A_i \times B_i$ be a rectangle written as a countable disjoint union of rectangles. Then*

$$l(A)m(B) = \sum_{i=1}^{\infty} l(A_i)m(B_i)$$

Proof. We have

$$\chi_{A\times B} = \sum_{i=1}^{\infty} \chi_{A_i\times B_i}$$

$$\implies \chi_A(x)\chi_B(y) = \sum_{i=1}^{\infty} \chi_{A_i}(x)\chi_{B_i}(y)$$

Fix x and integrate w.r.t. m term by term using the Monotone Convergence Theorem:

$$\chi_A(x)m(B) = \sum_{i=1}^{\infty} \chi_{A_i}(x)m(B_i)$$

Now integrate w.r.t. x using MCT to get:

$$l(A)m(B) = \sum_{i=1}^{\infty} l(A_i)m(B_i)$$

as required. □

Definition Let $\mathcal{A}$ be the collection of all finite unions of rectangles in $X \times Y$. Each element of $\mathcal{A}$ is a finite disjoint union of rectangles. For each $E \in \mathcal{A}$ such that

$$E = \bigcup_{i=1}^{\infty} A_i \times B_i$$

is a countable disjoint union of rectangles we define:

$$\pi(E) = \sum_{i=1}^{\infty} l(A_i)m(B_i)$$

Theorem 3.1.2. *π is well-defined and is a measure on $\mathcal{A}$.*

Proof. 1. *well-defined*

Suppose

$$E = \bigcup_{i=1}^{\infty} A_i \times B_i = \bigcup_{j=1}^{\infty} C_j \times D_j$$

then

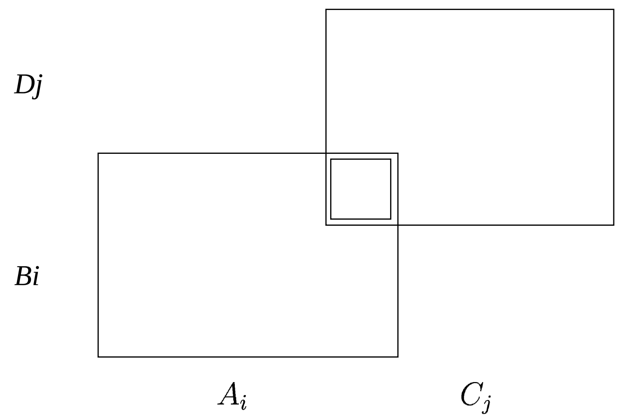
$$A_i \times B_i = \bigcup_{j=1}^{\infty}(A_i \cap C_j) \times (B_i \cap D_j)$$

Therefore, by Lemma 3.1.1

$$l(A_i)m(B_i) = \sum_{j=1}^{\infty} l(A_i \cap C_j)m(B_i \cap D_j)$$

and hence,

$$\begin{aligned} \textstyle\sum_{i=1}^{\infty} l(A_i)m(B_i) &= \textstyle\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} l(A_i \cap C_j)m(B_i \cap D_j) \\ &= \textstyle\sum_{j=1}^{\infty} l(C_j)m(D_j) \end{aligned}$$



2. *countably additive* Let $E = \bigcup_{i=1}^{\infty} E_i$ be a countable disjoint union with $E, E_i \in \mathcal{A}$.

$$E_i = \bigcup_{j=1}^{n_i} A_{ij} \times B_{ij}$$

is a finite disjoint union (say). Then

$$E = \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{n_i} A_{ij} \times B_{ij}$$

Therefore

$$\pi(E) = \sum_{i=1}^{\infty} \sum_{j=1}^{n_i} l(A_{ij})m(B_{ij}) = \sum_{i=1}^{\infty} \pi(E_i)$$

Therefore π is countably additive, as required.

$\square$

Definition The measure π on $\mathcal{A}$ extends to a measure (also denoted by π and called the *product* of the measures l and m) on the σ-algebra (denoted $\mathcal{L} \times \mathcal{M}$) of all subsets of $X \times Y$ which are measurable w.r.t. π.

Similarly, if

$$(X_1, M_1, m_1), \ldots, (X_n, M_n, m_n)$$

is a sequence of measure spaces then we have a product measure on a σ-algebra of subsets of

$$X_1 \times \cdots \times X_n$$

s.t.

$$\pi(E_1 \times \cdots \times E_n) = m_1(E_1) m_2(E_2) \ldots m_n(E_n)$$

In particular, starting with 1-dimensional Lebesgue measure on $\mathbb{R}$ we get *n-dimensional Lebesgue Measure* on

$$\mathbb{R} \times \cdots \times \mathbb{R} = \mathbb{R}^n$$

Example If we have two successive, independent events with independent probability measures, P_1, P_2, then the probability of the first event E and the second event F is:

$$\begin{aligned} P(E \times F) &= P[(E \times Y) \cap (X \times F)] \\ &= P(E \times Y) P(X \times F) \qquad \text{by independence} \\ &= P_1(E) P_2(F) \end{aligned}$$

(i.e. product measure)

3.2 Monotone Class

Definition A non-empty collection M of subsets of X is called a *monotone class* if

1. $E_n \uparrow E, E_n \in M \implies E \in M$
2. $E_n \downarrow E, E_n \in M \implies E \in M$

i.e. M is closed under countable increasing unions and under countable decreasing intersections.

We have: M is a σ-algebra $\implies$ M is a monotone class. Therefore, there are more monotone classes than σ-algebras.

Definition If $\mathcal{V}$ is a non-empty collection of subsets of X and if M is the intersection of all the monotone classes of subsets of X which contain $\mathcal{V}$, then M is a monotone class, called the monotone class *generated by* $\mathcal{V}$.

M is the smallest monotone class containing $\mathcal{V}$ and

$$\mathcal{V} \quad \subset \quad \begin{array}{c}\text{monotone class}\\ \text{generated by } \mathcal{V}\end{array} \quad \subset \quad \begin{array}{c}\sigma\text{-algebra}\\ \text{generated by } \mathcal{V}\end{array}$$

3.3 Ring of Subsets

Definition A non-empty collection R of subsets of X is called a *ring of subsets* of X if:

$$E, F \in R \implies \begin{cases} E \cup F \in R \\ \\ E \cap F' \in R \end{cases}$$

i.e. R is closed under finite unions and under relative complements.

Example The collection of all finite unions of rectangles contained in the interior X of a fixed circle in $\mathbb{R}^2$ is a ring of subsets of X.

Note:

1. if R is a ring then R is closed under finite intersections since $E \cap F = E \cap (E \cap F')'$.
2. if R is a ring then $\emptyset \in R$ since $E \cap E' = \emptyset$
3. if R is a ring of subsets of X and if $X \in R$ then R is an algebra of subsets of X since $E' = X \cap E'$, so R is closed under complements.

Theorem 3.3.1. (Monotone class lemma)

Let R be a ring of subsets of X and let M be the monotone class generated by R.

Then M is a ring.

Proof. We have to show M is closed under finite unions and relative complements. i.e. that:

$$E \cap F', \; E \cup F, \; E' \cap F \quad \in M$$

for all $E, f \in M$. So for each $E \in M$ put

$$M_E = \{F \in M : E \cap F', E \cup F, E' \cap F \in M\}$$

and we must show

$$M_E = M$$

for all $E \in M$. Now

1. $M_E \subset M$ by definition

2. M_E is a monotone class, because:

$$F_n \uparrow F, F_n \in M_E$$

$$\implies (E \cap F_n{}') \downarrow (E \cap F') \quad (E \cup F_n) \uparrow (E \cup F) \quad (E' \cap F_n) \uparrow (E' \cap F)$$

with $E \cap F_n{}'$ $\quad E \cup F_n$ $\quad E' \cap F_n$ $\quad$ all $\in M$

$\implies E \cap F'$ $\quad E \cup F$ $\quad E' \cap F$ $\quad$ all $\in M$ (Mi monotone)

$$\implies F \in M_E$$

therefore M_E is closed under countable increasing unions, and similarly M_E is closed under countable decreasing intersections.

3. $E \in R$

$\implies R \subset M_E$ since R is a ring. Therefore $M_E = M$ by (i), (ii) since M is smallest monotone class containing R.

4. $F \in M_E \ \ \forall E \in R, F \in M$ by (iii). Therefore $E \in M_F \ \ \forall E \in R, F \in M$.

Therefore

$$R \subset M_F \ \ \forall F \in M$$

and therefore

$$M_F = M \ \ \forall F \in M$$

as required.

□

Corollary 3.3.2. *Let M be the monotone class generated by a ring R, and let $X \in M$.*

Then $M = \mathcal{G}(R)$ the σ-algebra generated by R.

Proof. M is a ring and $X \in M$

$$E \in M \implies E' = E' \cap X \in M$$

therefore M closed under compliments, and therefore M is an algebra.

Also, if E_n is a sequence in M and $E = \bigcup_1^\infty E_n$ put

$$F_n = E_1 \cup \cdots \cup E_n \quad \in M$$

then $F_n \uparrow E$, and therefore $E \in M$.

Therefore, M is a σ-algebra, hence the result. □

Corollary 3.3.3. *Let M be the monotone class generated by an algebra $\mathcal{A}$. Then $M = \mathcal{G}(\mathcal{A})$ the σ-algebra generated by $\mathcal{A}$.*

Proof. $X \in \mathcal{A} \implies X \in M$, and hence the result by previous corollary. □

3.4 Integration using Product Measure

Let l be a measure on a σ-algebra $\mathcal{L}$ of subsets of X.

Let m be a measure on a σ-algebra $\mathcal{M}$ of subsets of Y.

Let $\mathcal{A}$ be the algebra of finite unions of rectangles $A \times B$; $A \in \mathcal{L},\ B \in \mathcal{M}$.

Let π be the product measure on the σ-algebra $\mathcal{G}(\mathcal{A})$.

If

$$f : X \longrightarrow [0, \infty]$$

$$g : Y \longrightarrow [0, \infty]$$

$$F : X \times Y \longrightarrow [0, \infty]$$

are non-negative measurable, write:

$$\int f\, dl = \int_X f(x)\, dx$$

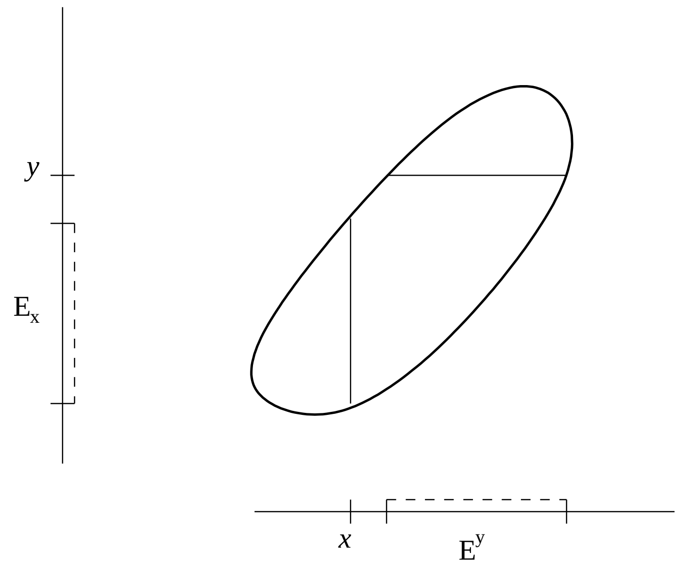
$$\int f\, dm = \int_Y g(y)\, dy$$

$$\int F\, d\pi = \int_{X\times Y} F(x, y)\, dx\, dy$$

If $E \subset X \times Y$ write

$$E_x = \{y \in Y : (x, y) \in E\}$$

$$E^y = \{x \in X : (x, y) \in E\}$$



Theorem 3.4.1. *if $E \in G(A)$ and l, m are σ-finite then:*

$$\pi(E) = \int_X m(E_x)\,dx = \int_Y l(E^y)\,dy$$

Proof. To show

$$\pi(E) = \int_X m(E_x)\,dx \tag{3.1}$$

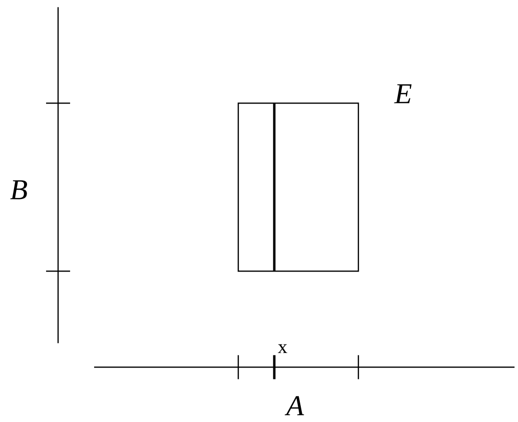
1. Suppose l, m are finite measures: $l(X) < \infty$, $m(Y) < \infty$. Let $\mathcal{N}$ be the collection of all $E \in \mathcal{G}(\mathcal{A})$ s.t. Equation (3.1) holds. We will show that N is a monotone class containing $\mathcal{A}$:

$$\mathcal{A} \subset N \subset \mathcal{G}(\mathcal{A})$$

   and hence $N = \mathcal{G}(\mathcal{A})$ since by the corollary to the monotone class lemma, $\mathcal{G}(\mathcal{A})$ is the monotone class generated by $\mathcal{A}$.

   (a) $\mathcal{A} \in N$.
   If $E = A \times B$ is a rectangle then

$$m(E_x) = \chi_A(x)m(B)$$



then

$$\int_X m(E_x)\,dx = l(A)m(B) = \pi(E)$$

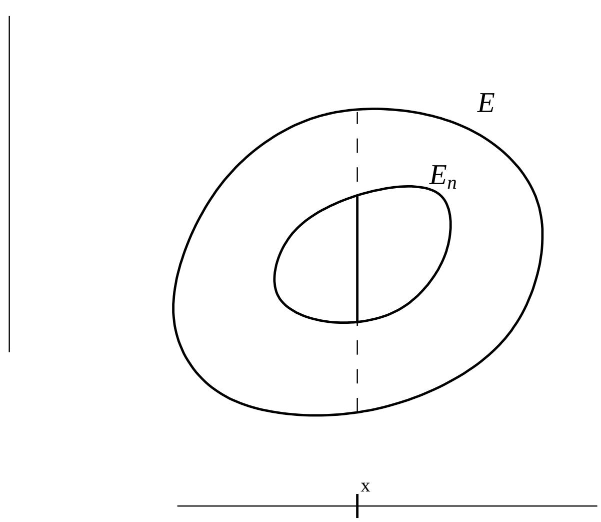
therefore $E \in N$. Now each element of $\mathcal{A}$ can be written as a finite disjoint union of rectangles, therefore

$$\mathcal{A} \subset N$$

(b) Let $E_n \uparrow E, E_n \in N$. So,

$$(E_n)_x \uparrow E_x$$

for each $x \in X$.



Then

$$\begin{aligned}
\int_X m(E_x)\,dx &= \int_X \lim m((E_n)_x)\,dx \\
&\overset{MCT}{=} \lim \int_X m((E_n)_x)\,dx \\
&= \lim \pi(E_n) \\
&= \pi(E)
\end{aligned}$$

therefore $E \in N$, and N is closed under converging unions.

(c) Let $E \in N$. Then

$$\begin{aligned}
\int_X m((E')_x)\,dx &= \int_X [m(Y) - m(E_x)]\,dx \\
&= l(X)m(Y) - \pi(E)x \\
&= \pi(X \times Y) - \pi(E) \\
&= \pi(E').
\end{aligned}$$

Therefore $E' \in N$, and N is closed under complements. Therefore N is a monotone class, and therefore $N = \mathcal{G}(\mathcal{A})$.

2. Suppose l, m are σ-finite,

$$A_n \uparrow X, \quad B_n \uparrow Y$$

(say) with

$$l(A_n) < \infty, \quad m(B_n) < \infty$$

Then

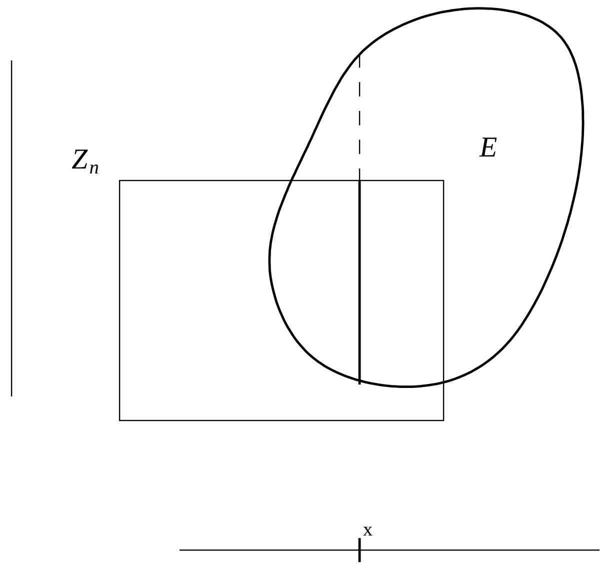
$$Z_n \uparrow (X \times Y)$$

where $Z_n = A_n \times B_n$ and $\pi(Z_n) < \infty$.

Let $E \in \mathcal{G}(\mathcal{A})$. Then $(E \cap Z_n) \uparrow E$, and

$$(E \cap Z_n)_x \uparrow E_x$$

for all $x \in X$.



So:

$$\int_X m(E_x)\,dx = \int_X \lim m((E\cap Z_n)_x)\,dx$$

$$\overset{MCT}{=} \lim \int_X m((E\cap Z_n)_x)\,dx$$

(since Z_n has finite measure)

$$= \lim \pi(E\cap Z_n)$$

$$= \pi(E)$$

as required.

□

3.5 Tonelli's Theorem

Theorem 3.5.1. (Repeated integral of a non-negative funtcion)
Let $F : X \times Y \longrightarrow [0,\infty]$ be non-negative measurable. Then

$$\int_X \left[\int_Y F(x,y)\,dy\right]\,dx = \int_{X\times Y} F(x,y)\,dx\,dy = \int_Y \left[\int_X F(x,y)\,dx\right]\,dy$$

(notation as before.)

Proof. To show

$$\int_X \left[\int_Y F(x,y)\,dy\right] dx = \int_{X\times Y} F(x,y)\,dx\,dy \tag{3.2}$$

1. Equation 3.2 holds for $F = \chi_E$, since

$$\int_X \left[\int_Y \chi_E(x,y)\,dy\right] dx = \int_X m(E_x)\,dx = \pi(E) = \int_{X\times Y} \chi_E(x,y)\,dx\,dy$$

   therefore Equation 3.2 also holds for any simple function.

2. *(General Case)*

   $\exists$ monotone increasing sequence of non-negative measurable functions with $F = \lim F_n$

$$\begin{aligned}
\textstyle\int_X \left[\int_Y F(x,y)\,dy\right] dx &= \textstyle\int_X \left[\int_Y \lim F_n(x,y)\,dy\right] dx \\
&\overset{MCT}{=} \textstyle\int_X \left[\lim \int_Y F_n(x,y)\,dy\right] dx \\
&\overset{MCT}{=} \textstyle\lim \int_X \left[\int_Y F_n(x,y)\,dy\right] dx \\
&\overset{(1.)}{=} \textstyle\lim \int_{X\times Y} F_n(x,y)\,dx\,dy \\
&\overset{MCT}{=} \textstyle\int_{X\times Y} \lim F_n(x,y)\,dx\,dy \\
&= \textstyle\int_{X\times Y} F(x,y)\,dx\,dy
\end{aligned}$$

□

3.6 Fubini's Theorem

Theorem 3.6.1. (Repeated Integral of an integrable function)
Let F be an integrable function on $X \times Y$. Then:

$$\int_X \left[\int_Y F(x,y)\,dy\right] dx = \int_{X\times Y} F(x,y)\,dx\,dy = \int_Y \left[\int_X F(x,y)\,dx\right] dy$$

Where $\int_Y F(x,y)\,dy$ is equal to an integrable function of x a.e., and $\int_X F(x,y)\,dx$ is equal to an integrable function of y a.e.

Proof.

$$\begin{aligned} \int_{X\times Y} F(x,y)\,dx\,dy &= \int_{X\times Y} F^+(x,y)\,dx\,dy - \int_{X\times Y} F^-(x,y)\,dx\,dy \\ &\overset{Tonelli}{=} \int_X \left[\int_Y F^+(x,y)\,dy\right]\,dx - \int_X \left[\int_Y F^-(x,y)\,dy\right]\,dx \end{aligned}$$

Therefore $\int_Y F^+(x,y)\,dy$, and $\int_Y F^-(x,y)\,dy$ are each finite a.a.x., and each has a finite integral w.r.t. x. (*)

So:

$$F(x,y) = F^+(x,y) - F^-(x,y)$$

is integrable w.r.t y a.e., and:

$$\int_Y F(x,y)\,dy = \int_Y F^+(x,y)\,dy - \int_Y F^-(x,y)\,dy$$

for a.a.x, and is an integrable function of x (by (*)), with:

$$\begin{aligned} \int_X \left[\int_Y F(x,y)\,dy\right]\,dx &= \int_X \left[\int_Y F^+(x,y)\,dy\right]\,dx - \int_X \left[\int_Y F^-(x,y)\,dy\right]\,dx \\ &= \int_{X\times Y} F(x,y)\,dx\,dy \end{aligned}$$

□

Theorem 3.6.2. *Let f be an integrable function $\mathbb{R} \longrightarrow \mathbb{R}$. Then*

1. $\int f(x+c)\,dx = \int f(x)\,dx$
2. $\int f(cx)\,dx = \frac{1}{|c|}\int f(x)\,dx \quad (c \neq 0)$

Proof. These are true for $f = \chi_E$, because:

1.

$$\int \chi_E(x+c)\,dx = \int \chi_{E-c}(x)\,dx = m(E-c) = m(E) = \int \chi_E(x)\,dx$$

2.

$$\int \chi_E(cx)\,dx = \int \chi_{\frac{1}{c}E}(x)\,dx = m(\frac{1}{c}E) = \frac{1}{|c|}m(E) = \frac{1}{|c|}\int \chi_E(x)\,dx$$

Therefore, these are true for f simple, and therefore true for f non-negative measurable (by MCT), since $\exists f_n$ simple, s.t. $f_n \uparrow f$. Therefore, these are true for $f = f^+ - f^-$ integrable. □

We now see how to deal with integration on change of variable on $\mathbb{R}^n$. Recall $\int f(cx)\,dx = \frac{1}{|c|}\int f(x)\,dx$.

Theorem 3.6.3. (Linear change of variable)

Let $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n$ *be an invertable matrix, then*

$$\int f(Ax)\,dx = \frac{1}{|\det A|}\int f(x)\,dx$$

Proof. we can reduce A to the unit matrix I by a sequence of elementary row operations.

1. To replace row i by row $i + c$ row j, multiply A by:

$$N = \begin{pmatrix} 1 & & & & \vdots & & \\ & \ddots & & & \vdots & & \\ \cdots & \cdots & 1 & \cdots & c & \cdots & \cdots \\ & & & & \vdots & & \\ & & & & 1 & & \\ & & & & \vdots & & \end{pmatrix}$$

   with c in the i,j position.

2. To interchange row i and row j, multiply A by:

$$P = \begin{pmatrix} 1 & & \vdots & & & & \vdots & & & \\ & 1 & \vdots & & & & \vdots & & & \\ \cdots & \cdots & 0 & \cdots & & \cdots & 1 & \cdots & & \\ & & \vdots & 1 & & & \vdots & & & \\ & & & & \ddots & & & & & \\ & & \vdots & & & 1 & \vdots & & & \\ \cdots & \cdots & 1 & \cdots & & \cdots & 0 & \cdots & & \\ & & & & & & & 1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{pmatrix}$$

3. To replace row i by c row j, multiply A by:

$$D = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Therefore, there exists matrices $B_1, \dots, B_k$, each of type N, P or D s.t.

$$B_1 B_2 \cdots B_k A = I$$

and therefore,

$$A = B_k^{-1} \cdots B_2^{-1} B_1^{-1}$$

is a product of matrices of type N, P or D.

Now, if the theorem holds for matrices A, B then it also holds for AB since:

$$\int f(ABx)\,dx = \frac{1}{|\det B|}\int f(Ax)\,dx = \frac{1}{|\det B||\det A|}\int f(x)\,dx = \frac{1}{|\det AB|}\int f(x)\,dx$$

therefore, it is sufficient to prove it for matrices of type N, P, D.

1. Let

$$N = \begin{pmatrix} 1 & c & 0 & \cdots & 0 & & \\ 0 & 1 & 0 & \cdots & 0 & & \\ 0 & 0 & 1 & 0 & \cdots & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & & 1 \end{pmatrix}$$

(say). $\det N = 1$. Then

$$\begin{aligned} \int f(Nx)\,dx &= \int f(x_1 + cx_2, x_2, \dots, x_n)\,dx_1\,dx_2 \cdots dx_n \\ &= \int f(x_1, x_2, \dots, x_n)\,dx_1\,dx_2 \cdots dx_n \\ &\quad \text{(by Fubini, and translation invariance)} \\ &= \tfrac{1}{|\det N|}\int f(x)\,dx \end{aligned}$$

2. Let

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \\ 1 & 0 & 0 & \cdots & 0 & \\ 0 & 0 & 1 & 0 & \cdots & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

(say). $\det P = -1$. Then

$$\begin{aligned} \int f(Px)\,dx &= \int f(x_2, x_1, \ldots, x_n)\,dx_1\,dx_2 \cdots dx_n \\ &= \int f(x_1, x_2, \ldots, x_n)\,dx_1\,dx_2 \cdots dx_n \\ &\quad \text{(by Fubini)} \\ &= \frac{1}{|\det P|} \int f(x)\,dx \end{aligned}$$

3. Let

$$D = \begin{pmatrix} c & 0 & 0 & \cdots & 0 & \\ 0 & 1 & 0 & \cdots & 0 & \\ 0 & 0 & 1 & 0 & \cdots & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

(say). $\det N = c$. Then

$$\begin{aligned} \int f(Dx)\,dx &= \int f(cx_1, x_2, \ldots, x_n)\,dx_1\,dx_2 \cdots dx_n \\ &= \frac{1}{|c|} \int f(x_1, x_2, \ldots, x_n)\,dx_1\,dx_2 \cdots dx_n \\ &\quad \text{(by Fubini)} \\ &= \frac{1}{|\det D|} \int f(x)\,dx \end{aligned}$$

□

Corollary 3.6.4. *if $E \subset \mathbb{R}^n$ is measurable and $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n$ is a linear homomorphism then*

$$m(A(E)) = |\det A| m(E)$$

Proof.

$$\begin{aligned} m(E) &= \int \chi_E(x)\,dx \\ &= \int \chi_{A(E)}(AX)\,dx \\ &= \frac{1}{|\det A|} \int \chi_{A(E)}(x)\,dx \\ &= \frac{1}{|\det A|} m(A(E)) \end{aligned}$$

as required. □

Chapter 4

Differentiation

4.1 Differentiation

If $\mathbb{R} \stackrel{f}{\longrightarrow} \mathbb{R}$ is a real valued function od a real variable then the derivative of f at a is defined to be:

$$f'(a) = \lim_{h \longrightarrow 0} \frac{f(a+h) - f(a)}{h} \tag{4.1}$$

We want to define the derivative $f'(a)$ when f is a vector-valued function of a vector variable:

$$f : M \longrightarrow N$$

where m, N are real (or complex) vector spaces.

We cannot use Equation (4.1) directly since we don't know how to divide $f(a+h) - f(a)$, which is a vector in N, by h, which is a vector in M

So, we rewrite Equation (4.1) as:

$$f(a+h) = \underbrace{f(a)}_{\text{(constant)}} + \underbrace{f'(a)h}_{\text{(linear in } h)} + \underbrace{\phi(h)}_{\text{(remainder)}}$$

where

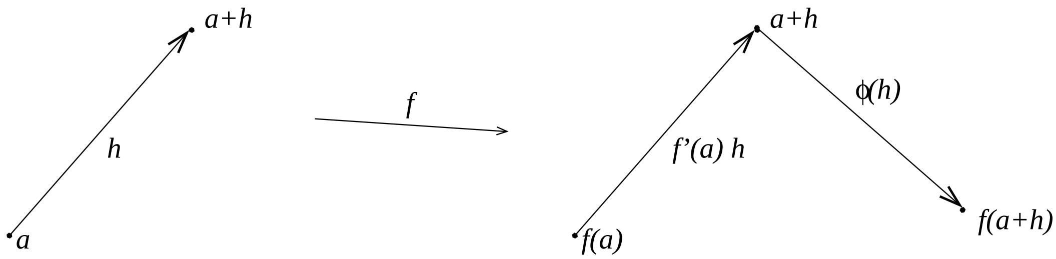
$$\lim_{h \longrightarrow 0} \frac{\phi(h)}{h} = \frac{f(a+h) - f(a)}{h} - f'(a)$$

This suggests that we take M, N to be normed spaces and define $f'(a)$ to be a linear operator such that

$$f(a+h) = f(a) + f'(a)h + \phi(h)$$

where

$$\lim_{||h|| \longrightarrow 0} \frac{\|\phi(h)\|}{\|h\|} = 0$$



This $f'(a)h$ is the linear approx to the change in f when variable changes by h from a to $a+h$.

4.2 Normed Space

Definition let M be a real or complex vector space. Then M is called a *normed space* if a function $\|.\|$ exists,:

$$M \longrightarrow \mathbb{R}$$

$$x \longrightarrow \|x\|$$

is given on M (called the *norm* on M), such that:

1. $\|x \ \geq 0$
2. $\|x\| = 0 \Leftrightarrow x = 0$
3. $\|\alpha x\| = |\alpha| \, \|x\| \, \forall$ scalar α
4. $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)

Example 1. $\mathbb{R}^n$ with $\|(\alpha_1, \ldots, \alpha_n)\| = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}$ is called the *Euclidean Norm* on $\mathbb{R}^n$.

2. $\mathbb{C}^n$ with $\|(\alpha_1, \ldots, \alpha_n)\| = \sqrt{|\alpha_1|^2 + \cdots + |\alpha_n|^2}$ is called the *Hilbert Norm* on $\mathbb{C}^n$.

3. $\mathbb{C}^n$ with $\|(\alpha_1, \ldots, \alpha_n)\| = \max\{|\alpha_1|, \ldots, |\alpha_n|\}$ is called the *sup norm* on $\mathbb{C}^n$.

4. if (X, M, m) is a measure space the the set of integrable functions $\mathcal{L}'(X, \mathbb{R}, m)$ with

$$\|f\| = \int |f| \, dm$$

is called the $\mathcal{L}'$*-norm* (functions are to be regarded as equal if they are equal a.e.)

4.3 Metric Space

Definition a set X is called a *metric space* if a function

$$D : X \times X \longrightarrow \mathbb{R}$$

is given (called a *metric* on X) such that

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \Leftrightarrow x = y$
3. $d(x, z) \leq d(x, y) + d(y, z)$. This is known as the *triangle inequality*

Definition If $a \in X$ and $r > 0$ the we write

$$B_X(a, r) = \{x \in X : d(a, x) < r\}$$

and call it the *ball in* X centre a, radius r.

Example if M is a normed space then M is also a metric space with

$$d(x, y) = \|x - y\|$$

4.4 Topological space

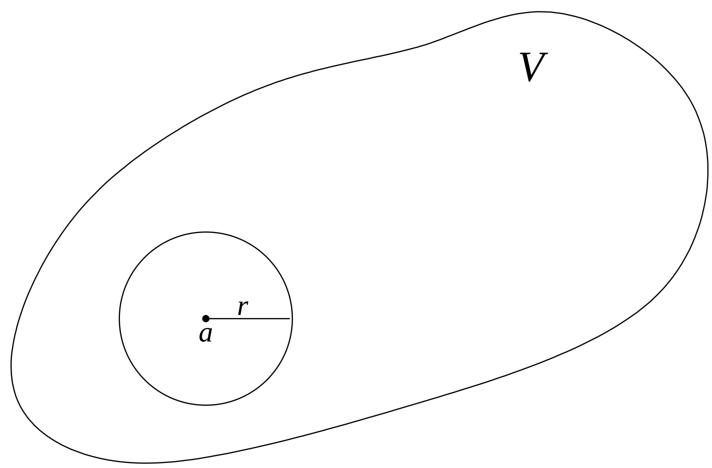
Definition a set X is called a *topological space* id a collection $\mathcal{V}$ of subsets of X is given (called the *topology* on X) such that:

1. $\emptyset$ and X belong to $\mathcal{V}$
2. if $\{V_i\}_{i \in I}$ is any family of elements of $\mathcal{V}$ then $\bigcup_{i \in I} v_i$ belongs to $\mathcal{V}$. i.e. $\mathcal{V}$ is closed under unions
3. if U and V belong to $\mathcal{V}$ then $U \cap V$ belongs to $\mathcal{V}$. i.e. $\mathcal{V}$ is closed under finite intersections.

We call the elements of $\mathcal{V}$ the *open sets* of the topological space X, or *open in* X.

Example Let X be a metric space. Then X is a topological space where we define a set V to be *open in* V if $V \subset X$ and each $a \in V \exists r > 0$ s.t.

$$B_X(a, r) \subset V$$



Theorem 4.4.1. *if X is a metric space then each $c \in X$ and $s > 0$ the ball $B_X(c, s)$ is open in X.*

Proof. Let $a \in B_X(c, s)$. Put $r = s - d(a, c) > 0$.

Then

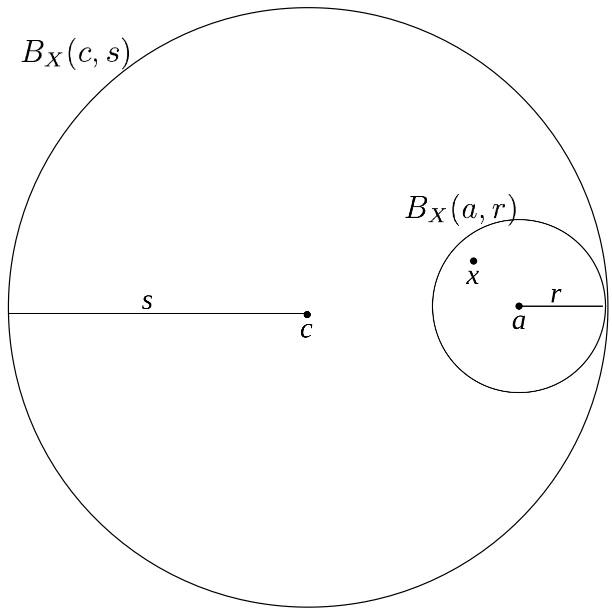
$$x \in B_X(a, r)$$

$$\implies d(x, a) < r = s - d(a, c)$$

$$\implies d(x, a) + d(a, c) < s$$

$$\implies d(x, c) < s$$

$$\implies x \in B_X(c, s)$$



and, therefore $BX(a,r) \subset B_X(c,s)$, as required. ☐

4.5 Continuous map of topological spaces

Definition a map $f : X \longrightarrow Y$ of topological spaces is called *continuous* if:

$$V \text{ open in } Y \implies f^{-1}V \text{ open in } X$$

Theorem 4.5.1. *let X, Y be metric spaces, then:*

$$f : X \longrightarrow Y$$

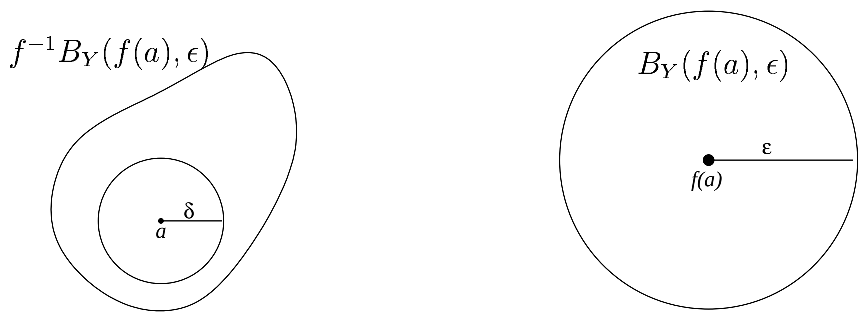
is continuous if and only if,

$$\begin{array}{c} \textit{for each } a \in X,\ \epsilon > 0 \\ \exists\, \delta > 0 \ s.t. \\ d(x,a) < \delta \implies d(f(x), f(a)) < \epsilon \end{array} \tag{4.2}$$

i.e.

$$fB_X(a,\delta) \subset B_Y(f(a),\epsilon)$$

Proof. 1. Let f be continuous. Let $a \in X$, $\epsilon > 0$. Then $B_Y(f(a),\epsilon)$ is open in Y. Therefore, $f^{-1}B_Y(f(a),\epsilon)$ is open in X.



$\exists \delta > 0$ s.t.

$$B_X(a,\delta) \subset f^{-1}B_Y(f(a),\epsilon)$$

and

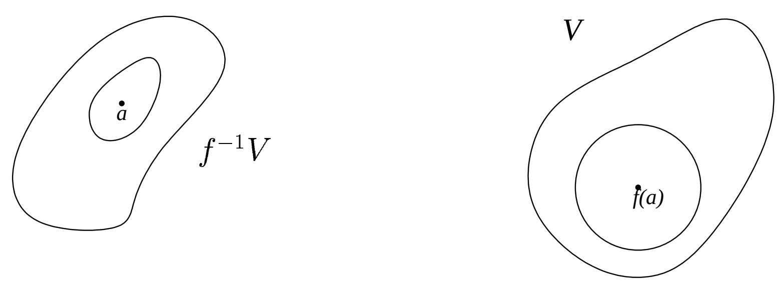
$$fB_X(a,\delta) \subset B_Y(f(a),\epsilon)$$

therefore Equation (4.2) holds.

2. Let Equation (4.2) hold. Let V be open in Y, $a \in f^{-1}V$, and so $f(a) \in V$.

   Now, $\exists \epsilon > 0$ s.t.

$$B_Y(f(a),\epsilon) \subset V$$



and $\exists \delta > 0$ s.t.

$$fB_X(a,\delta) \subset B_Y(f(a),\epsilon) \subset V$$

therefore,

$$B_X(a,\delta) \subset f^{-1}V$$

so, $f^{-1}V$ is open in X, and f is continuous.

□

4.6 Homeomorphisms

Definition a map $f : X \longrightarrow Y$ of topolgical spaces is called *homeomorphism* if:

1. f is bijective
2. f and f^{-1} are continuous

Definition X is *homeomorphic to* Y (*topologically equivalent*) if $\exists$ a homeomorphism $X \longrightarrow Y$. i.e. $\exists$ a beijective map under which the open sets of X correspond to the open sets of Y.

Definition a property P of a topological space is called a topological property if X has property P and X homeomorphic to $Y \implies Y$ has property P.

Example *'compactness'* is a topolgical property

Note: we have a category with topological spaces as objects, and continuous maps as morphisms.

4.7 Operator Norm

Definition Let M, N be finite dimensional normed spaces. Then we can make the vector space

$$\mathcal{L}(M, N)$$

of all linear operators $T : M \longrightarrow N$, into a normed space by defining:

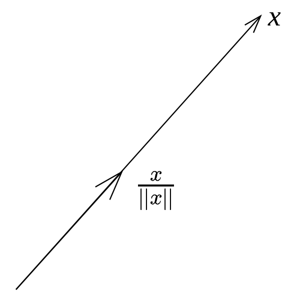
$$\|T\| = \sup_{\substack{x \in M \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}$$

$\|T\|$ is called the *operator norm* of T.

We have:

1. If $x \neq 0$ and $y = \frac{x}{\|x\|}$ then

$$\|y\| = \frac{1}{\|x\|}\|x\| = 1$$

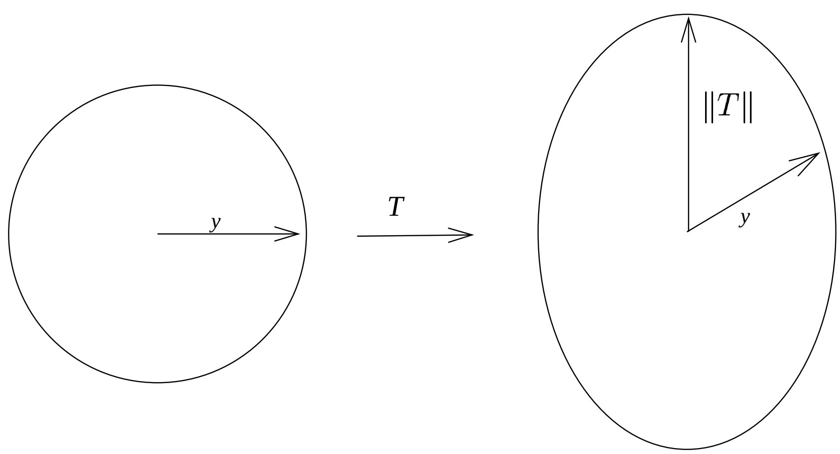


i.e. y has unit norm. Therefore

$$\frac{\|Tx\|}{\|x\|} = \left\| T\frac{x}{\|x\|} \right\| = \|Ty\|$$

and

$$\|T\| = \sup_{\substack{y \in M \\ ||y||=1}} \|Ty\|$$



2. if α is a scalar then

$$\|\alpha T\| = \sup_{||y||=1} \|\alpha Ty\| = |\alpha| \sup_{||y||=1} \|Ty\| = |\alpha| \|T\|$$

3. if $S, T : M \longrightarrow N$ then

$$\|S+T\| = \sup_{||y||=1} \|(S+T)y\| \leq \sup_{||y||=1} (\|Sy\|+\|Ty\|) \leq \sup \|Sy\|+\sup \|Ty\| = \|S\|+\|T\|$$

4.
$$\begin{aligned} \|T\| = 0 &\iff \sup_{x\neq 0} \tfrac{\|Tx\|}{\|x\|} = 0 \\ &\iff \|Tx\| = 0 \,\forall x \\ &\iff Tx = 0 \,\forall x \\ &\iff T = 0 \end{aligned}$$

(by 2,3,4 the operator norm **is** a norm).

Theorem 4.7.1. *1. if $T : M \longrightarrow N$ and $x \in M$ then*

$$\|Tx\| \leq \|T\| \, \|x\|$$

2. If $L \xrightarrow{T} M \xrightarrow{S} N$ then

$$\|ST\| \leq \|S\| \, \|T\|$$

Proof. 1.

$$\frac{\|Tx\|}{\|x\|} \leq \|T\|$$

$\forall x \neq 0$, be definition. therefore,

$$\|Tx\| \leq \|T\| \, \|x\|$$

for all x

2.
$$\begin{aligned} \|ST\| &= \sup_{\|y\|=1} \|STy\| \\ &\leq \sup \|S\| \, \|Ty\| && \text{(by 1.)} \\ &\leq \sup_{\|y\|=1} \|S\| \, \|T\| \, \|y\| && \text{(by 1.)} \\ &= \|S\| \, \|T\| \end{aligned}$$

∎

Note:

1. if M is finite dimensional then every choice of norm on M defines the *same* topology on M
2. a sequence x_r of points in a topological space X is said to *converge to* $a \in M$ if, for each open set V containing $a \,\exists N$ s.t. $x_r \in V \,\forall\, r \geq N$.

   For a normed space M this is the same as: for each $\epsilon > 0 \,\exists N$ s.t. $\|x_r - a\| < \epsilon \,\forall\, r \geq N$.

4.8 Differentiation

Definition let

$$M \supset V \xrightarrow{f} N$$

where M, N are normed spaces and V is open in M. Let $a \in V$. THen f is *differentiable at* a if $\exists$ a linear operator

$$M \xrightarrow{f'(a)} N$$

such that

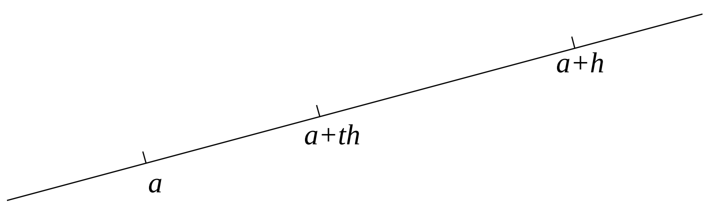
$$f(a+h) = f(a0) + f'(a)h + \phi(h)$$

where

$$\frac{\|\phi(h)\|}{\|h\|} \longrightarrow 0 \quad \text{as } \|h\| \longrightarrow 0$$

Theorem 4.8.1. *if f is differentiable at a the the operator $f'(a)$ is uniquely determined by:*

$$\begin{aligned} f'(a)h &= \lim_{t \longrightarrow 0} \tfrac{f(a+th)-f(a)}{t} \\ &= \textit{directional derivative of } f \textit{ at } a \textit{ along } h \\ &= \tfrac{d}{dt} f(a+th)\big|_{t=0} \end{aligned}$$



$f'(a)$ is called the derivative of f at a

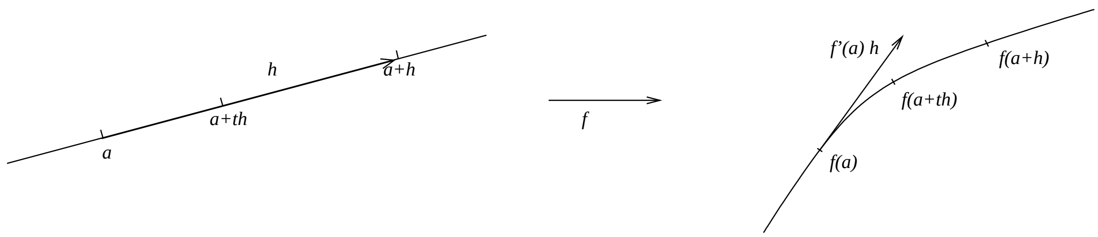
Proof.

$$f(a+th) = f(a) + f'(a)th + \phi(th)$$

therefore,

$$\left\| \frac{f(a+th)-f(a)}{t} - f'(a)h \right\| = \left\| \frac{\phi(th)}{t} \right\| = \frac{\|\phi(th)\|}{\|th\|}\|h\|$$

tends to 0 as $t \longrightarrow 0$.



□

Example 1. $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x^3$. Then

$$f(a+h) = (a+h)^3 = \underbrace{a^3}_{f(a)} + \underbrace{3a^2h}_{f'(a)h} + \underbrace{3ah^2 + h^3}_{\phi(h)}$$

so $f'(a) = 3a^2$.

2. $f : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}^{n\times n}$, $f(X) = X^3$.

$$\begin{aligned} f(A+H) &= (A+H)^3 \\ &= \underbrace{A^3}_{f(A)} + \underbrace{A^2H + AHA + HA^2}_{f'(A)H} \\ &\quad + \underbrace{AH^2 + HAH + H^2A + H^3}_{\phi(H)} \end{aligned}$$

taking (say) Euclidean norm on $\mathbb{R}^n$ and operator norm on $\mathbb{R}^{n\times n} = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.

$$\begin{aligned} \frac{||\phi(H)||}{||H||} &= \frac{||AH^2+HAH+H^2A+H^3||}{||H||} \\ &\leq \frac{3||A||\,||H||^2}{||H||} \\ &= 3\|A\|\,\|H\| \longrightarrow 0 \\ &\quad \text{as } \|H\| \longrightarrow 0 \end{aligned}$$

therefore, f is differentiable at A, and

$$f'(A) : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}^{n\times n}$$

is given by

$$f'(A)H = A^2H + AHA + HA^2$$

Theorem 4.8.2. *Let* $\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}^m$ *be a differentiable function with*

$$f = (f^1, \dots, f^m) = (f^i)$$

Then

$$f' = \left(\frac{\partial f^i}{\partial x^j}\right)$$

Proof. For each $a \in V$ we have

$$f'(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

is a linear operator. Therefore, $f'(a)$ is an $m \times n$ matrix whose j^{th} column is

$$\begin{aligned} f'(a)e_j &= \lim_{t \longrightarrow 0} \tfrac{f(a+te_j)-f(a)}{t} \\ &= \tfrac{\partial f}{\partial x^j}(a) \\ &= \left(\tfrac{\partial f^1}{\partial x^j}(a), \dots, \tfrac{\partial f^m}{\partial x^j}(a)\right) \end{aligned}$$

therefore,

$$f'(a) = \left(\frac{\partial f^i}{\partial x^j}(a)\right)$$

for all $a \in V$. Therefore,

$$f' = \left(\frac{\partial f^i}{\partial x^j}\right)$$

which is the *Jacobian matrix.* □

Theorem 4.8.3. *Let* $M \supset V \xrightarrow{f} N_1 \times \cdots \times N_k$ *where*

$$f(x) = \left(f^1(x), \dots, f^k(x)\right)$$

Then $f^1, \dots, f^k$ *differentiable at* $a \in V \implies f$ *is differentiable at* a*, and* $f'(a)h = \left({f^1}'(a)h, \dots, {f^k}'(a)h\right)$.

Proof. $(k = 2)$

$$M \supset V \xrightarrow{f} N_1 \times N_2$$

$$f(x) = \left(f^1(x), f^2(x)\right)$$

Take any norms on M, N_1, N_2 and define a norm on $N_! \times N_2$ by

$$\|(y_1, y_2)\| = \|y_1\| + \|y_2\|$$

Then

$$\begin{aligned} f(a+h) &= (f^1(a+h), f^2(a+h)) \\ &= \left(f^1(a) + f^{1'}(a)h + \phi^1(h), f^2(a) + f^{2'}(a)h + \phi^2(h)\right) \\ &= (f^1(a), f^2(a)) + \underbrace{\left(f^{1'}(a)h, f^{2'}(a)h\right)}_{\text{linear in } h} + \underbrace{\left(\phi^1(h), \phi^2(h)\right)}_{\text{remainder}} \end{aligned}$$

Now:

$$\frac{\| (\phi^1(h), \phi^2(h)) \|}{\|h\|} = \frac{\|\phi^1(h)\|}{\|h\|} + \frac{\|\phi^2(h)\|}{\|h\|}$$

tends to 0 as $h \longrightarrow 0$. Therefore, f is differentiable at am and

$$f'(a)h = \left(f^{1'}(a)h, f^{2'}(a)h\right)$$

as required. □

4.9 Notation

1. Given a function of n variables

$$\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$$

   V open, we shall denote by

$$x^1, x^2, \ldots, x^n$$

   the usual co-ordinate functions on $\mathbb{R}^n$ and shall often denote the *partial derivative of f at a w.r.t. j^{th} variable* by:

$$\begin{aligned} \frac{\partial f}{\partial x^j}(a) &= \lim_{t \longrightarrow 0} \frac{f(a_1, a_2, \ldots, a_j + t, \ldots, a_n) - f(a_1, a_2, \ldots, a_j, \ldots, a_n)}{t} \\ &= \frac{d}{dt} f(a + te_j)\Big|_{t=0} \end{aligned}$$

   Notice that the symbol x^j does not appear in the definition: it is a *'dummy symbol'* indicating deriv w.r.t. j^{th} 'slot', sometimes written as $f_{,j}(a)$.

2. given a function

$$\mathbb{R}^2 \supset V \xrightarrow{f} \mathbb{R}$$

V open, we often denote by x, y the usual co-ordinate functions instead of x^1, x^2 and write

$$\frac{\partial f}{\partial x} \text{ for } \frac{\partial f}{\partial x^1}$$

$$\frac{\partial f}{\partial y} \text{ for } \frac{\partial f}{\partial x^2}$$

etc.

3. given

$$M \supset V \xrightarrow{f} N$$

V open, if f is differentiable at a, $\forall a \in V$, we say that f *is differentiable on* V and call the function on V:

$$f' : a \longrightarrow f'(a)$$

the *derivative of f.* We write:

$$f^{(2)} = (f')'$$

$$f^{(3)} = ((f')')'$$

and call f C^r if $f^{(r)}$ exists and is continuous, f C^∞ if $f^{(r)}$ exists $\forall r$.

4.10 C^r Functions

Theorem 4.10.1. *Let*

$$\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$$

V open. Then f is C^1 $\iff$ $\frac{\partial f}{\partial x^i}$ exists and is continuous for $i = 1, \ldots, n$.

Proof. (case $n = 2$)

$$\mathbb{R}^2 \supset V \xrightarrow{f} \mathbb{R}, \qquad f(x, y)$$

1. if f is C^1, then $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ exist and

$$f' = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

therefore, f' is continuous, and therefore, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous

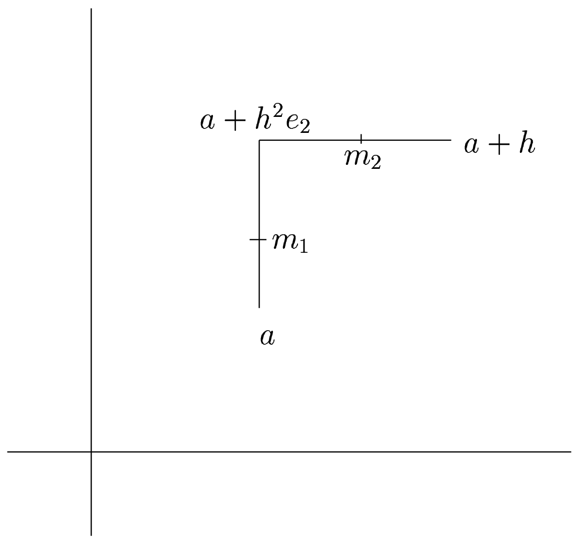
2. Let $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ exists and be continuous on V. Then

$$f(a+h) = f(a) + \frac{\partial f}{\partial x}(a)h^1 + \frac{\partial f}{\partial y}(a)h^2 + \phi(h)$$

(say) where $h = (h^1, h^2)$. Now

$$\begin{aligned}\phi(h) &= f(a+h) - f(a+h^2e_2) - \tfrac{\partial f}{\partial x}(a)h^1 \\ &\quad + f(a+h^2e_2) - f(a) - \tfrac{\partial f}{\partial y}(a)h^2 \\ &= \tfrac{\partial f}{\partial x}(m_1)h^1 - \tfrac{\partial f}{\partial x}(a)h^1 \\ &\quad + \tfrac{\partial f}{\partial y}(m_2)h^2 - \tfrac{\partial f}{\partial x}(a)h^2\end{aligned}$$

(say) by mean value theorem.



Therefore,

$$\frac{\|\phi(h)\|}{\|h\|} \le \left|\frac{\partial f}{\partial x}(m_1) - \frac{\partial f}{\partial x}(a)\right| \frac{|h^1|}{\|h\|} + \left|\frac{\partial f}{\partial y}(m_2) - \frac{\partial f}{\partial y}(a)\right| \frac{|h^2|}{\|h\|} \longrightarrow 0$$

as $h \longrightarrow 0$ since $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ are continuous at a.

Therefore, f is differentiable at a and $f'(a) = \left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a)\right)$. Also, $f' = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ is continuous.

□

Corollary 4.10.2. *f is C^r $\iff$ $\frac{\partial^r f}{\partial x^{i_1} \cdots \partial x^{i_r}}$ exists and is continuous for each $i_1, \ldots, i_r$.*

Theorem 4.10.3. *if*

$$\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R} \qquad V \text{ open}$$

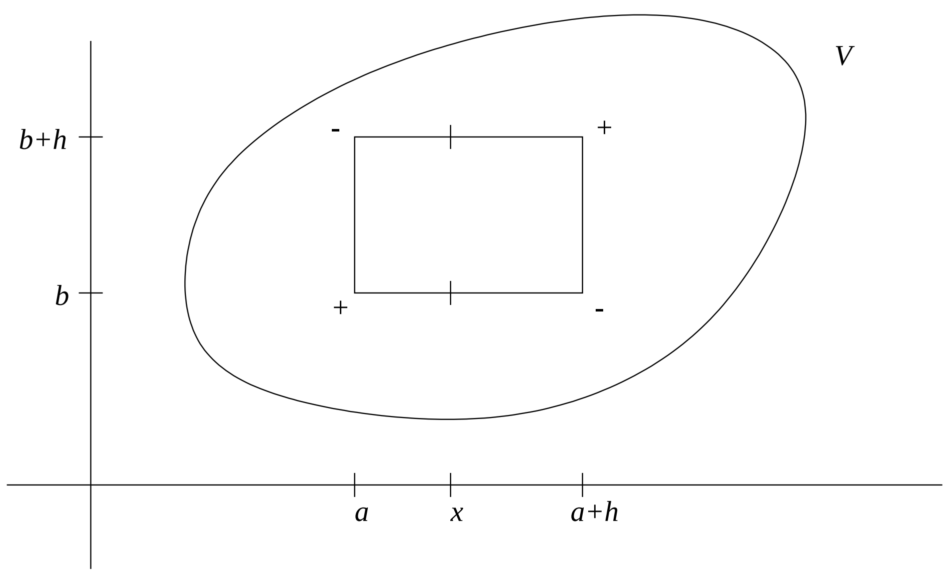
is C^2 then $\frac{\partial^2 f}{\partial x^i \partial x^j}$ is a symmetric matrix.

Proof. Let

$$\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$$

be C^2. Need to show:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$



Let $(a, b) \in V$. Let $h \neq 0$, $k \neq 0$ be such that the closed rectangle

$$(a, b),\ (a + h, b),\ (a + h, b + k),\ (a, b + k)$$

is contained in V. Put

$$g(x) = f(x, b + k) - f(x, b)$$

Then

$$\begin{aligned} f(a+h,b+k)-f(a+h,b) \\ -f(a,b+k)+f(a,b) &= g(a+h)-g(a) \\ &= hg'(c) \qquad \text{some } a \le c \le a+h \text{ by MVT} \\ &= h\left[\tfrac{\partial f}{\partial x}(c,b+k)-\tfrac{\partial f}{\partial x}(c,b)\right] \\ &= hk\tfrac{\partial^2 f}{\partial y \partial x}(c,d) \qquad \text{some } b \le d \le b+k \text{ by MVT} \end{aligned}$$

and, similarly:

$$\begin{aligned} f(a+h,b+k)-f(a+h,b) \\ -f(a,b+k)+f(a,b) &= kh\tfrac{\partial^2 f}{\partial x \partial y}(c',di') \end{aligned}$$

for some $a \le c' \le a+h$ and $b \le d' \le b+k$. Therefore

$$\frac{\partial^2 f}{\partial y \partial x}(c,d) = \frac{\partial^2 f}{\partial x \partial y}(c',d')$$

Now, let $(h,k) \longrightarrow (0,0)$. Then $(c,d) \longrightarrow (a,b)$, and $(c',d') \longrightarrow (a,b)$. Therefore,

$$\frac{\partial^2 f}{\partial y \partial x}(a,b) = \frac{\partial^2 f}{\partial x \partial y}(a,b)$$

by continuity of $\frac{\partial^2 f}{\partial y \partial x}$ and $\frac{\partial^2 f}{\partial x \partial y}$ □

4.11 Chain Rule

Theorem 4.11.1. (Chain Rule for functions on finite dimensional real or complex vector spaces)

Let L, M, N be finite dimensional real or complex vector spaces. Let U be open in L, V open in M and let

$$U \xrightarrow{g} V \xrightarrow{f} N$$

Let g be differentiable at $a \in U$, f be differentiable at $g(a)$. Then the composition $f \cdot g$ is differentiable at a and

$$(f \cdot g)'(a) = f'(g(a))\, g'(a) \qquad \textit{operator product}$$

Proof. we have

$$
\begin{array}{rcll}
f[g(a_h)] & = & f\left[g(a) + Th + \phi(h)\right] & \text{where } T = g'(a) \\
& & & \text{and } \frac{\|\phi(h)\|}{\|h\|} \longrightarrow 0 \\
& & & \text{as } \|h\| \longrightarrow 0 \\
& = & f[g(a) + y] & \text{where } y = Th + \phi(h) \\
& & & \text{so } \|y\| \leq \|T\| \, \|h\| + \|\phi(h)\| \longrightarrow 0 \\
& & & \text{as } \|h\| \longrightarrow 0 \\
& = & f(g(a)) + Sy + \|y\|\psi(y) & \text{where } S = f'(g(a)) \\
& & & \text{and } \|\psi(h)\| \longrightarrow 0 \\
& & & \text{as } \|y\| \longrightarrow 0 \\
& = & f(g(a) + S(Th + \phi(h) + \|y\|\psi(y) & \\
& = & f(g(a)) + STh + S\phi(h) + \|y\|\psi(y) &
\end{array}
$$

Now

$$\frac{\|S\phi(h) + \|y\|\psi(y)\|}{\|h\|} \leq \|S\| \, \frac{\|\phi(h)\|}{\|h\|} + \left(\|T\| + \frac{\|\phi(h)\|}{\|h\|}\right) \|\psi(y)\|$$

tends to 0 as $\|h\| \longrightarrow 0$. Therefore, $f \cdot g$ is differentiable at a, and

$$(f \cdot g)'(a) = ST = f'(g(a)) \, g'(a)$$

□

Example

$$\frac{d}{dt} f(g^1(t), \ldots, g^n(t)) = (f \cdot g)'(t) = \underbrace{f'(g(t))}_{1\times n} \, \underbrace{g'(t)}_{n\times 1}$$

where

$$\mathbb{R} \supset U \xrightarrow{g} V \xrightarrow{f} \mathbb{R}$$

V open in $\mathbb{R}^n$, and f, g differentiable. Then

$$
\begin{array}{rcl}
f'(g(t))g'(t) & = & \left(\frac{\partial f}{\partial x^1}(g(t)), \ldots \frac{\partial f}{\partial x^n}(g(t))\right) \begin{pmatrix} \frac{d}{dt} g^1(t) \\ \vdots \\ \frac{d}{dt} g^n(t) \end{pmatrix} \\
& = & \frac{\partial f}{\partial x^1}(g(t)) \frac{d}{dt} g^1(t) + \cdots + \frac{\partial f}{\partial x^n}(g(t)) \frac{d}{dt} g^n(t)
\end{array}
$$

is the usual chain rule.

Chapter 5

Calculus of Complex Numbers

5.1 Complex Differentiation

Definition let

$$\mathbb{C} \supset V \xrightarrow{f} \mathbb{C}$$

be a complex valued function of a complex variable defined on open V. $\mathbb{C}$ is a 1-dimensional complex normed space.

Let $a \in V$. Then f is differentiable at a as a function of a complex variable if $\exists$ a linear map of complex spaces:

$$\mathbb{C} \xrightarrow{f'(a)} \mathbb{C}$$

s.t.

$$f(a+h) = f(a) + f'(a)h + \phi(h)$$

where $\frac{|\phi(h)|}{|h|} \longrightarrow 0$ as $h \longrightarrow 0$. $f'(a)$ is a 1×1 complex matrix, i.e. a complex number, and

$$\left| \frac{f(a+h) - f(a)}{h} - f'(a) \right| = \frac{|\phi(h)|}{|h|}$$

which tends to 0 as $h \longrightarrow 0$. Therefore

$$f'(a) = \lim_{h \longrightarrow 0} \frac{f(a+h) - f(a)}{h}$$

We call $f'(a)$ the *derivative of* f *at* a. f' is called *holomorphic* on V if $f'(a)$ exists $\forall\, a \in V$. We write $f = \frac{df}{dy}$.

Example

$$f : \mathbb{C} \longrightarrow \mathbb{C}, \qquad f(z) = z^n$$

then,

$$\begin{aligned}\lim_{h\longrightarrow 0}\tfrac{f(z+h)-f(z)}{h} &= \lim_{h\longrightarrow 0}\tfrac{(z+h)^n-z^n}{h}\\ &= \lim_{h\longrightarrow 0}[nz^{n-1}+\tfrac{n(n-1)}{2}z^{n-2}h+\text{ higher powers } h]\\ &= nz^{n-1}\end{aligned}$$

therefore, f is holomorphic in $\mathbb{C}$ and $f'(z)=nz^{n-1}$.

Example

$$f(z)=\overline{z}$$

then,

$$\frac{f(z+h)-f(z)}{h}=\frac{\overline{h}}{h}=\begin{cases}1 & h \text{ real}\\ -1 & h \text{ pure imaginary}\end{cases}$$

which does not converge as $h\longrightarrow 0$.

The usual rules for differentiation apply:

1. $\frac{d}{dy}(f+g)=\frac{df}{dy}+\frac{dg}{dy}$
2. $\frac{d}{dy}fg=f\frac{dg}{dy}+g\frac{df}{dy}$
3. $\frac{d}{dy}\frac{f}{g}=\frac{g\frac{df}{dy}-f\frac{dg}{dy}}{g^2}$ if $g\neq 0$
4. $\frac{d}{dz}f(g(z))=f'(g(z))\,g'(z)$ chain rule

By definition $\mathbb{C}=\mathbb{R}^2$

$$z=x+iy=(x,y)$$

and the operation of mult by $i=\sqrt{-1}$ $\mathbb{C}\xrightarrow{i}\mathbb{C}$ is the operator $\mathbb{R}^2\xrightarrow{i}\mathbb{R}^2$ with matrix $\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$, because

$$ie_1=i=(0,1)=0e_1+e_2$$

$$ie_2=ii=-1=-e_1+0e_2=(-1,0)$$

Let f is (real) differentiable on V then each $a\in v$ the operator

$$\mathbb{C}\xrightarrow{f'(a)}\mathbb{C}\qquad\text{i.e}\qquad\mathbb{R}^2\xrightarrow{f'(a)}\mathbb{R}^2$$

preserves addition and commutes with mult by real scalars, for each $a \in V$, and f is holomorphic in V iff $f'(a)$ also commutes with multiplication by complex scalars. $\iff$ $f'(a)$ also commutes with multiplication by i.

$$\iff \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

$$\iff \begin{pmatrix} \frac{\partial u}{\partial y} & -\frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial y} & -\frac{\partial v}{\partial x} \end{pmatrix} = \begin{pmatrix} -\frac{\partial v}{\partial x} & -\frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \end{pmatrix}$$

$$\iff \left.\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial v}{\partial x} &= -\frac{\partial u}{\partial y} \end{aligned}\right\} \quad \text{Cauchy-Riemann Equations}$$

so, therefore: $f = u + iv$ is holomorphic in V $\iff$ f is real-differentiable on V and satisfies the Cauchy-Riemann Equations.

Example $f(z) = z^2 = (x+iy)^2 = x^2 - y^2 + 2ixy$. then

$$u = x^2 - y^2 \quad \frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y}$$

$$v = 2xy \qquad \frac{\partial v}{\partial x} = 2y = -\frac{\partial u}{\partial x}$$

Note, if $f = u + iv$ is holomorphic in V then

$$\begin{aligned} \frac{\partial f}{\partial y}(a) &= \lim_{h \longrightarrow 0} \frac{f(a+h)-f(a)}{h} \\ &= \lim_{t \longrightarrow 0} \frac{f(a+t)-f(a)}{t} = \lim_{t \longrightarrow 0} \frac{f(a+it)-f(a)}{it} \\ &= \frac{\partial}{\partial x}[u+iv] = \frac{1}{i}\frac{\partial}{\partial y}[u+iv] \end{aligned}$$

therefore,

$$\frac{\partial f}{\partial z} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}$$

this also gives the Cauchy-Riemann Equations.

5.2 Path Integrals

Definition C^1 map

$$[t_1, t_2] \xrightarrow{\alpha} V \subset \mathbb{R}^2 \qquad V \text{ open}$$

$$t \longrightarrow \alpha(t)$$

is called a (parametrical) *path* in $\mathbb{R}$ fom $\alpha(t_1)$ to $\alpha(t_2)$ and if f, g are complex-valued continuous on V we write

$$\int_\alpha (f\,dx + g\,dy) = \int_{t_1}^{t_2} [f(\alpha(t))\frac{d}{dt}x(\alpha(t)) + g(\alpha(t))\frac{d}{dt}y(\alpha(t))]\,dt$$

and call it the *integral* of the differential form $f\,dx + g\,dy$ over the path α. If

$$[s_1, s_2] \xrightarrow{\sigma} [t_1, t_2]$$

is C^1 with $\sigma(s_1) = t_1$, $\sigma(s_2) = t_2$ and $\beta(s) = \alpha(\sigma(s))$ then β is called a *reparameterisation* of α.

We have

$$\begin{aligned}
\int_\beta (f\,dx + g\,dy) &= \int_{s_1}^{s_2} [f(\beta(s))\tfrac{d}{ds}x(\beta(s)) + g(\beta(s))\tfrac{d}{ds}y(\beta(s))]\,ds \\
&= \int_{s_1}^{s_2} [f(\alpha(\sigma(s)))\tfrac{d}{ds}x(\alpha(\sigma(s))) + g(\alpha(\sigma(s)))\tfrac{d}{ds}y(\alpha(\sigma(s)))]\,ds \\
&= \int_{s_1}^{s_2} [f(\alpha(\sigma(s)))(x\cdot\alpha)'(\sigma(s)) + g(\alpha(\sigma(s)))(y\cdot\alpha)'(\sigma(s))]\sigma'(s)\,ds \\
&= \int_{t_1}^{t_2} [f(\alpha(t))\tfrac{d}{dt}x(\alpha(t)) + g(\alpha(t))\tfrac{d}{dt}y(\alpha(t))]\,dt \\
&= \int_\alpha (f\,dx + g\,dy)
\end{aligned}$$

therefore, the integral over α is independent of parmaterisation.

If we take

$$[t_1, t_2] \xrightarrow{\sigma} [t_1, t_2]$$

with $\sigma(s) = t_1 + t_2 - s$, and $\beta(s) = \alpha(\sigma(s))$ we have $\sigma(t_2) = t_1$, $\sigma(t_1) = t_2$ then β is same path but traversed in the opposite direction and

$$\begin{aligned}
\int_\beta (f\,dx + g\,dy) &= \int_{t_2}^{t_1} [f(\alpha(t))\tfrac{d}{dt}x(\alpha(t)) + g(\alpha(t))\tfrac{d}{dt}y(\alpha(t))]\,ds \\
&= -\int_\alpha (f\,dx + g\,dy)
\end{aligned}$$

Definition If f is C^1 on V we write

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

A differential form of this type is called *exact*, and is called the *differential* of f.

If α is a path from a to b, $\alpha(t_1) = a$, $\alpha(t_2) = b$, then

$$\begin{aligned}
\int_\alpha df &= \int_{t_1}^{t_2} [f(\alpha(t))\tfrac{d}{dt}x(\alpha(t)) + g(\alpha(t))\tfrac{d}{dt}y(\alpha(t))]\,dt \\
&= \int_{t_1}^{t_2} [\tfrac{d}{dt}f(\alpha(t))]\,dt \\
&= f(\alpha(t_2)) - f(\alpha(t_1)) \\
&= f(b) - f(a) \qquad \text{so we have path independence} \\
&= \text{change of value of f along } \alpha
\end{aligned}$$

If α is a *closed path*, (i.e. $a = b$) then

$$\int_\alpha df = 0$$

If $f(z) = u(x, y) + iv(x, y)$, for $z = x + iy$ with u, v real, put $dz = dx + idy$ and

$$\begin{aligned}
f(z)\,dz &= [u + iv][dx + idy] \\
&= (u\,dx - v\,dy) + i(v\,dx + u\,dy)
\end{aligned}$$

then

$$\begin{aligned}
\int_\alpha f(z)\,dz &= \int_\alpha (u\,dx - v\,dy) + i\int_\alpha (v\,dx + u\,dy) \\
&= \int_{t_1}^{t_2} \left\{[u(\alpha(t)) + iv(\alpha(t))]\tfrac{d}{dt}[x(\alpha(t)) + iy(\alpha(t))]\right\}\,dt \\
&= \int_{t_1}^{t_2} f(\alpha(t))\,\alpha'(t)\,dt
\end{aligned}$$

If f is holomorphic then

$$\begin{aligned}
df &= du + i\,dv = \tfrac{\partial u}{\partial x}dx + \tfrac{\partial u}{\partial y}dy + i\,\tfrac{\partial v}{\partial x}dx + i\,\tfrac{\partial v}{\partial y}dy \\
&= \tfrac{\partial u}{\partial x}dx - \tfrac{\partial v}{\partial x}dx + i\,\tfrac{\partial v}{\partial x}dx + i\,\tfrac{\partial u}{\partial x}dy \\
&= \left(\tfrac{\partial u}{\partial x} + i\,\tfrac{\partial v}{\partial x}\right)(\,dx + i\,dy) \\
&= f'(z)\,dz
\end{aligned}$$

therefore, if f is holomorphic on V open then

1. for any path α in V from a to b:

$$\int_\alpha f'(z)\,dz = f(b) - f(a)$$

   path independence

2. for any closed path α in V:

$$\int_\alpha f'(z)\,dz = 0$$

Example 1. if α path from a to b

$$\int_\alpha z^n\,dz = \int_\alpha [\frac{d}{dz}\frac{z^{n+1}}{n+1}]\,dz = \frac{b^{n+1}-a^{n+1}}{n+1} \qquad n \neq 1$$

in particular,

2. if α is a closed path,

$$\int_\alpha z^n\,dz = 0 \qquad n \neq 1$$

But:

3. for the closed path $\alpha(t) = e^{it}$, $0 \leq t \leq 2\pi$

$$\int_\alpha \frac{dz}{z} = \int_0^{2\pi} \frac{1}{e^{it}} ie^{it}\,dt = \int_0^{2\pi} i\,dt = 2\pi i \neq 0$$

therefore, $\frac{1}{z}$ cannot be the derivative of a holomorphic function on $\mathbb{C}$.

Definition if $[t_1, t_2] \xrightarrow{\alpha} \mathbb{C}$ is a path in $\mathbb{C}$ then we write

$$L(\alpha) = \int_{t_1}^{t_2} |\alpha'(t)|\,dt$$

and call it the *length of* α.

If $\beta = \alpha \cdot \sigma$ is a reparameterisation of α with $\alpha'(s) > 0$

$$\begin{aligned} L(\beta) &= \int_{s_1}^{s_2} |\beta'(s)|\,ds \\ &= \int_{s_1}^{s_2} |\alpha'(\sigma(s))\sigma'(s)|\,ds \\ &= \int_{t_1}^{t_2} |\alpha'(t)|\,dt \\ &= L(\alpha) \end{aligned}$$

Theorem 5.2.1. (Estimating a complex integral)
Let f be bounded on α:

$$|f(\alpha(t))| \leq M$$

(say), $t_1 \leq t \leq t_2$. then

$$\left|\int_\alpha f(z)\,dz\right| \leq ML(\alpha)$$

Proof.

$$\begin{aligned} \left|\int_\alpha f(z)\,dz\right| &= \left|\int_{t_1}^{t_2} f(\alpha(t))\alpha'(t)\,dt\right| \\ &\leq \int_{t_1}^{t_2} |f(\alpha(t))|\,|\alpha'(t)|\,dt \\ &\leq M\int_{t_1}^{t_2} |\alpha'(t)|\,dt \\ &= ML(\alpha) \end{aligned}$$

□

5.3 Cauchy's Theorem for a triangle

If $f(z) = u(x,y) + i\,v(x,y)$ is holomorphic then

$$\int_\alpha f(z)\,dz = \int_\alpha (u\,dx - v\,dy) + i\int (v\,dx + u\,dy)$$

and $\frac{\partial}{\partial y}u = \frac{\partial}{\partial x}(-v)$, $\frac{\partial}{\partial y}v = \frac{\partial}{\partial x}u$ (by Cauchy-Riemann).
Therefore, the neccessary conditions for path-independence are satisfied, but not always sufficient, e.g.$f(z) = \frac{1}{z}$. However, we have

Theorem 5.3.1. (Cauchy's Theorem for a triangle)
Let $\mathbb{C} \supset V \xrightarrow{f} \mathbb{C}$ be holomorphic on open V, and let T be a triangle (interior plus boundary δT) $\subset V$. Then

$$\int_{\delta T} f(z)\,dz = 0$$

Proof. write

$$i(T) = \int_{\delta t} f(z)\,dz$$

and join the mid-points of the sides to get 4 triangles

$$S_1, S_2, S_3, S_4$$

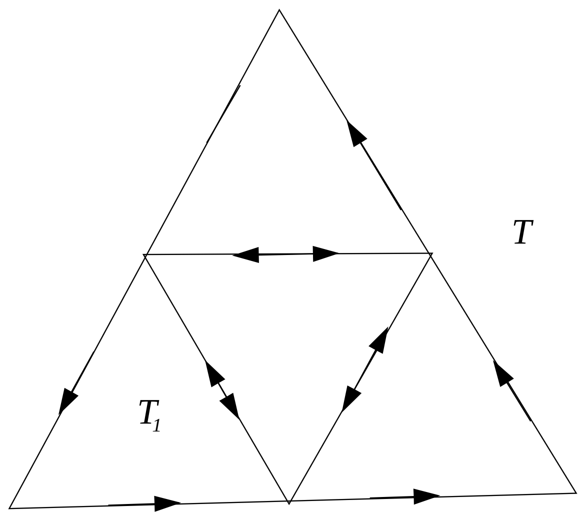
then

$$i(T) = \sum_{j=1}^{4} i(S_j)$$

and therefore

$$|i(T)| \leq \sum_{j=1}^{4} |i(S_j)| \leq 4|i(T_1)|$$

(say) where T_1 is one of $S_1, \ldots, S_4$.



Repeat the process to get a sequence of triangles

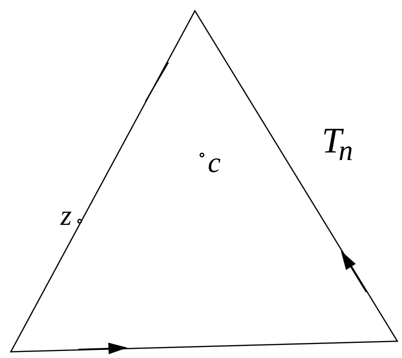
$$T_1, T_2, \ldots, T_n, \ldots$$

with

$$|i(T)| \leq 4^n |i(T_n)|$$

Let $\bigcap_{j=1}^{\infty} T_j = \{c\}$. Then

$$\begin{aligned} i(T_n) &= \int_{\delta T_n} f(z)\,dz \\ &= \int_{\delta T_n} [f(c) + f'(c)(z-c) + |z-c|\phi(z-c)]\,dz \\ &= \int_{\delta T_n} |z-c|\,\phi(z-c)\,dz \qquad \text{where } |\phi(z-c)| \longrightarrow 0 \text{ as } |z-c| \longrightarrow 0 \end{aligned}$$



Let $\epsilon > 0$. Choose $\delta > 0$ s.t. $|\phi(z-c)| < \epsilon \, \forall \, |z-c| < \delta$. Let L =length of T. Choose n s.t. length of $T_n = \frac{l}{"n} < \delta$.

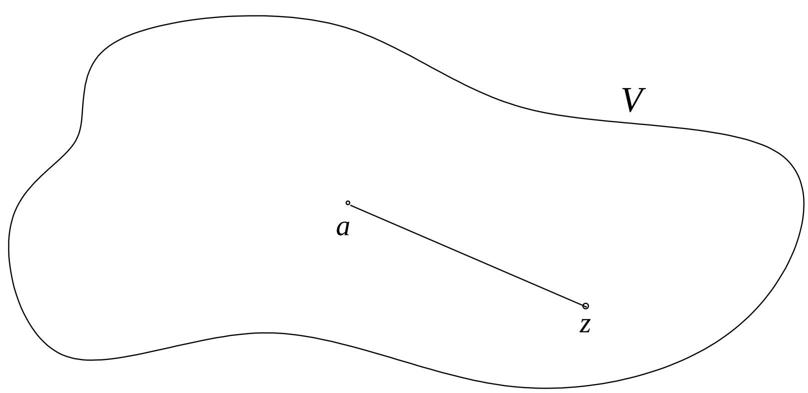
Then

$$|i(T_n)| \leq \frac{l}{2^n}\epsilon\frac{l}{2^n} = \frac{l^2}{4^n}\epsilon$$

therefore $|i(T)| \leq l^2\epsilon$, and therefore $i(T) = 0$. □

Definition A set $V \subset \mathbb{C}$ is called *star-shaped* if $\exists a \in V$ s.t.

$$[a, z] \subset V \qquad \forall \, z \in V$$

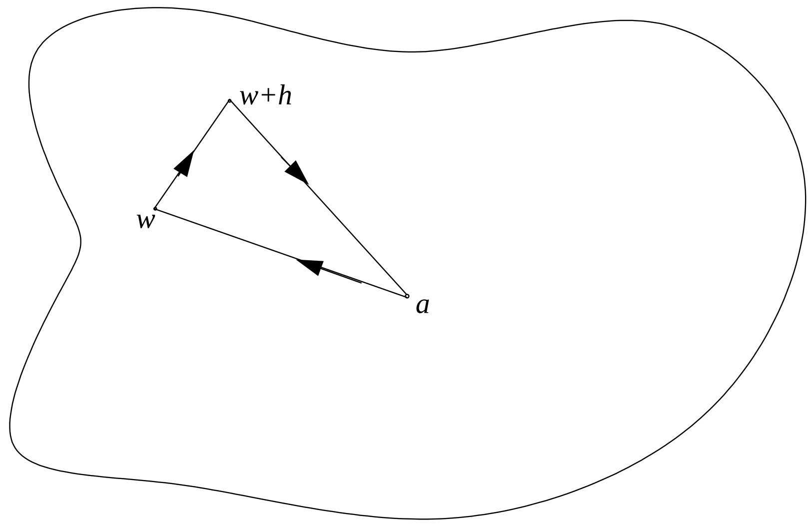


Theorem 5.3.2. *Let* $\mathbb{C} \supset V \xrightarrow{f} \mathbb{C}$ *be a holomorphic function on an open star-shaped set. Then* $f = F'$ *for some complex-differentable function* F *on* V *and hence:*

$$\int_\alpha f(z)\, dz = 0$$

for each closed curve α *in* V.

Proof. Choose $a \in V$ s.t. $[a, z] \subset V \, \forall \, z \in V$. Put $F(w) = \int_a^w f(z)\, dz$. Then



$$
\begin{aligned}
F(w+h) &= \textstyle\int_a^{w+h} f(z)\,dz \\
&= \textstyle\int_a^{w} f(z)\,dz + \int_w^{w+h} f(z)\,dz \qquad \text{by Cauchy} \\
&= F(w) + f(w)h + \underbrace{\int_w^{w+h} [f(z) - f(w)]\,dy}_{\phi(h)}
\end{aligned}
$$

Let $\epsilon > 0$. Choose $\delta > 0$ w.t. $|f(z) - f(w)| < \epsilon \ \ \forall |z - w| < \delta$. Therefore

$$|\phi(h)| \leq |h|\,\epsilon$$

for all $|h| < \delta$, and therefore

$$\lim_{h \longrightarrow 0} \frac{|\phi(h)|}{|h|} = 0$$

Therefore, F is differentiable at w and $F'(w) = f(w)$ as required.

□

5.4 Winding Number

We have seen that

$$\int_{\text{circle about } o} \frac{dz}{z} = 2\pi i$$

More generally:

Theorem 5.4.1. *Let* $a \in \mathbb{C}$ *ad* $\alpha : [t_1, t_2] \longrightarrow \mathbb{C} - \{a\}$ *be a closed path. The*

$$\frac{1}{2\pi i} \int_\alpha \frac{dz}{z - a}$$

is an integer, called the winding number of α about a

Proof. put

$$\beta(t) = \int_{t_1}^{t} \frac{\alpha'(s)}{\alpha(s) - a}\, ds$$

so $\beta'(t) = \frac{\alpha'(t)}{\alpha(t)-a}$. Then

$$\begin{aligned} \tfrac{d}{dt}\left[[\alpha(t) - a]e^{-\beta(t)}\right] &= [\alpha'(t) - \beta'(t)\{\alpha(t) - a\}]\, e^{-\beta(t)} \\ &= 0 \end{aligned}$$

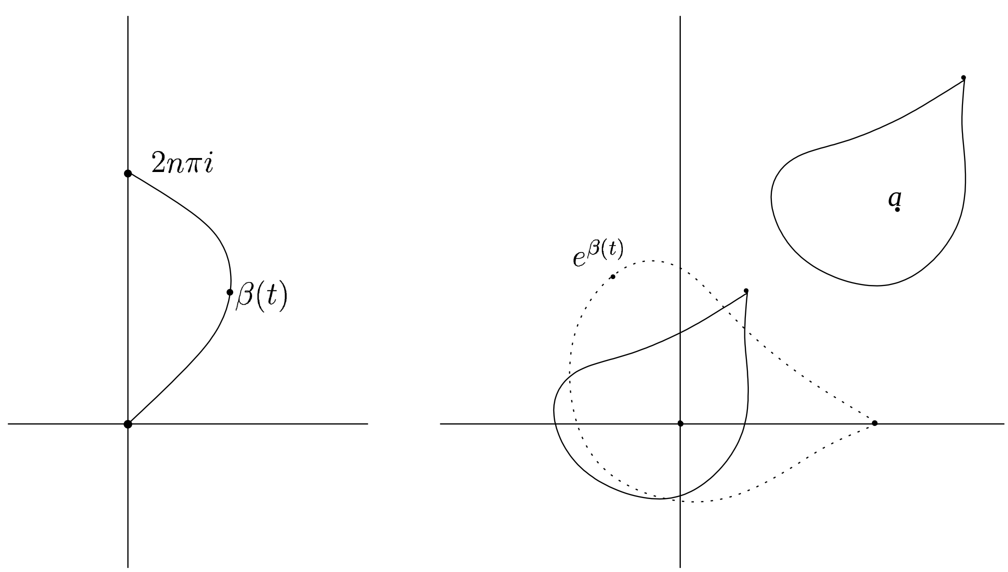
therefore, $[\alpha(t) - a]e^{-\beta(t)}$ is a constant function of t and is equal to $[\alpha(t_1) - a]$, since $\beta(t_1) = 0$. Therefore

$$e^{\beta(t)} = \frac{\alpha(t) - a}{\alpha(t_1) - a}$$

therefore,

$$e^{\beta(t_2)} = \frac{\alpha(t_2) - a}{\alpha(t_1) - a} = 1$$

and therefore, $\beta(t_2) = 2n\pi i$, with n an integer.



Therefore,

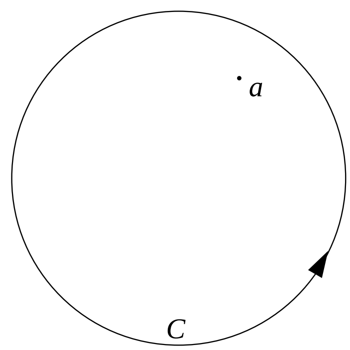
$$\int_\alpha \frac{dz}{z-a} = \int_{t_1}^{t_2} \frac{\alpha'(s)}{\alpha(s)-a}\,ds = \beta(t_2) = 2n\pi i$$

as required. □

Theorem 5.4.2. *Let C be a circle and $a \in \mathbb{C}$. Then*

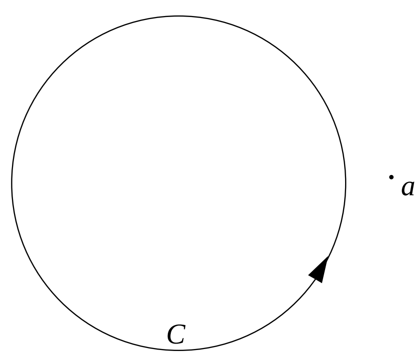
1. *if a is inside C, then C has winding number 1 about a:*

$$\int_C \frac{dz}{z-a} = 2\pi i$$

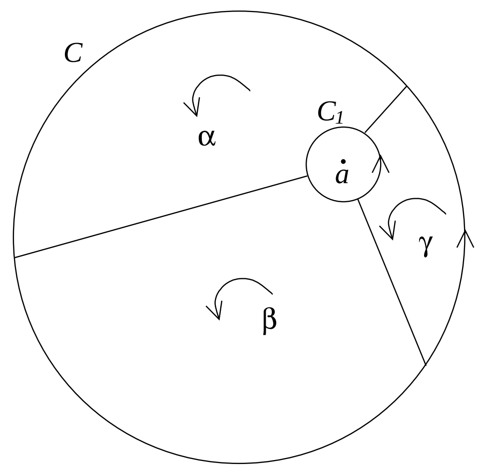


2. *if a is outside C then winding number about a is 0.*

$$\int_C \frac{dz}{z-a} = 0$$



Proof. 1. Let a be inside C, let C_1 be a circle, centre a inside C. Let α, β, γ be the closed paths shown, each is contained in an open star shaped set on which $\frac{1}{z-a}$ is holomorphic.



$$\int_\alpha \frac{dz}{z-a} + \int_\beta \frac{dz}{z-a} + \int_\gamma \frac{dz}{z-a} = \int_C \frac{dz}{z-a} - \int_{C_1} \frac{dz}{z-a}$$

$$0 + 0 = 0 = \int_C - \int_{C_1}$$

therefore,

$$\int_C \frac{dz}{z-a} = \int_{C_1} \frac{dz}{z-a} = \int_0^{2\pi} \frac{ire^{i\theta}}{re^{i\theta}}\, d\theta = 2\pi i$$

(put $z = a + re^{i\theta}$).

2. Let a be outside C, then C is contained in an open star-shaped set on which $\frac{1}{z-a}$ is holomorphic. Therefore,

$$\int_C \frac{dz}{z-a} = 0$$

□

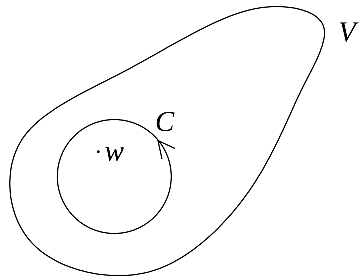
5.5 Cauchy's Integral Formula

Theorem 5.5.1. (Cauchy's Integral Formula)

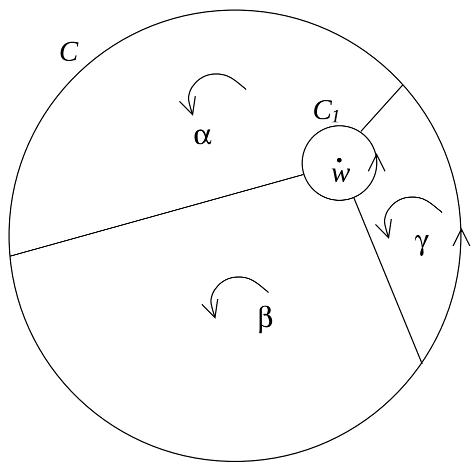
Let f be holomorphic on open V in $\mathbb{C}$. Let $w \in V$. Then, for any circle C around w, such that C and it's interior is contained in V, we have:

$$f(w) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z-w}\, dz$$

(Thus the values of f on any circle uniquely determine the values of f inside the circle)



Proof. Let $\epsilon > 0$. Choose a circle C_1 centre w, radius r (say) inside C such that $|f(z) - f(w)| \leq \epsilon \quad \forall z \in C_1$ by continuity of f.



Let α, β, γ be as indicated. Then, by integrating $\frac{f(z)}{z-w}$:

$$\int_C - \int_{C_1} = \int_\alpha + \int_\beta + \int_\gamma = 0 + 0 + 0 = 0$$

since each of α, β, γ are contained in a star-shaped set on which $\frac{f(z)}{z-w}$ is a holomorphic function of z. Therefore,

$$\begin{aligned} \int_C \tfrac{f(z)}{z-w}\, dz &= \int_{C_1} \tfrac{f(z)}{z-w}\, dz \\ &= \int_{C_1} \tfrac{f(w)}{z-w}\, dz + \int_{C_1} \tfrac{f(z)-f(w)}{z-w}\, dz \\ &= 2\pi i\, f(w) + 0 \end{aligned}$$

because:

$$\left| \int_{C_1} \frac{f(z) - f(w)}{z - w}\, dz \right| \leq \frac{\epsilon}{r} 2\pi r = 2\pi\epsilon$$

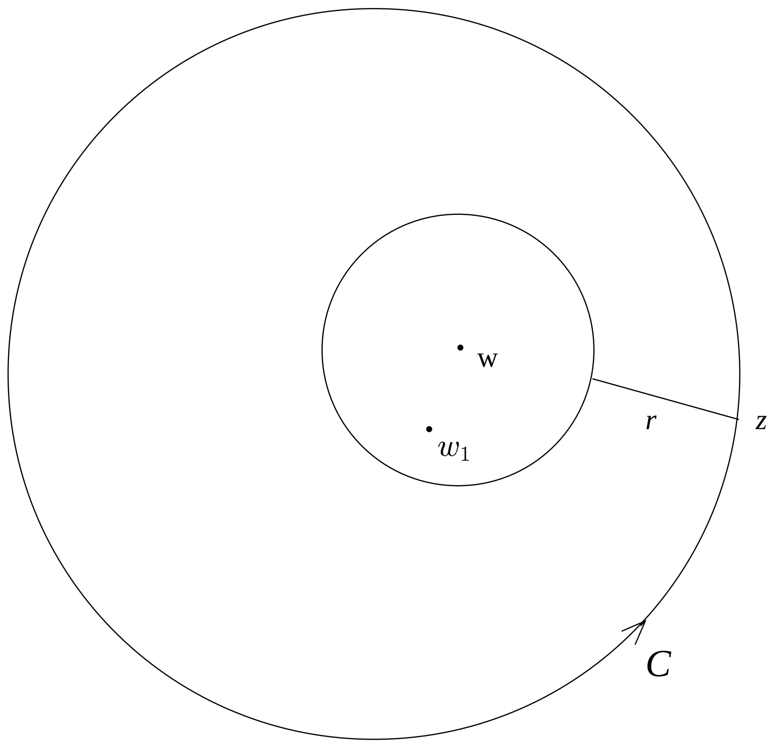
for all $\epsilon > 0$. Hence result. $\square$

Let $\mathbb{C} \supset V \xrightarrow{f} \mathbb{C}$ be holomorphic on V open. Let $w \in V$. Pick a circle C around w such that C and it's interior $\supset V$. We have:

$$f(w) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - w}\, dz$$

differentiating w.r.t. w under the integral sign:

$$f'(w) = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - w)^2}\, dz$$



Justified since $\exists\, r > 0$ s.t.

$$\left| \frac{f(z)}{(z - w_1)^2} \right| \leq \frac{|f(z)|}{r^2}$$

for all w_1 on an open set containing w_1, which is *integrable* w.r.t. z.

Repeating n times gives:

$$f^{(n)}(w) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z - w)^{n+1}}\, dz \tag{5.1}$$

Thus, f holomorphic on open $V \implies f$ has derivatives of *all* orders and $f^{(n)}(w)$ is given by Equation(5.1) for any suitable circle C (or any suitable closed curve) around w.

5.6 Term-by-term differentiation, analytic functions, Taylor series

Theorem 5.6.1. *Let $\{f_n(z)\}$ be a sequence of functions which is uniformly convergent to the function $f(z)$ on a path α. Then*

$$\lim_{n\longrightarrow\infty}\int_\alpha f_n(z)\,dz = \int_\alpha f(z)\,dz$$

Proof. Let $\epsilon > 0$. Choose N s.t.

$$|f_n(z) - f(z)| \leq \epsilon \qquad \forall n \geq N \quad \forall\, z = \alpha(t) \quad t_1 \leq t \leq t_2$$

(definition of uniform convergence). Then

$$\begin{aligned} \left|\int_\alpha f_n(z)\,dz - \int_\alpha f(z)\,dz\right| &\leq \int_\alpha |f_n(z) - f(z)\,dz| \\ &\leq \epsilon\, l(\alpha) \qquad \forall\; n \geq N \end{aligned}$$

hence the result. □

Theorem 5.6.2. (term by term differentiation)

Let $\sum_{i=1}^\infty f_n(z)$ be a series of holomorphic functions on an open set V, which converges uniformly on a circle C which, together with it's interior, is contained in V.

Then the series converges inside C to a holomorphic function F (say) and

$$f^{(k)} = \sum_{n=1}^\infty f_n^{(k)}$$

inside C

Proof. Let w be inside C, then $\lambda = \inf_{z\in C}|z-w| > 0$.

$$\Longrightarrow \quad \frac{1}{|z-w|^{k+1}} \leq \frac{1}{\lambda^{k+1}} \qquad \forall\, z \in C$$

Therefore, $\sum_{n=1}^\infty \frac{f_n(z)}{(z-w)^{k+1}}$ converges uniformly to $\frac{f(z)}{(z-w)^{k+1}}$ $z \in C$. Therefore,

$$\sum_{n=1}^\infty \int_C \frac{f_n(z)}{(z-w)^{k+1}}\,dz = \int_C \frac{f(z)}{(z-w)^{k+1}}\,dz$$

and therefore,

$$\sum_{n=1}^\infty f_n^{(k)}(w) = f^{(n)}(w)$$

as required. □

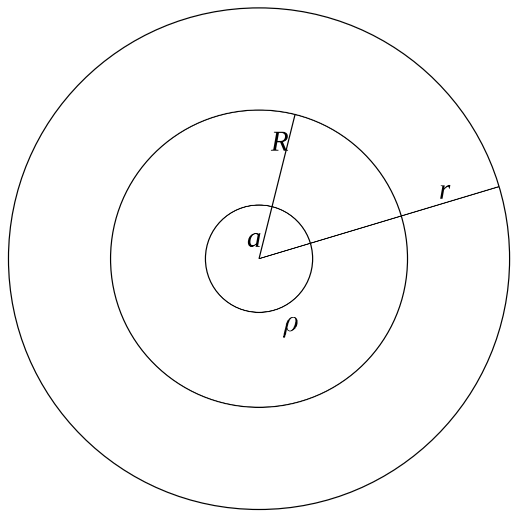
Definition Let $\mathbb{C} \supset V \xrightarrow{f} \mathbb{C}$ with V opn. Then f is called *analytic* on V if, for each $a \in V$ $\exists r > 0$ and constants $c_0, c_1, c_2, \ldots \in \mathbb{C}$, such that:

$$f(z) = c_0 + c_1(z-a) = c_2(z-a)^2 + \cdots \qquad \forall z \text{ s.t. } |z-a| < r$$

The R.H.S. is called the *Taylor series* of f about a.

Lemma 5.6.3. *If* $0 < \rho < r$ *then the series converges uniformly on*

$$\{z : |z-a| \le \rho\}$$



Proof. Let $0 < \rho < R < r$. $\sum c_n(z-a)^n$ converges for $z - a = R$, and therefore, $\sum c_n R^n$ converges. Therefore, $\exists K$ s.t. $|c_n R^n| \le K$ $\forall n$. Therefore

$$|c_n(z-a)^n| \le |c_n|\,|z-a|^n \le \frac{K}{R^n}\rho^n = K\left(\frac{\rho}{R}\right)^n$$

for all $|z-a| \le \rho$. □

Therefore, we can differentiate term by term n times:

$$f^{(n)}(z) = n!c_n + (n+1)!c_{n+1}|z-a| + \cdots \text{ higher powers of } z-a$$

Therefore, $f^{(n)}(a) = n!c_n$, and therefore, $c_n = \frac{1}{n!}f^{(n)}(a)$.

The Taylor coefficents are uniquely determined, f is C^∞, and

$$f(z) = \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(z-a)^n \qquad \text{on } |z-a| < r$$

Thus, f analytic on $V \implies f$ holomorphic on V.

Conversely,

Theorem 5.6.4. *Let $\mathbb{C} \supset V \xrightarrow{f} \mathbb{C}$ be holomorphic on open V. Let C be any circle, centre a s.t. C and it's interior are contained in V. Then f has a Taylor series about a convergent inside C. Hence f is analytic on V.*

Proof. Let w be inside C. Then

$$\begin{aligned}
f(w) &= \tfrac{1}{2\pi i}\int_C \tfrac{f(z)}{z-w}\,dz \\
&= \tfrac{1}{2\pi i}\int_C \tfrac{f(z)}{(z-a)-(w-a)}\,dz \\
&= \tfrac{1}{2\pi i}\int_C \tfrac{f(z)}{(z-a)\left[1-\frac{w-a}{z-a}\right]}\,dz \\
&= \tfrac{1}{2\pi i}\int_C \tfrac{f(z)}{z-a}\textstyle\sum_{n=1}^{\infty}\left(\frac{w-a}{z-a}\right)^n\,dz \\
&= \textstyle\sum_{n=1}^{\infty}(w-a)^n \underbrace{\frac{1}{2\pi i}\int_C \frac{f(z)}{(z-a)^{n=1}}\,dz}_{c_n} \\
&= \textstyle\sum_{n=0}^{\infty}\frac{1}{n!}f^{(n)}(c)(w-a)^n
\end{aligned}$$

as required. ☐

Chapter 6

Further Calculus

6.1 Mean Value Theorem for Vector-valued functions

Theorem 6.1.1. (Mean value theorem for vector valued functions)
Let M, N be finite dimensional normed spaces and let

$$M \supset V \xrightarrow{f} N$$

be a C^1 function, where V is open in M. Let $x, y \in V$ and $[x, y] \subset V$. Then

$$\|f(x - f(y)\| \leq k \|x - y\|$$

where $k = \sup_{z \in [x,y]} \|f'(z)\|$

Proof.

$$\begin{aligned} f(y) - f(x) &= \int_0^1 \tfrac{d}{dt} f[ty + (1-t)x]\, dt \\ &= \int_0^1 \underbrace{f'[ty + (1-t)x]}_{\text{operator}} \underbrace{(y - x)}_{\text{vector}}\, dt \end{aligned}$$

Therefore,

$$\begin{aligned} \|f(y) - f(x)\| &\leq \int_0^1 \|f'[ty + (1-t)x]\| \, \|y - x\|\, dt \\ &\leq \int_0^1 k \|y - x\|\, dt \\ &= k \|y - x\| \end{aligned}$$

as required. □

6.2 Contracting Map

Theorem 6.2.1. *Let M be a finite dimensional real vector space with norm. Let*

$$M \supset V \xrightarrow{f} M$$

be a C^1 function, V open in M, $f(0) = 0$, $f'(0) = 1$.

Let $0 < \epsilon < 1$, and let B be a closed ball centre O s.t.

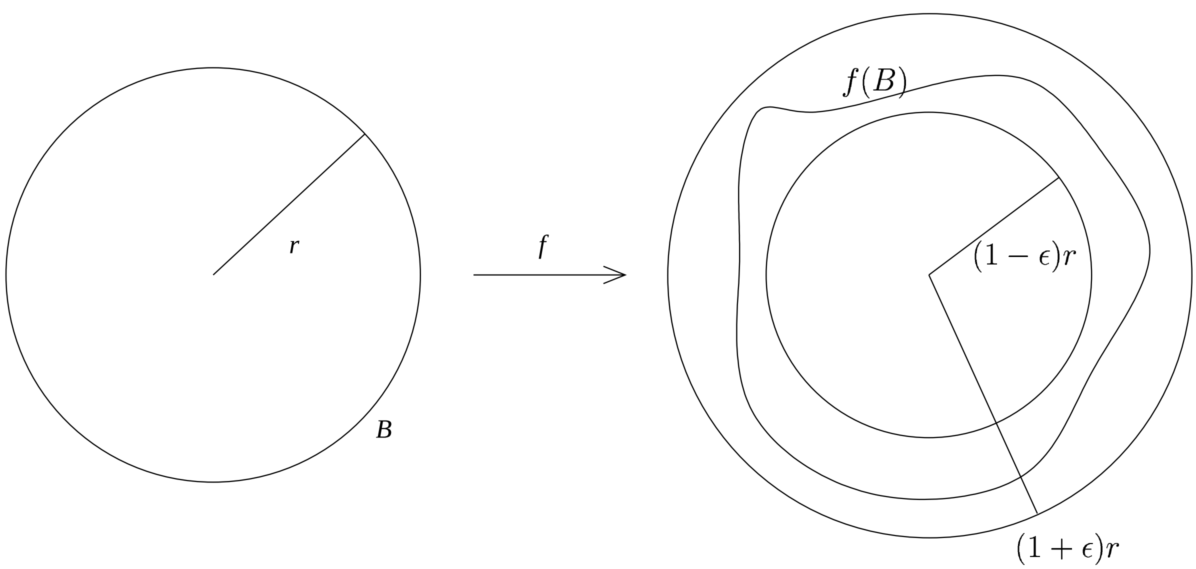
$$\|1 - f'(x)\| \leq \epsilon \qquad \forall x \in B$$

Then

1. $$\|f(x) - f(y)\| \geq (1 - \epsilon) \|x - y\| \qquad \forall\, x, y \in B \tag{6.1}$$

   Thus f_B is injective.

2. $(1 - \epsilon)B \subset f(B) \subset (1 + \epsilon)B$



Proof. Let r =radius of B

1. $$\|f'(x)\| = \|1 + (f'(x) - 1)\| \leq 1 + \epsilon \qquad \forall x \in B$$

   therefore,

   $$\|f(x)\| = \|f(x) - f(0)\| \overset{MVT}{\leq} (1 + \epsilon)\|x\| \leq (1 + \epsilon)r$$

   for all $x \in B$. Therefore, $f(B) \subset (1 + \epsilon)B$

2.

$$\|(1-f)(x)-(1-f)(y)\| \overset{MVT}{\leq} \epsilon\|x-y\| \qquad \forall\, x,y \in B$$

therefore, $\|x-y| - \|f(x)-f(y)\| \leq \epsilon\, \|x-y\|$, and so

$$\|f(x)-f(y)\| \geq (1-\epsilon)\|x-y\|$$

and hence Eqn(6.1).

3. To show $(1-\epsilon)B \subset f(B)$. Let $a \in (1-\epsilon)B$, define $g(x) = x - f(x) + a$. Then

$$\|g'(x)\| = \|1 - f'(x)\| \leq \epsilon \qquad \forall x \in B$$

therefore,

$$\|g(x)-g(y)\| \overset{MVT}{\leq} \epsilon\|x-y\|$$

therefore, g is a *contracting map* (shortens distances).

Also,

$$\begin{aligned} \|g(x)\| &= \|g(x)-g(0)+a\| \qquad \text{since } a = g(0) \\ &\leq \|g(x)-g(0)\| + \|a\| \\ &\leq \epsilon\|x\| + \|a\| \\ &\leq \epsilon r + (1-\epsilon)r \qquad \forall\, x \in B \\ &= r \end{aligned}$$

therefore, $g(x) \in B\ \forall\, x \in B$. So g maps B into B and is contracting. Therefore, by the contraction mapping theorem $\exists x \in B$ s.t. $g(x) = x$.

i.e. $x - f(x) + a = x$

i.e. $f(x) = a$. Therefore, $a \in f(B)$, and $(1-\epsilon)B \subset f(B)$ as required.

□

6.3 Inverse Function Theorem

Definition Let $M \supset V \overset{f}{\longrightarrow} W \subset N$ where M, N are finite dimensional normed spaces. Then f is called a C^r *diffeomorphism* if

1. V open in M, W open in N
2. $V \overset{f}{\longrightarrow} W$ is bijective

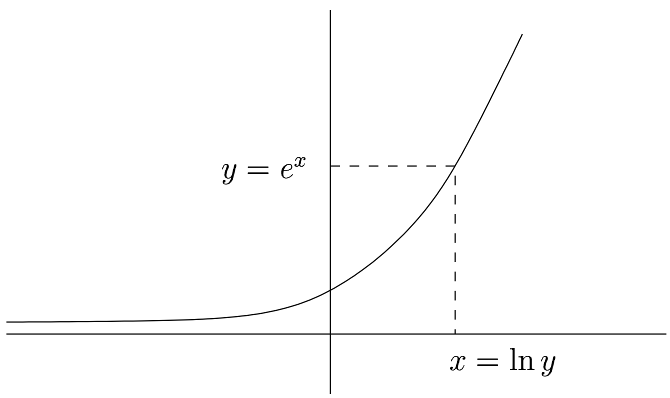
3. f and f^{-1} are C^r

V is C^r *diffeomorphic* to W if $\exists$ a C^r-diffeomorphism $V \longrightarrow W$.

Example

$$\mathbb{R} \xrightarrow{f} (0, \infty)$$

$f(x) = e^x$ is C^∞, and $f^{-1}(x) = \ln x$ is C^∞. Therefore, f is a C^∞ diffeomorphism.



Theorem 6.3.1. (Inverse Function Theorem)

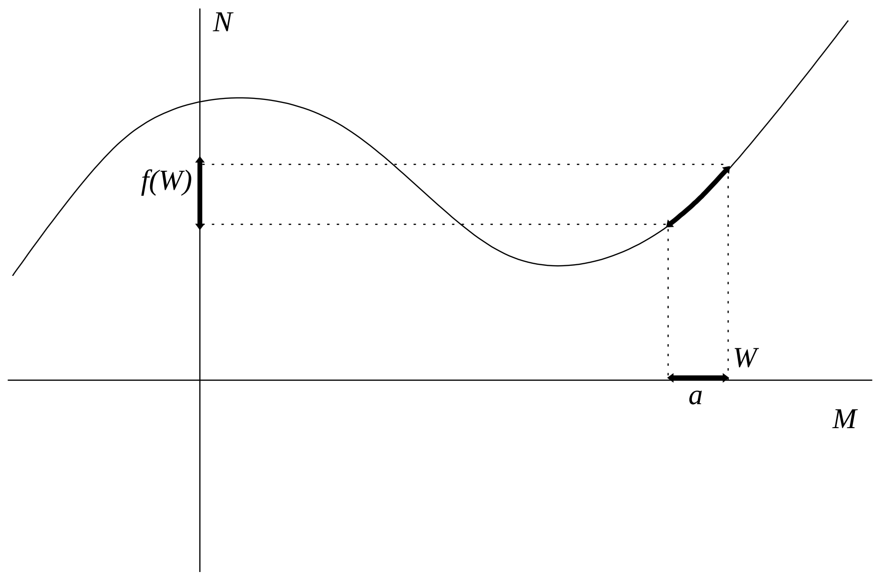
Let $M \supset V \xrightarrow{f} N$ be C^r where M, N are finite dimensional normed spaces and V is open in M. Let $a \in V$ be a point at which

$$f'(a) : M \longrightarrow N$$

is invertible. Then $\exists$ open neighbourhood W of a such that

$$f_W : W \longrightarrow f(W)$$

is a C^r diffeomorphism.

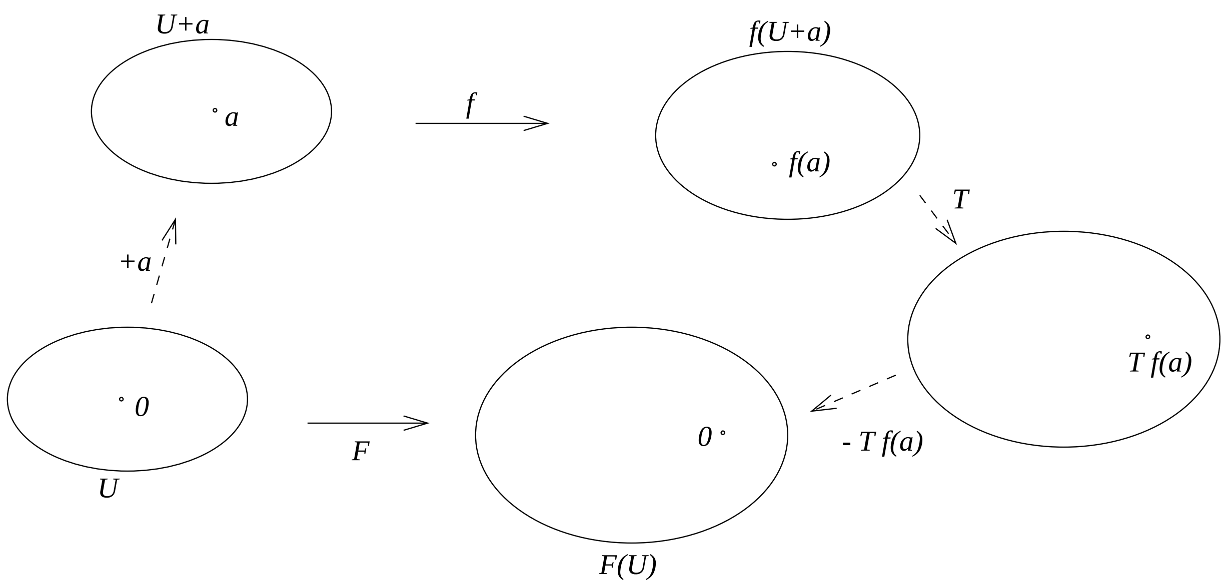


Proof. Let T be the inverse of $f'(a)$ and let F be defined by

$$F(x) = Tf(x + a) - Tf(a)$$

We have:

$$F(0) = Tf(a) - Tf(a) = 0$$



We prove that F maps an open neighbourhood U of 0 onto an open neighbourhood $F(U)$ of 0. It the follows that f maps open $U + a$ onto open

$f(U+a)$ by a C^r diffeomorphism. Now

$$F'(x) = T \cdot f'(x+a)$$

$$\implies F'(0) = T \cdot f'(a) = 1_M$$

Choose a closed ball B centre 0 of positive radius s.s.

$$\|F'(x) - 1_M\| \leq \frac{1}{2} \qquad \forall x \in B$$

and also s.t. $\det F'(x) \neq 0$. Then by the previous theorem (with $\epsilon = \frac{1}{2}$) we have:

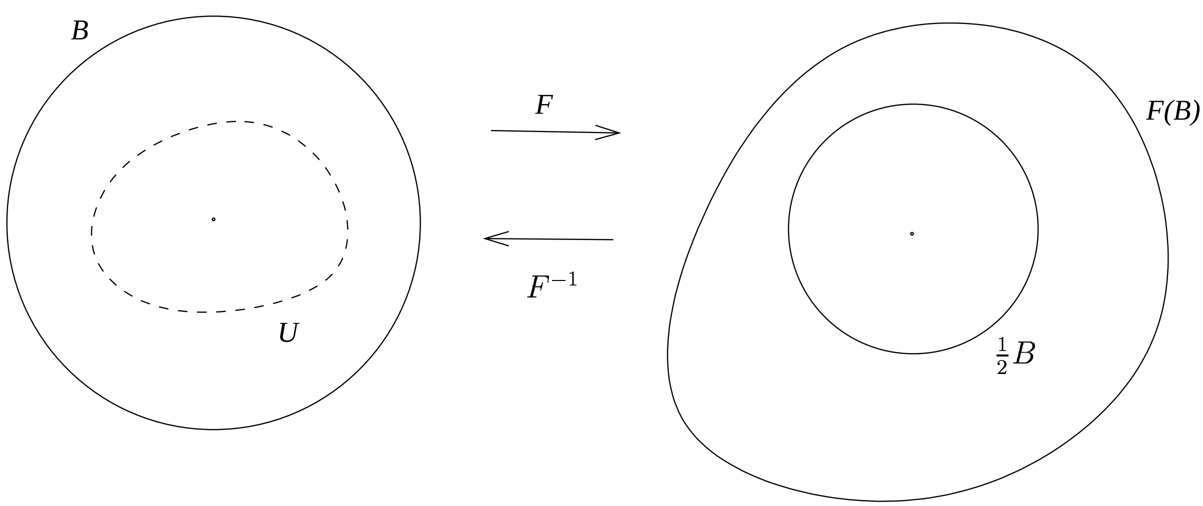
$$F_B \text{ is injective}$$

and

$$\|F(x) - F(y)\| \geq \frac{1}{2}\|x-y\| \qquad \forall x, y, \in B$$

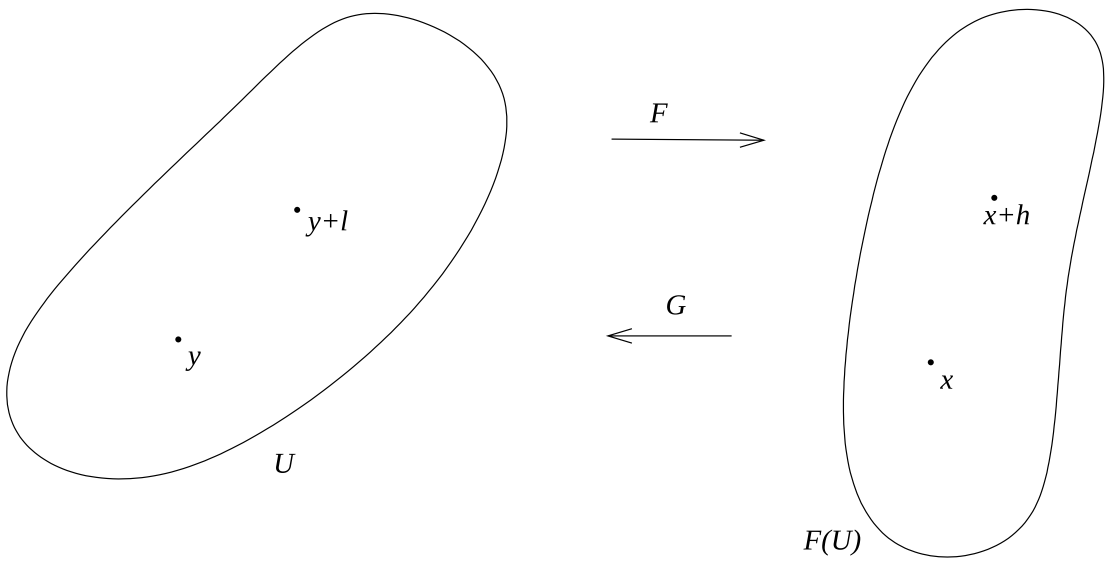
and

$$\frac{1}{2}B \subset F(B)$$



therefore, $F^{-1} : \frac{1}{2}B \longrightarrow B$ is well-defined and continuous. Let B^0 be the interior of B, an open set. Put $U = F^{-1}(\frac{1}{2}B^0) \cap B^0$, an open set. $F(U)$ is open since F^{-1} is continuous. So $F_U : U \longrightarrow F(U)$ is a homeomorphism of open U onto open $F(U)$.

Let G be it's inverse. To show G is C^r.



Let $x, x+h \in F(U)$, $G(x) = y$, $G(x+h) = y+l$ (say), and $\|h\| \geq \frac{1}{2}\|l\|$.

Let $F'(y) = S$. Then $F(y+h) = F(y) + Sl + \phi(l)$, where $\frac{\|\phi(l)\|}{\|l\|} \longrightarrow 0$ as $\|l\| \longrightarrow 0$.

$\implies x+h = x + Sl + \phi(l)$,

$\implies l = S^{-1}h - S^{-1}\phi(l)$

$\implies G(x+h) = y + l = G(x) + S^{-1}h - S^{-1}\phi(l)$

Now,

$$\frac{\|S^{-1}\phi(l)\|}{\|h\|} \leq \|S^{-1}\| \frac{\|\phi(l)\|}{\|l\|} \frac{\|l\|}{\|h\|} \longrightarrow 0$$

as $\|h\| \longrightarrow 0$ since $\|l\| \leq 2\|h\| \implies \frac{\|l\|}{\|h\|} \leq 2$.

So, G is differentiable at x $\forall x \in F(U)$ and

$$G'(x) = S^{-1} = [F'(y)]^{-1} = [F'(G(x))]^{-1}$$

It follows that if G is C^s for some $0 \leq s < r$ then G' is C^s since G' is a composition of C^s functions

$$F',\ G,\ [\cdot]^{-1}$$

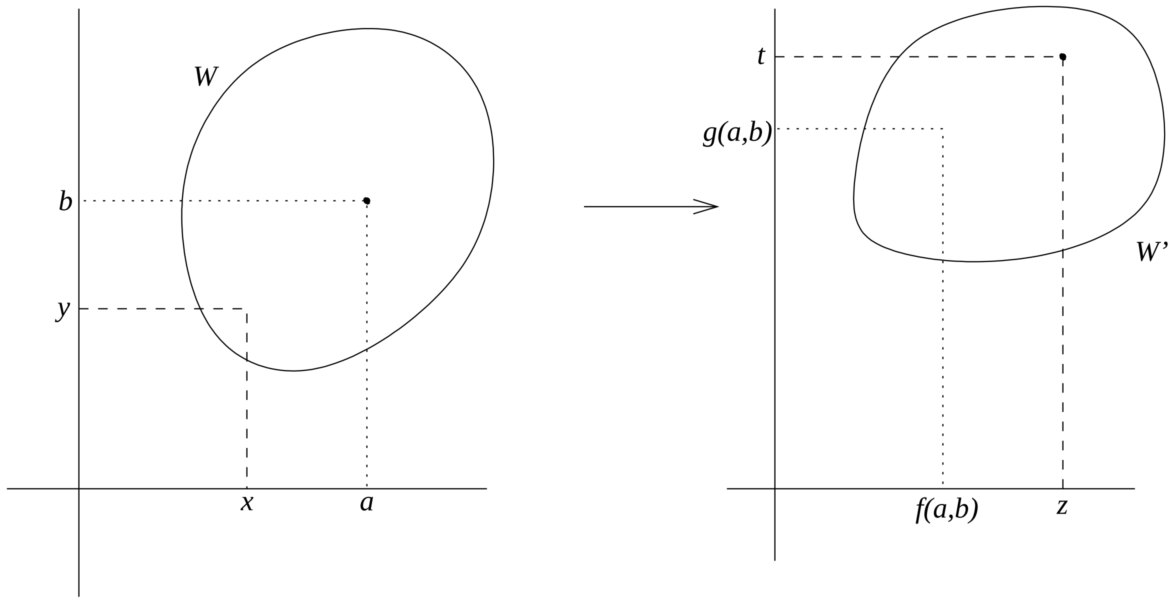
and therefore G is C^{s+1}. So

$$G\ C^s \implies G\ C^{s+1} \qquad \forall 0 \leq s < r$$

therefore G is C^r as required.

$\square$

Example Let f, g be C^r on an open set containing (a, b)



Let

$$\frac{\partial(f,g)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{vmatrix} \neq 0$$

at (a, b). Then (f, g) maps an open neighbourhood W of (a, b) onto an open neighbourhood W' of $(f(a, b), g(a, b)$ by a C^r diffeomorphism. Therefore, for each $z, t \in W'$ $\exists$ unique $(x, y) \in W$ such that

$$z = f(x, y)$$

$$t = g(x, y)$$

and $x = h(z, t)$, $y = k(z, t)$ where h and k are C^r.

Chapter 7

Coordinate systems and Manifolds

7.1 Coordinate systems

Definition Let X be a topological space. Let V be open in X. A sequence

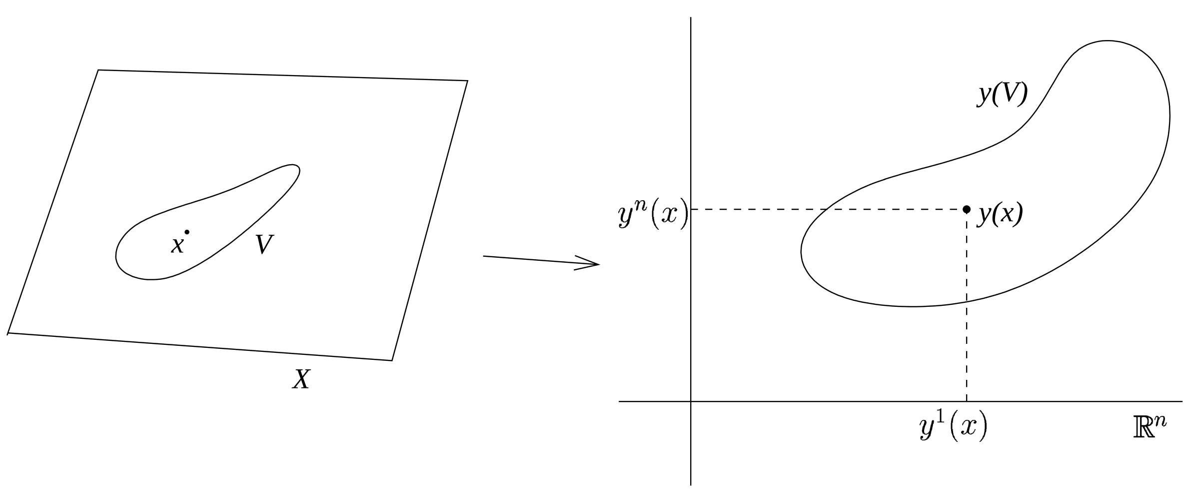
$$y = (y^1, \ldots, y^n)$$

of real-valued functions on V is called an *n-dimensional coordinate system* on X with domain V if

$$X \supset V \xrightarrow{y} y(V) \subset \mathbb{R}^n$$

$$x :\longrightarrow y(x) = (y^1(x), \ldots, y^n(x))$$

is a homeomorphism of V onto an open set $y(V)$ in $\mathbb{R}^n$.



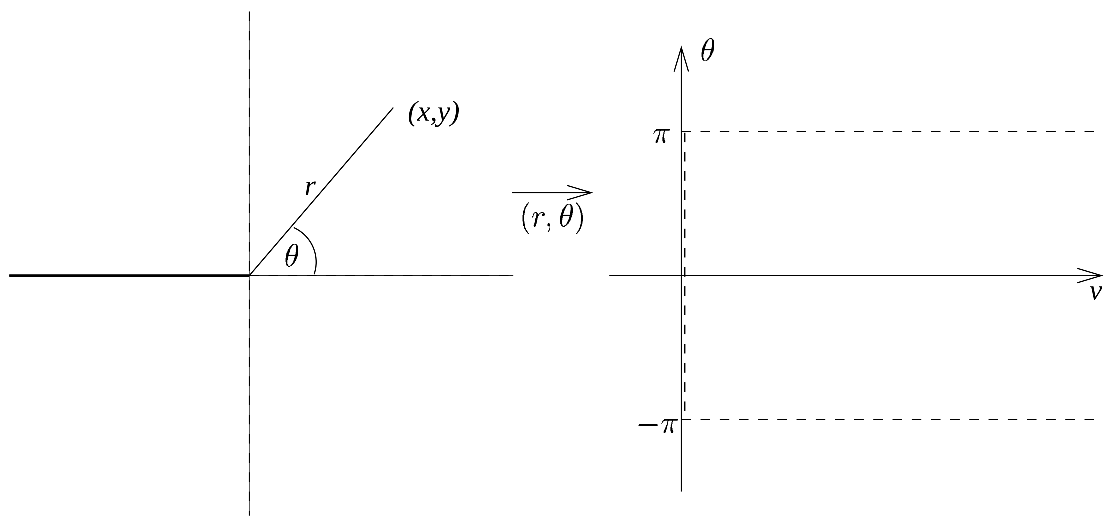
Example On the set $V = \{y \neq 0 \text{ or } x > 0\}$ in $\mathbb{R}^2$ the functions r, θ given by

$$x = r\cos\theta$$

$$y = r\sin\theta$$

$$-\pi < \theta < \pi$$

map V homomorphically onto an open set in $\mathbb{R}^2$.



Therefore, (r, θ) is a 2-dimensional coordinate system on $\mathbb{R}^2$ with domain V.

Definition If $y = (y^1, \dots, y^n)$ is a coordinate system with domain V and if f is a real-valued function on V then

$$f = F(y^1, \dots, y^n)$$

for a unique function F on $y(V)$ s.t. $F = f \cdot y^{-1}$

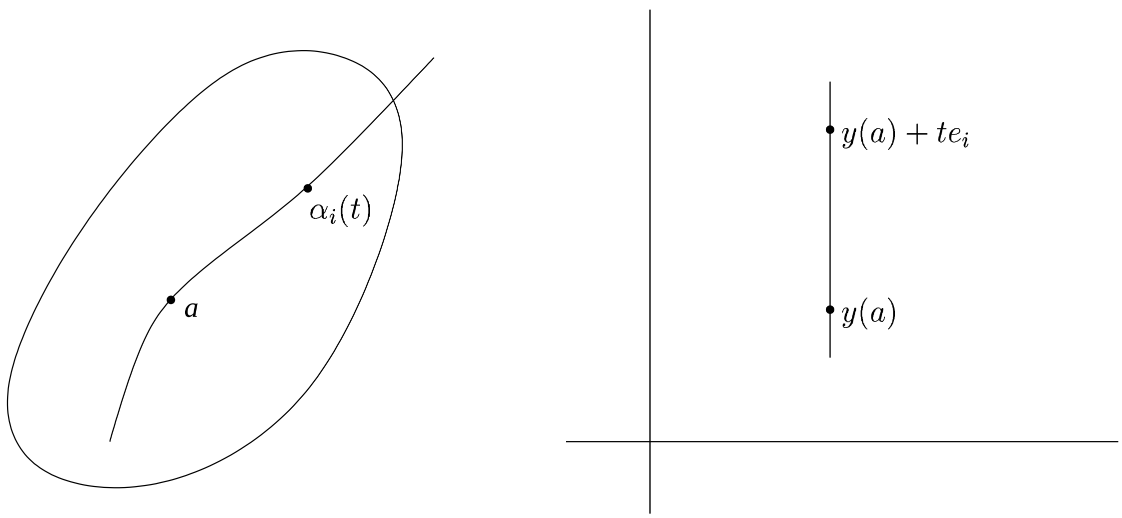
We call f a C^r *function of* $y^1, \dots, y^n$ if F is C^r and we write

$$\frac{\partial f}{\partial y^i} = \frac{\partial F}{\partial x^i}(y^1, \dots, y^n)$$

and call it the partial derivative w.r.t Y^i in the coord system $y^1, \dots, y^n$.

Note:

$$
\begin{aligned}
\frac{\partial f}{\partial y^i}(a) &= \frac{\partial F}{\partial x^i}(y^1(a),\dots,y^n(a)) \\
&= \frac{d}{dt}F(y(a)+te_i)\Big|_{t=0} \\
&= \frac{d}{dt}f(\alpha_i(t))\Big|_{t=0} \\
&= \text{rate of change of } f \text{ along curve } \alpha_i \text{ in } V
\end{aligned}
$$



where α_i is the curve given by:

$$y(\alpha_i(t)) = y(a) + te_i = (y^1(a),\dots,y^i(a)+t,\dots,y^n(a))$$

therefore, along curve α_i all coords $y^1,\dots,y^n$ are constant except i^{th} coord y^i and $\alpha_i(t)$ is parameterised by change t in i^{th} coord y^i.

α_i is called the *i^{th} coordinate curve at a.*

7.2 C^r-manifold

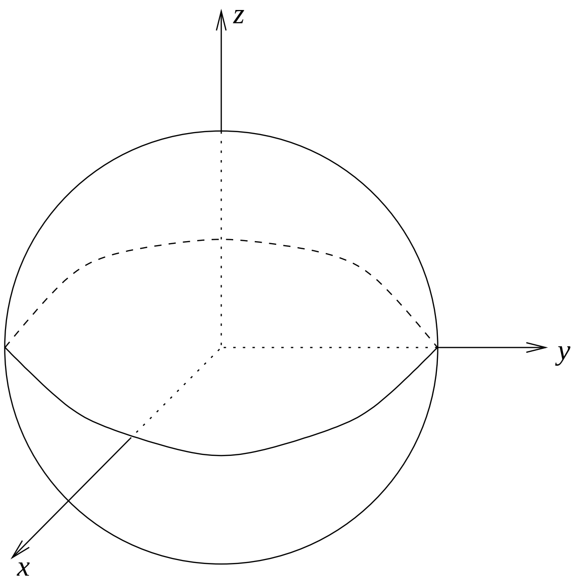
Definition Let $y^1,\dots,y^n$ with domain V, and $z^1,\dots,z^n$ with domain W be two coordinate systems on X. Then these two systems are called *C^r-compatible* if each z^i is a C^r function of $y^1,\dots,y^n$, and each y^i is a C^r function of $z^1,\dots,z^n$ on $V\cap W$.

We call X an *n-dimensional C^r-manifold* if a collection of n-dimensional coordinate systems is given whose domains cover X and which are C^r-compatible.

Example the 2-sphere S^2 given by:

$$x^1 + y^1 + z^2 = 1 \qquad \text{in } \mathbb{R}^3$$

The function x, y with domain $\{z > 0\}$ is a 2-dimensional coordinate system on S^2.



Similarly the functions x, z with domain $\{y > 0\}$ is a 2-dimensional co-ordinate system

On the overlap $\{y > 0, z > 0\}$ we have:

$$x = x \qquad\qquad x = x$$

$$y = \sqrt{1 - x^2 - y^2} \qquad\qquad z = \sqrt{1 - x^2 - y^2}$$

these systems are C^∞-compatible. In this way, we make S^2 into a C^∞-manifold.

7.3 Tangent vectors and differentials

Definition Let $a \in X$, X a manifold. Let $y^1, \ldots, y^n$ be co-ords on X at a (i.e. domain an open neighbourhood of a). Then, the linear operators

$$\left.\frac{\partial}{\partial y^1}\right|_a, \cdots, \left.\frac{\partial}{\partial y^n}\right|_a \tag{7.1}$$

act on the differentiable functions f at a (i.e. functions f which are real-valued differentiable functions of $y^1, \ldots, y^n$ on an open neighbourhood of a) and are defined by:

$$\left.\frac{\partial}{\partial y^j}\right|_a f = \frac{\partial f}{\partial y^j}(a)$$

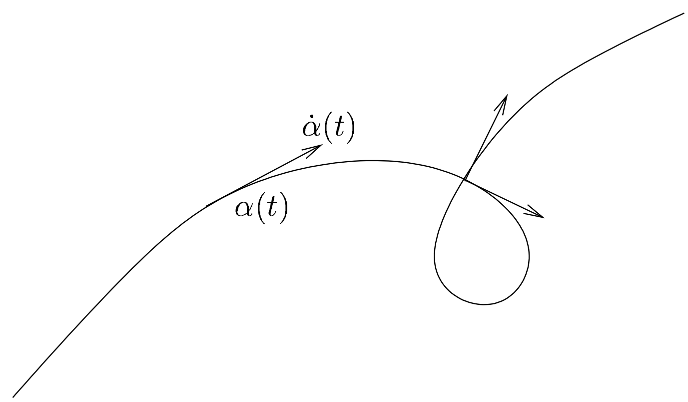
The operators (7.1) are linearly independent since

$$\left[\alpha^j \left.\frac{\partial}{\partial y^j}\right|_a\right] y^i = \alpha^j \frac{\partial y^i}{\partial y^j}(a) = \alpha^j \delta^i_j = \alpha^i \tag{7.2}$$

$$\Longrightarrow \alpha^j \left.\frac{\partial}{\partial y^j}\right|_a = 0 \Longrightarrow \alpha^i = 0 \qquad \forall i \tag{7.3}$$

The real vector space with basis (7.1) is denoted T_aX and is called the *tangent space* of X at a. If $v \in T_aX$ then (7.2) shows that v has components vy^i w.r.t. basis (7.1)

Definition If $\alpha(t)$ is a curve in X then the *velocity vector* $\dot{\alpha}(t) \in T_{\alpha(t)}X$ is the tangent vector at $\alpha(t)$ given by taking rate of change along $\alpha(t)$ w.r.t. t:



i.e.

$$\begin{aligned}
\dot{\alpha}(t)f &= \tfrac{d}{dt} f(\alpha(t)) \\
&= \tfrac{d}{dt} F(y^1(\alpha(t)), \ldots, y^n(\alpha(t))) \qquad \text{if } f = F(y^1, \ldots, y^n) \\
&= \tfrac{\partial F}{\partial x^j}(y^1(\alpha(t)), \ldots, y^n(\alpha(t))) \tfrac{d}{dt} y^j(\alpha(t)) \\
&= \tfrac{\partial f}{\partial y^j}(\alpha(t)) \tfrac{d}{dt} y^j(\alpha(t)) \\
&= \left[\tfrac{d}{dt} y^j(\alpha(t)) \left.\tfrac{\partial}{\partial y^j}\right|_{\alpha(t)}\right] f
\end{aligned}$$

therefore, $\dot{\alpha(t)}$ is the tangent vector with components

$$\frac{d}{dt}y^j(\alpha(t)) = \dot{y}^j(t)$$

i.e.

$$\dot{\alpha}(t) = \dot{y}^j(t)\left.\frac{\partial}{\partial y^j}\right|_{\alpha(t)}$$

The tangent space T_aX does not depend on the choice of (compatible) coordinates at a since if $z^1, \ldots, z^n$ is another coordinate system at a, then

$$\begin{aligned} \left.\tfrac{\partial}{\partial z^j}\right|_a &= \text{velocity vector of } j^{\text{th}} \text{ coordinate curve of } z^1, \ldots, z^n \text{ at } a \\ &= \tfrac{\partial y^i}{\partial z^j}\left.\tfrac{\partial}{\partial y^i}\right|_a \quad \in T_aX \end{aligned}$$

therefore,

$$\left.\frac{\partial}{\partial z^1}\right|_a, \cdots, \left.\frac{\partial}{\partial z^n}\right|_a$$

is also a basis for T_aX, and $\frac{\partial y^i}{\partial z^j}(a)$ is the transition matrix from $z^1, \ldots, z^n$ to $y^1, \ldots, y^n$.

Note also that the curve $\alpha(t)$ with coordinates:

$$y(\alpha(t)) = (y^1(a) + \alpha^1 t, \ldots, y^n(a) + \alpha^n t)$$

has velocity vector

$$\dot{\alpha}(0) = \alpha^1\left.\frac{\partial}{\partial y^1}\right|_a + \cdots \alpha^n\left.\frac{\partial}{\partial y^n}\right|_a$$

Therefore, every tangent vector (at a) is the velocity vector of some curve, and vice versa.

Definition Let $a \in X$ and f be a differentiable function at a. Then for each $v \in T_aX$ with $v = \dot{\alpha}(t)$ (say), define

$$\langle df_a, v\rangle = vf = \dot{\alpha}(t)f = \frac{d}{dt}f(\alpha(t)) = \text{rate of change of } f \text{ along } v$$

Thus df_a is a linear form on T_aX, called the *differential of f at a* df_a measures the rate of change of f at a. If $y^1, \ldots, y^n$ are coordinates on X at a then

$$\left\langle dy^i_a, \frac{\partial}{\partial y^j}_a\right\rangle = \left.\frac{\partial}{\partial y^j}\right|_a y^i = \frac{\partial y^i}{\partial y^j}(a) = \delta^i_j$$

therefore, $dy_a^1, \ldots, dy_a^n$ is the basis of T_aX^* which is dual to the basis $\frac{\partial}{\partial y^1}\Big|_a, \cdots, \frac{\partial}{\partial y^n}\Big|_a$ of T_aX.

Thus the linear form dy_a^i gives the i^{th} component of tangent vectors at a:

$$\langle dy_a^i, v\rangle = i^{\text{th}} \text{ component of } v = vy^i$$

Also, df_a has components:

$$\left\langle df_a, \frac{\partial}{\partial y^j}_a \right\rangle = \frac{\partial}{\partial^j}_a f = \frac{\partial f}{\partial y^j}(a)$$

therefore,

$$df_a = \frac{\partial f}{\partial y^j}(a)\, dy_a^j$$

(which is called the chain rule for differentials.)

7.4 Tensor Fields

From now on assume manifolds, functions are C^∞.

Definition Let W be an open set in a manifold X. A *tensor field* S on X with domain W is a function on W:

$$x \longmapsto S_x$$

which assigns to each $x \in W$ a tensor of fixed type over the tangent space T_aX at x. e.g.

$$S_x : T_xX \times (T_xX)^* \times T_xX \longrightarrow \mathbb{R}$$

We can add tensor fields, contract them, form tensor products and wedge products by carrying out these operations at each point $x \in W$. e.g.

$$(R + S)_x = R_x + S_x$$

$$(R \otimes S)_x = R_x \otimes S_x$$

Definition A tensor field with no indices is called a *scalar field* (i.e. a real-valued function); a tensor field with one upper index is called a *vector field*; a tensor field with r skew-symmetric lower indices is called a *differential r-form*; a tensor field with two lower indices is called a *metric tensor* if it is symmetric and non-singular at each point.

If $y^1, \ldots, y^n$ is a coordinate system with domain W and S is a tensor field on W with (say) indices of type down-up-down, then

$$S = \alpha_i{}^j{}_k dy^i \otimes \frac{\partial}{\partial y^k} \otimes dy^k$$

where the scalar fields $\alpha_i{}^j{}_k$ are the *components* of S w.r.t. the coordinates y^i.

If f is a scalar field, then df is a differential 1-form. If v is a vector field, then vf is the scalar field defined by:

$$(vf)_x = v_x f = \langle df_x, v_x \rangle = \langle df, v \rangle_x$$

i.e. $vf = \langle df, v \rangle =$ rate of change of f along v.

If $(\cdot|\cdot)$ is a metric tensor on X then for any two vector fields u, v with common domain W we define the scalar field $(u|v)$ by

$$(u|v)_x = (u_x|v_x)_x$$

if v is a vector field then we can lower it's index to get a differential 1-form ω such that

$$\langle \omega, u \rangle = (v|u)$$

Conversely, raising the index of a 1-form gives a vector field. Raising the index of df gives a vector field $\mathrm{grad} f$ called the *gradient of f*:

$$(\mathrm{grad} f|u) = \langle df, u \rangle = uf = \text{rate of change of } f \text{ along } u$$

If $(\cdot|\cdot)$ is positive definite then for $\|u\| = 1$ we have:

$$|(\mathrm{grad} f|u)| \leq \|\mathrm{grad} f\|$$

Thus the maximum rate of change of f is $\|\mathrm{grad} f\|$ and is attained in the direction of $\mathrm{grad} f$.

A metric tensor $(\cdot|\cdot)$ defines a field ds^2 of quadratic forms called the associated *line element* by:

$$ds^2(v) = (v|v) = \|v\|^2 \quad \text{if metric is positive definite}$$

If y^i are coordinates with domain W then, on W:
each vector field $u = \alpha^i \frac{\partial}{\partial y^i}$ with components α^i
each differential r-form $\omega = \omega_{i_1 \ldots i_r} dy^{i_1} \wedge \cdots \wedge dy^{i_r}$
each differential 1-form $\omega = \omega_i dy^i$

$$\langle \omega, u \rangle = \alpha^i \omega_i$$

if f is a scalar field $df = \frac{\partial f}{\partial y^i} dy^i$

$$\langle df, u\rangle = uf = \alpha^i \frac{\partial f}{\partial y^i}$$

if

$$\left(\frac{\partial}{\partial y^i} \middle| \frac{\partial}{\partial y^j} \right) = g_{ij}$$

then

$$(\cdot|\cdot) = g_{ij} dy^i \otimes dy^j$$

and

$$ds^2 = g_{ij} dy^i dy^j$$

grad$f$ has components

$$g^{ij} \frac{\partial f}{\partial y^j}$$

therefore,

$$\text{grad} f = g^{ij} \frac{\partial f}{\partial y^j} \frac{\partial}{\partial y^i}$$

so

$$\|\text{grad} f\| = (\text{grad} f | \text{grad} f) = \langle df, \text{grad} f\rangle = g^{ij} \frac{\partial f}{\partial y^i} \frac{\partial f}{\partial y^j}$$

Example On $\mathbb{R}^3$ with the usual coordinate functions x, y, z the *usual metric tensor* is

$$(\cdot|\cdot) = dx \otimes dx + dy \otimes dy + dz \otimes dz$$

with components

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = g^{ij}$$

$\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ are orthonormal vector fields

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial y} dy$$

and

$$\text{grad} f = \frac{\partial f}{\partial x} \frac{\partial}{\partial x} + \frac{\partial f}{\partial y} \frac{\partial}{\partial y} + \frac{\partial f}{\partial z} \frac{\partial}{\partial z}$$

The line element is

$$ds^2 = (dx)^2 + (dy)^2 + (dz)^2$$

If r, θ, ϕ are spherical polar cordinates on $\mathbb{R}^3$ then

$$\begin{aligned} x &= r \sin\theta \cos\phi \\ y &= r \sin\theta \sin\phi \\ z &= r \cos\theta \end{aligned}$$

therefore,

$$\begin{aligned} ds^2 &= (dx)^2 + (dy)^2 + (dz)^2 \\ &= (\sin\theta\cos\phi dr + r\cos\theta\cos\phi d\theta - r\sin\theta\sin\phi d\phi)^2 \\ &\quad +(\sin\theta\sin\phi dr + r\cos\theta\sin\phi d\theta + r\sin\theta\cos\phi d\phi)^2 \\ &\quad +(\cos\theta dr - r\sin\theta d\theta)^2 \\ &= (dr)^2 + r^2(d\theta)^2 + r^2\sin^2\theta(d\phi)^2 \end{aligned}$$

therefore,

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

and

$$g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2\sin^2\theta} \end{pmatrix}$$

so

$$df = \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial \theta}d\theta + \frac{\partial f}{\partial \phi}d\phi$$

and df has components:

$$\left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \phi}\right)$$

therefore, grad$f$ has components

$$\left(\frac{\partial f}{\partial r}, \frac{1}{r^2}\frac{\partial f}{\partial \theta}, \frac{1}{r^2\sin^2\theta}\frac{\partial f}{\partial \phi}\right)$$

therefore,

$$\text{grad}f = \frac{\partial f}{\partial r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial f}{\partial \theta}\frac{\partial}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial f}{\partial \phi}\frac{\partial}{\partial \phi}$$

7.5 Pull-back, Push-forward

Definition Let X and Y be manifolds and

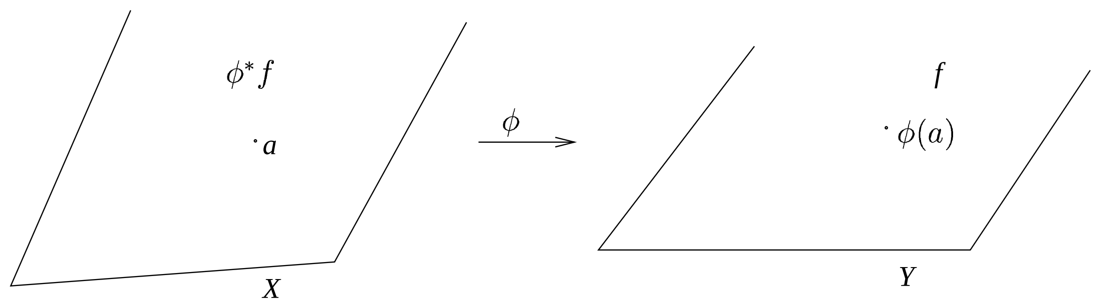
$$X \xrightarrow{\phi} Y$$

a continuous map. Then for each scalar field f on Y we have a scalr field

$$\phi^* f = f \cdot \phi$$

on X (we assume that $\phi^* f$ is C^∞ for each C^∞ f).

$\phi^* f$ is called the *pull-back* of f to X under ϕ.



$$(\phi^* f)(x) = f(\phi(x))$$

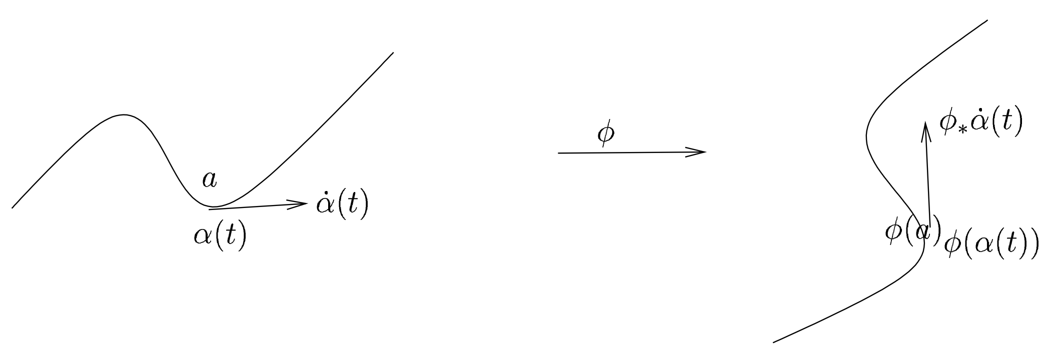
For each $a \in X$ we have a linear operator

$$T_a X \xrightarrow{\phi_*} T_a Y$$

If $v = \dot{\alpha}(t) \in T_a X$ then we define $\phi_* v$ by:

$$[\phi_* v] f = \frac{d}{dt} f(\phi(\alpha(t))) = \dot{\alpha}(t)[f \cdot \phi] = v[f \cdot \phi] = v[\phi^* f]$$

for each scalar field f on Y at $\phi(a)$.



Thus the velocity vector of $\alpha(t)$ pushes forward to the velocity vector of $\phi(\alpha(t))$

Theorem 7.5.1. *Let* $X \xrightarrow{\phi} Y$ *and let* $y^1, \ldots, y^n$ *be coordinates on* X *at* a *and let*

$$\phi^1(x), \ldots, \phi^n(x)$$

be the coordinates of $\phi(x)$ *w.r.t. a coordinate system* $z^1, \ldots, z^n$ *on* Y *at* $\phi(a)$. *Then the push-forward*

$$T_aX \xrightarrow{\phi_*} T_{\phi(a)}Y$$

has matrix

$$\frac{\partial \phi^i}{\partial y^j}(x)$$

Proof.

$$\begin{aligned} i^{\text{th}} \text{ component of } \phi_* \frac{\partial}{\partial y^j}_a &= \left[\phi_* \frac{\partial}{\partial y^j}_a\right] z^i \\ &= \frac{\partial}{\partial y^j}_a [\phi^* z^i] \\ &= \frac{\partial \phi^i}{\partial y^j}(a) \end{aligned}$$

since $\phi^i(x) = z^i(\phi(x))$, so $\phi^i = \phi^* z^i$. □

Thus the matrix of ϕ_* is the same as the matrix of the derivative in the case $\mathbb{R}^n \longrightarrow \mathbb{R}^n$. The push-forward ϕ_* is often called the *derivative* of ϕ at a.

The chain rule for vector spaces

$$(\psi \cdot \phi)'(x) = \psi'(\phi(x))\, \phi'(x)$$

corresponds, for manifolds, to:

Theorem 7.5.2. (Chain rule for maps of manifolds; functorial property of the push-forward)

Let $X \xrightarrow{\phi} Y \xrightarrow{\psi} Z$ *be maps of manifolds. Then*

$$(\psi \cdot \phi)_* = \psi_* \cdot \phi_*$$

Proof. Let $\dot\alpha(t)$ be a tangent vector on X, then:

$$\begin{aligned} (\psi \cdot \phi)_* \dot\alpha(t) &= \text{velocity vector of } \psi(\phi(\alpha(t))) \\ &= \psi_*[\text{velocity vector of } \phi(\alpha(t))] \\ &= \psi_*[\phi_* \dot\alpha(t)] \end{aligned}$$

as required. □

Definition Let $X \xrightarrow{\phi} Y$ be a map of manifolds and ω a tensor field on Y with r lower indices. Then we define the *pull-back* of ω to X under ϕ to be the tensor field $\phi^*\omega$ on X having r lower indices and defined by:

$$(\phi^*\omega)_x[v_1, \ldots, v_r] = \omega_{\phi(x)}[\phi_* v_1, \ldots, \phi_* v_r]$$

all $v_1, \ldots, v_r \in T_x X$, $\forall x \in X$. i.e.

$$(\phi^*\omega)_x = \phi^*[\omega_{\phi(x)}]$$

We note the ϕ^* on tensor fields preserves all the algebraic operations such as additions, tensor products, wedge products.

We have:

Theorem 7.5.3. *if $X \xrightarrow{\phi} Y$ and f is a scalar field on Y, then*

$$\phi^* df = d\phi^* f$$

i.e., the pull-back commutes with differentials

Proof. Let $v \in T_x X$. Then

$$\begin{aligned} \langle(\phi^* df)_x, v\rangle &= \langle(df)_{\phi(x)}, \phi_* v\rangle \\ &= [\phi_* v] f \\ &= v[\phi^* f] \\ &= \langle(d\phi^* f)_x, v\rangle \end{aligned}$$

therefore, $\phi^* df)_x = (d\phi^* f)_x \; \forall x \in X$, and therefore $\phi^* df = d\phi^* f$. □

7.6 Implicit funciton theorem

Theorem 7.6.1 (Implicit Function Theorem). (on the solution spaces of l equations in n variables)

Let $f = (f^1, \ldots, f^l)$ be a sequence of C^r real-valued functions on an open set V in $\mathbb{R}^n$. Let $c = (c^1, \ldots, c^l) \in \mathbb{R}^l$ and let

$$X = \{x \in V : f(x) = c, \, \text{rank} f'(x) = l\}$$

Then for each $a \in X$ we can select $n - l$ of the usual coordinate functions $x^1, \ldots, x^n$:

$$x^{l+1}, \ldots, x^n$$

(say) so that on an open neighbourhood of a in X they form a coordinate system on X. Any two such coordinate systems are C^r-compatible. Thus X is an $(n-l)$-dimensional C^r manifold.

Proof.

$$f' = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^l} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial f^l}{\partial x^1} & \cdots & \frac{\partial f^l}{\partial x^l} & \cdots & \frac{\partial f^l}{\partial x^n} \end{pmatrix}_{l\times n}$$

Let $a \in X$. THen $f'(a)$ has rank l, so the matrix $f'(a)$ has l linearly independent columns: the first l columns (say). Put

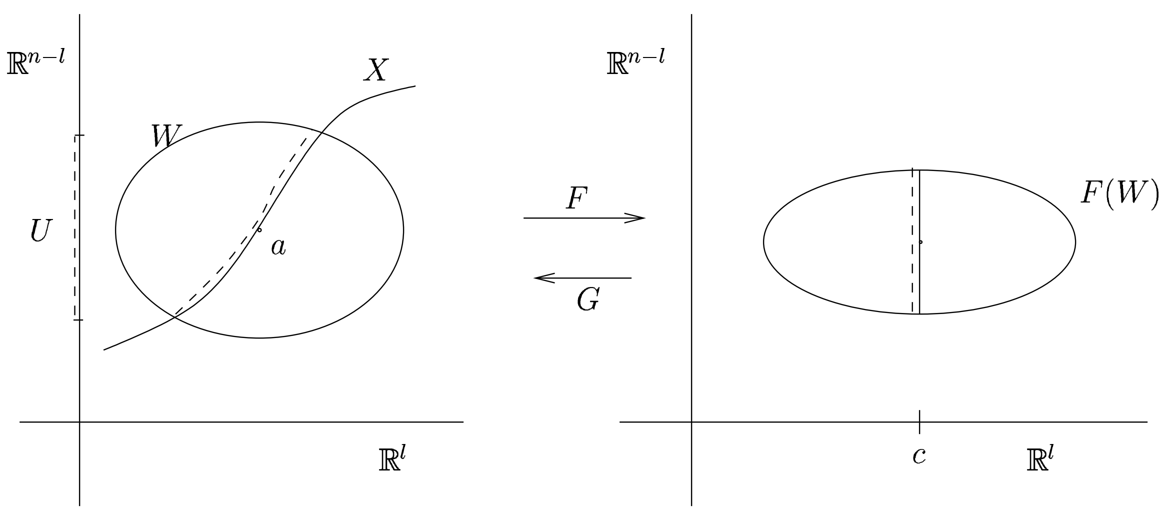
$$F = (f^1, \dots, f^l, x^l + 1, \dots, x^n)$$

then

$$F' = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^l} & 0 & \cdots & \\ \vdots & & \vdots & & & \\ \frac{\partial f^l}{\partial x^1} & \cdots & \frac{\partial f^l}{\partial x^l} & 0 & \cdots & \\ & & & & & \\ 0 & \cdots & 0 & 1 & & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \cdots & 0 & 0 & & 1 \end{pmatrix}_{l\times n}$$

therefore,

$$\det F'(a) = \begin{vmatrix} \frac{\partial f^1}{\partial x^1}(a) & \cdots & \frac{\partial f^1}{\partial x^l}(a) \\ \vdots & & \vdots \\ \frac{\partial f^l}{\partial x^1}(a) & \cdots & \frac{\partial f^l}{\partial x^l}(a) \end{vmatrix} \neq 0$$



By the inverse function theorem, F maps an open neighbourhood W of a onto an open neighbourhood $F(W)$ by a C^r-diffeomorphism with inverse G (say). Note that F, and hence also G leave the coordinates $x^{l+1}, \ldots, x^n$ unchanged.

F maps $W \cap X$ homeomorphically onto $\{c\} \times U$ where U is open in $\mathbb{R}^{n-l}$.

Thus $(x^{l+1}, \ldots, x^n)$ maps $W \cap X$ homeomorphically onto U and is therefore an $(n-l)$-dimensional coordinate system on X with domain $W \cap X$. Also, if

$$G = (G^1, \ldots, G^l, x^{l+1}, \ldots, x^n)$$

on $F(W)$ then

$$\begin{aligned} x^1 &= G^1(c^1, \ldots, c^l, x^{l+1}, \ldots, x^n) \\ &\vdots \\ x^l &= G^l(c^1, \ldots, c^l, x^{l+1}, \ldots, x^n) \end{aligned}$$

on $W \cap X$. Therefore $x^1, \ldots, x^l$ and C^r functions of $x^{l+1}, \ldots, x^n$ on $W \cap X$. Hence any two such coordinate systems on X are C^r compatible. $\square$

Note: the proof shows that if (say) the first l columns of the matrix $\frac{\partial f^i}{\partial x^j}(a)$ are linearly independent, then the l equations in n unknowns

$$\begin{aligned} f^1(x^1, \ldots, x^n) &= c^1 \\ &\vdots \\ f^l(x^1, \ldots, x^n) &= c^l \end{aligned}$$

determine $x^1, \ldots, x^n$ as C^r functions of $x^{l+1}, \ldots, x^n$ on an open neighbourhood of a in the solution space.

Example 1.

$$\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$$

a C^r function.

$$f' = \left(\frac{\partial f}{\partial x^i}\right)_{1\times n} = \left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}\right) = \nabla f$$

are the components of the gradient vector field.

If ∇f is non-zero at each point of the space X of solutions of the equations:

$$f(x^1, \ldots, x^n) = c$$

(one equation in n unknowns) then X is an $(n-1)$-dimensional C^r-manifold.

If (say) $\frac{\partial f}{\partial x^1}(a) \neq 0$ at $a \in X$ then

$$x^2, x^3, \ldots, x^n$$

are coordinates on an open neighbourhood W of a in X and x^1 is a C^r function of $x^2, \ldots, x^n$ on W.

2. $\mathbb{R}^n \supset V \xrightarrow{f^1, f^2} \mathbb{R}$ two C^r functions.

$$\left(\frac{\partial f^i}{\partial x^j}\right) = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \frac{\partial f^2}{\partial x^1} & \cdots & \frac{\partial f^2}{\partial x^n} \end{pmatrix}$$

If the two rows $\nabla f^1, \nabla f^2$ are linearly independent at each point of the space X of solutions of the two equations:

$$f^1(x^1, \ldots, x^n) = c^1$$

$$f^2(x^1, \ldots, x^n) = c^2$$

(two equations in n unknowns) then X is an $(n-2)$-dimensional C^r manifold. If (say)

$$\frac{\partial(f^1, f^2)}{\partial(x^1, x^2)} \neq 0 \qquad \text{at } a \in X$$

then $x^3, x^4, \ldots, x^n$ are coordinates on an open neighbourhood W of a in X and x^1 and x^2 are C^r functions of $x^3, \ldots, x^n$ on W.

7.7 Constraints

If $f^1, \ldots, f^l$ are C^r functions on an open set V in $\mathbb{R}^n$ and $c = (c^1, \ldots, c^n)$ we consider the equations

$$f^1 = c^1, \ldots, f^l = c^l$$

as a system of *constraints* which are satisfied by points on the *constraint manifold*:

$$X = \left\{x \in V : f^1(x) = c^1, \ldots, f^l(x) = c^l, \text{rank}\frac{\partial f^i}{\partial x^j}(x) = l\right\}$$

Let $X \xrightarrow{i} \mathbb{R}^n$ be the *inclusion map* $i(x) = x \, \forall x \in X$. Then for each scalar field f on V:

$$(i^* f)(x) = f(i(x)) = f(x) \qquad \forall x \in X$$

the pull-back i^*f is the restriction of f to X. We call i^*f the *constrained* function F. If ω is a differential r-form on V the $i^*\omega$ is called the *constrained* r-form ω. Similarly if $(\cdot|\cdot)$ is a metric tensor on V and ds^2 the line element the $i^*(\cdot|\cdot)$ is the *constrained* matrix and i^*ds^2 is the *constrained* line element.

For each $a \in X$ the push-forward

$$T_aX \xrightarrow{i_*} T_a\mathbb{R}^n$$

is injective and enables us to identify T_aX with a vector subspace of $T_a\mathbb{R}^n$. The constrained metric and constrained line element are just the restrictions to T_aX of the matrix and line element, for each $a \in X$.

Example Let X be the constraint surface in $\mathbb{R}^3$:

$$f(x, y, z) = c \qquad f\, C^r$$

We have:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \qquad \text{unconstrained}$$

Pulling back to X, where f is constant, we get:

$$0 = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \qquad \text{constrained}$$

If (say) $\frac{\partial f}{\partial z} \neq 0$ at $a \in X$ then, by the implicit function theorem, the constrained functions x, y are coordinates on X in a neighbourhood W of a and constrained z is a C^r function of x, y on W.

$$z = F(x, y)$$

say. Now

$$dz = -\left(\frac{\partial f}{\partial x}\Big/\frac{\partial f}{\partial z}\right)dx - \left(\frac{\partial f}{\partial y}\Big/\frac{\partial f}{\partial z}\right)dy$$

Therefore,

$$\left(\frac{\partial z}{\partial x}\right)_y = \frac{\partial F}{\partial x} = -\frac{\partial f}{\partial x}\Big/\frac{\partial f}{\partial z}$$

$$\left(\frac{\partial z}{\partial y}\right)_x = \frac{\partial F}{\partial y} = -\frac{\partial f}{\partial y}\Big/\frac{\partial f}{\partial z}$$

The usual line element on $\mathbb{R}^3$ is

$$ds^2 = (dx)^2 + (dy)^2 + (dz)^2 \qquad \text{unconstrained}$$

pulling back to X gives:

$$\begin{aligned} ds^2 &= (dx)^2 + (dy)^2 + \left[-\left(\frac{\partial f}{\partial x}\Big/\frac{\partial f}{\partial z}\right)dx - \left(\frac{\partial f}{\partial y}\Big/\frac{\partial f}{\partial z}\right)dy\right]^2 \\ &= \left[1+\left(\frac{\partial f}{\partial x}\Big/\frac{\partial f}{\partial z}\right)^2\right](dx)^2 + 2\frac{\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}}{\left(\frac{\partial f}{\partial y}\right)^2}dx\,dy + \left[1+\left(\frac{\partial f}{\partial y}\Big/\frac{\partial f}{\partial z}\right)^2\right](dy)^2 \end{aligned}$$

The coefficients give the components of the constrained metric on X with respect to the coordinates x and y

Example using spherical polar coordinates on $\mathbb{R}^3$ the usual line element is

$$ds^2 = (dr)^2 + r^2(d\theta)^2 + r^2\sin^2\theta(d\phi)^2$$

Pulling back to the sphere S^2 by the constraint:

$$r = \text{const}$$

we have:

$$ds^2 = r^2(d\theta)^2 + r^2\sin^2(d\phi)^2$$

Thus the constrained line element has components:

$$\begin{pmatrix} r^2 & 0 \\ 0 & r^2\sin^2\theta \end{pmatrix}$$

7.8 Lagrange Multipliers

A scalar field f on a manifold X has a *critical point* $a \in X$ if $df_a = 0$. i.e. $\frac{\partial f}{\partial y^i}(a) = 0$ for a coordinate system y^i at a. (e.g. a local maximum or minimum or saddle point)

Problem given a scalar field F on V open in $\mathbb{R}^n$, to find the critical point of constrained F, where the constraints are:

$$f^1 = c^1, \ldots, f^l = c^l$$

and f^i are C^r functions on V

Method Take $f^1, \ldots, f^l, x^{l+1}, \ldots, x^n$ (say) as coordinates on $\mathbb{R}^3$ as in the proof of the implicit function theorem, so

$$dF = \sum_{i=1}^{l} \frac{\partial f}{\partial f^i}df^i + \sum_{i=l+1}^{n} \frac{\partial f}{\partial x^i}dx^i \qquad \text{unconstrained}$$

Thus

$$dF = 0 + \sum_{i=l+1}^{n} \frac{\partial F}{\partial x^i} dx^i \qquad \text{constrained}$$

This is zero at a critical point a, so $\frac{\partial F}{\partial x^i}(a) = 0$, $i = l+1, \ldots, n$, and hence

$$dF_a = \sum_{i=1}^{l} \frac{\partial F}{\partial f^i}(a) df^i$$

at a critical point a of constrained F. If we put

$$\lambda_i = \frac{\partial F}{\partial f^i}(a) = \text{rate of change of } F \text{ at } a \text{ with respect to the } i^{\text{th}} \text{ constraint}$$

we have:

$$dF = \lambda_1 df^1 + \cdots + \lambda_l df^l$$

at a where the scalars $\lambda_1, \ldots, \lambda_l$ are called *Lagrange multipliers*.

Since dF is a linear combination of $df^1, \ldots, df^l$ at a we can take the wedge product to get the equations:

$$dF \wedge df^1 \wedge \cdots \wedge df^l = 0$$

$$f^1 = c^1, \ldots, f^l = c^l$$

which must hold at any critical point of constrained F.

7.9 Tangent space and normal space

If f is constant on the constraint manifold X then

$$df = 0 \qquad \text{constrained}$$

therefore,

$$\langle df, \dot{\alpha}(t) \rangle = 0$$

for all curves in X. Hence the system of l linear equations

$$df^1 = 0, \ldots, df^l = 0$$

give the tangent space to X at each point. Also if $(\cdot|\cdot)$ is a metric tensor with domain V then

$$(\mathrm{grad} f | \dot{\alpha}(t)) = \langle df, \dot{\alpha}(t) \rangle = 0$$

for all curves in X. Therefore grad$f$ is orthogonal to the tangent space to X at each point. Hence

$$\text{grad}f^1, \ldots, \text{grad}f^l$$

is a basis for the normed space to X at each point.

At a critical point of constrained F we have:

$$dF = \lambda_1 df^1 + \cdots + \lambda_l df^l$$

and raising the index gives:

$$\text{grad}F = \lambda_1 \text{grad}f^1 + \cdots + \lambda_l \text{grad}f^l$$

thus grad$F$ is normal to the constraint manifold X at each critical point of constrained F.

7.10 ?? Missing Page

1. $\Longrightarrow$ ω is closed.

   However ω is not exact because it's integral around circle $\alpha(t) = (\cos t, \sin t)$ $0 \leq t \leq 2\pi$ is

   $$\int_\alpha \omega = {}^{2\pi}_{0} \frac{[\cos t \cos t - (\sin t)(-\sin t)]}{\cos^2 t + \sin^2 t} dt = \int_0^{2\pi} dt = 2\pi \neq 0$$

   Note: on $\mathbb{R}^2 - \{\text{negative } x\text{-axis}\}$ we have:

   $$\begin{aligned} x &= r\cos\theta \qquad -\pi < \theta < \pi \\ y &= r\sin\theta \end{aligned}$$

   therefore,

   $$\omega = \frac{r\cos\theta[\sin\theta\, dr + r\cos\theta\, d\theta] - r\sin\theta[\cos\theta\, dr - r\sin\theta\, d\theta]}{r^2\cos^2\theta + r^2\sin^2\theta} = d\theta$$

   so. $\int_\alpha \omega$ is path-independent provided α does not cross the -ve x-axis. $\int_\alpha \omega = \theta(b) - \theta(a)$, and ω is called the *angle-form*

2. Let ω be a differential n-form with domain V open in $\mathbb{R}^n$

   $$\omega = f(x^1, \ldots, x^n) dx^1 \wedge \cdots \wedge dx^n \quad \text{(say)} \quad x^i \text{ usual coord}$$

   then we define:

   $$\int_V \omega = \int_V f(x_1, \ldots, x_n) dx_1\, dx_2 \ldots dx_n \qquad \text{Lebesgue integral } x_i \textit{ dummy}$$

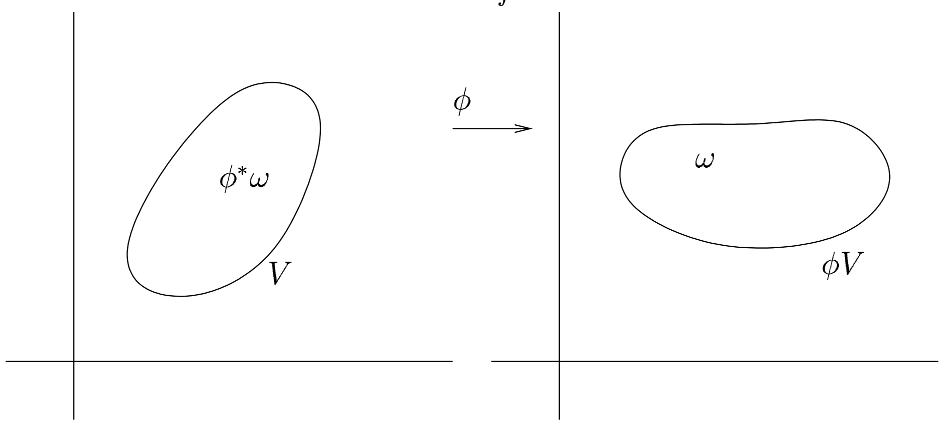
7.11 Integral of Pull-back

Theorem 7.11.1. *Let $V \xrightarrow{\phi} \phi(V)$ be a C^1 diffeomorphism of open V in $\mathbb{R}^n$ onto open $\phi(V)$ in $\mathbb{R}^n$, with* $\det \phi' > 0$. *Let ω be an n-form on $\phi(V)$. Then*

$$\int_V \phi^*\omega = \int_{\phi(V)} \omega$$

[so writing $\langle \omega, V\rangle$ to denote integral of ω over V we have: $\langle \phi^*\omega, V\rangle = \langle \omega, \phi V\rangle$ adjoint]

Proof.



Let

$$\begin{aligned} \omega &= f(x^1, \ldots, x^n) dx^1 \wedge \cdots \wedge dx^n \\ \phi &= (\phi^1, \ldots, \phi^n) \end{aligned}$$

i.e. $\phi^i(x) = x^i(\phi(x))$ i.e. $\phi^i = x^i \cdot \phi = \phi^* x^i$.

Then

$$\begin{aligned} \phi^*\omega &= f(\phi^1, \ldots, \phi^n) d\phi^1 \wedge \cdots \wedge d\phi^n \\ &= f(\phi(x)) \tfrac{\partial \phi^1}{\partial x^{i_1}} dx^{i_1} \wedge \cdots \wedge \tfrac{\partial \phi^n}{\partial x^{i_n}} dx^{i_n} \\ &= f(\phi(x)) \det\left(\tfrac{\partial \phi^i}{\partial x^j}(x)\right) dx^1 \wedge \cdots \wedge dx^n \end{aligned}$$

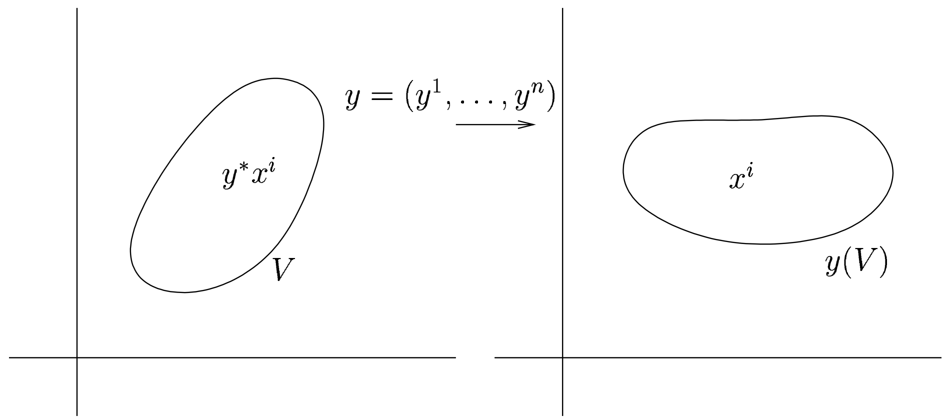
therefore,

$$\int_V \phi^*\omega = \int_V f(\phi(x)) \det \phi'(x) dx_1\, dx_2\, \ldots dx_n = \int_{\phi(V)} f(x)\, dx$$

by the general change of variable theorem for multiple integrals (still to be proved). $\square$

Corollary 7.11.2. *Let $y^1, \ldots, y^n$ be positively oriented coordinates with domain V open in $\mathbb{R}^n$. Then*

$$\underbrace{\int_V f(y^1, \ldots, y^n) dy^1 \wedge \cdots \wedge dy^n}_{\textit{non-dummy } y^i \textit{ with wedge}} = \underbrace{\int_{y(V)} f(y_1, \ldots, y_n) dy_1\, dy_2 \ldots dy_n}_{\textit{Lebesgue integral over } y(V).\ y_1, \ldots, y_n \textit{ dummy. No wedge}}$$



Proof.

$$y^*x^i = x^i \cdot y = y^i$$

therefore,

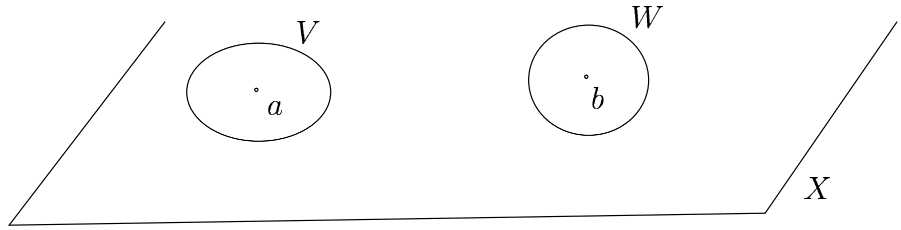
$$\begin{aligned}
\text{LHS} &= \textstyle\int_V y^*[f(x^1, \ldots, x^n)\, dx^1 \wedge cdots \wedge dx^n\,] \\
&= \textstyle\int_{y(V)} f(x^1, \ldots, x^n)\, dx^1 \wedge \cdots \wedge dx^n \\
&= \textstyle\int_{y(V)} f(x_1, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n \qquad x_1, \ldots, x_n \text{ dummy} \\
&= \text{RHS}
\end{aligned}$$

□

7.12 integral of differential forms

To define $\int_X \omega$ where ω is an n-form and X is an n-dimensional manifold (e.g. ω is a 2-form, X a surface) we need some topological notions:

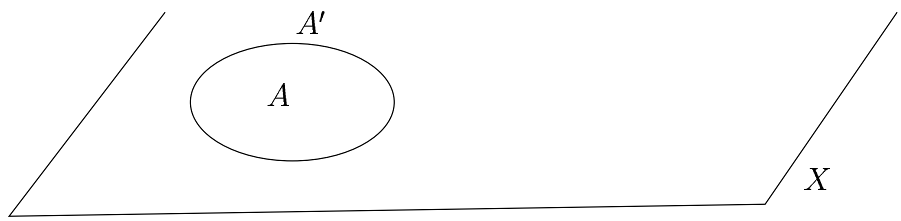
1. A topological space X is called *Hausdorff* if, for each $a, b \in X$, $a \neq b$, $\exists$ open disjoint V, W s.t. $a \in V$, $b \in W$.



2. A set $A \subset X$ is called *closed in* X if it's complement in X

$$A' = \{x \in X : x \notin Z\}$$

is open in X.

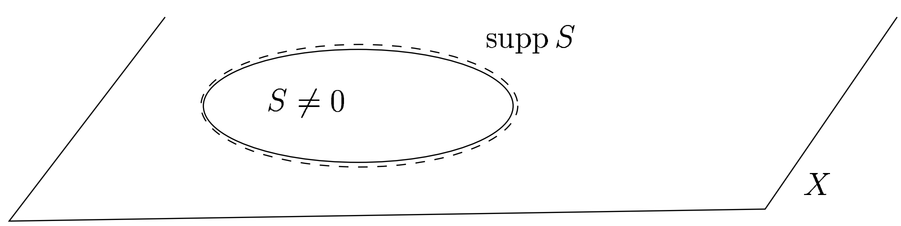


3. If $A \subset X$ then the intersection of all the closed subsets of X which contain A is denoted $\bar{A}$ and is called the *closure of* A *in* X; $\bar{A}$ is the smallest closed subset of X which contains A.

4. Let S be a tensor field on a manifold X. THe closure in X of the set

$$\{x \in X : S_x \neq 0\}$$

is called the *support of* S, denoted $\operatorname{supp} S$.



Theorem 7.12.1. *Let X be a Hausdorff manifold and let $A \subset X$ be compact. Then $\exists$ scalar fields*

$$F_1, \ldots, F_k$$

on X s.t.

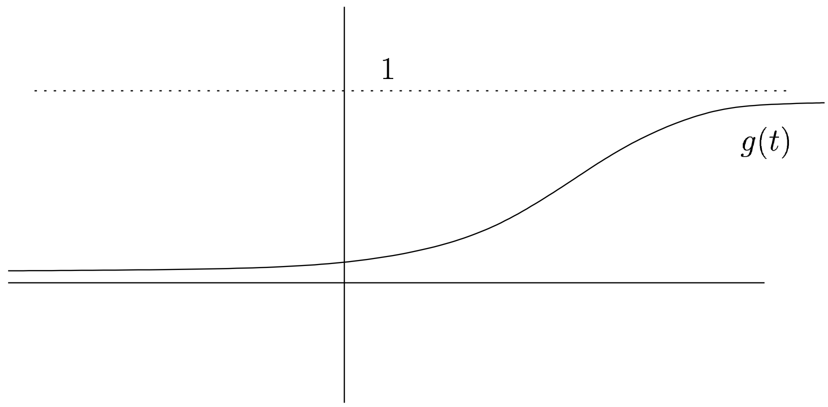
1. $F_i \geq 0$

2. $F_1 + \cdots + F_k = 1$ *on* A *(*partition of unity*)*

3. *each* $\operatorname{supp} F_i$ *is contained in the domain* V_i *of a coordinate system on* X*.*

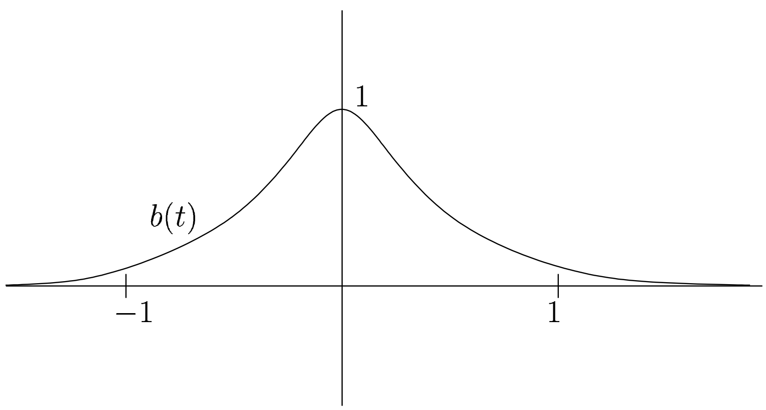
Proof. (sketch) Let

$$g(t) = \left\{ \begin{array}{ll} e^{-\frac{1}{t}} & t > 0 \\ 0 & t \leq 0 \end{array} \right\}$$



let

$$b(t) = \frac{g[1-t^2]}{g(1)}$$



b is C^∞, $\operatorname{supp} b = [-1, 1]$, $0 \leq b \leq 1$. (b is for bump).

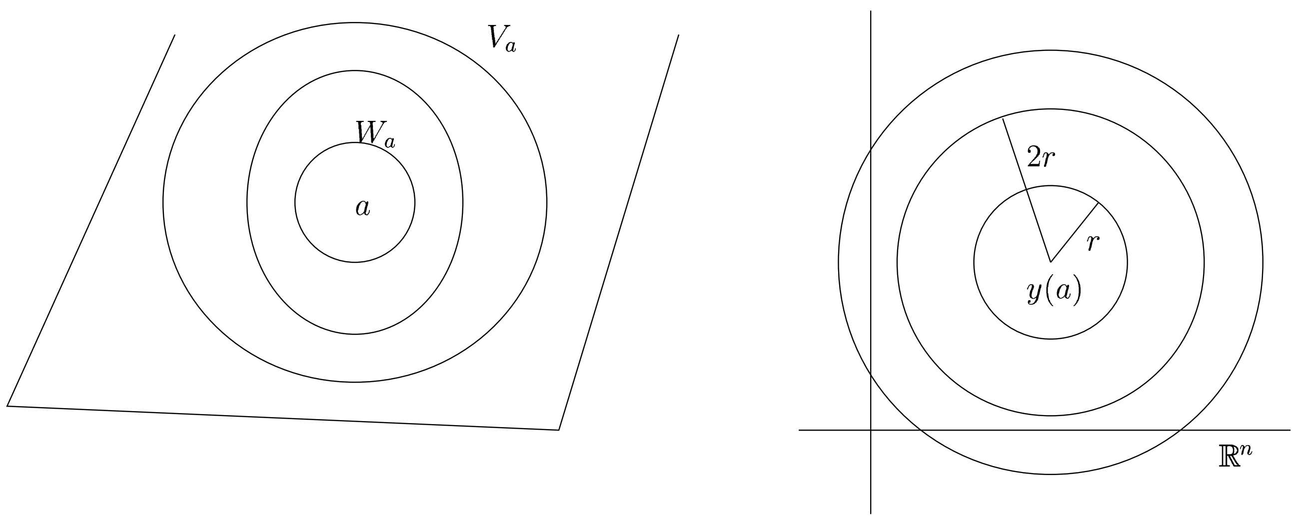
Let $a \in A$. Pick a coordinate system y on X at a with domain V. Pick a ball centre $y(a)$ radius $2r > 0$ contained in $y(V)$. Put

$$h(x) = \left\{ \begin{array}{ll} b\left(\frac{||y(x)-y(a)||}{r}\right) & \text{if } x \in V \\ 0 & \text{if } x \notin V \end{array} \right\}$$

then h is C^∞, $\operatorname{supp} h \subset V$, $0 \le h \le 1$, $h(a) = 1$; 'bump' at a.

Let $W = \{x \in X : h(x) > 0\}$. then W is an open neighbourhood of a and $h > 0$ on W.

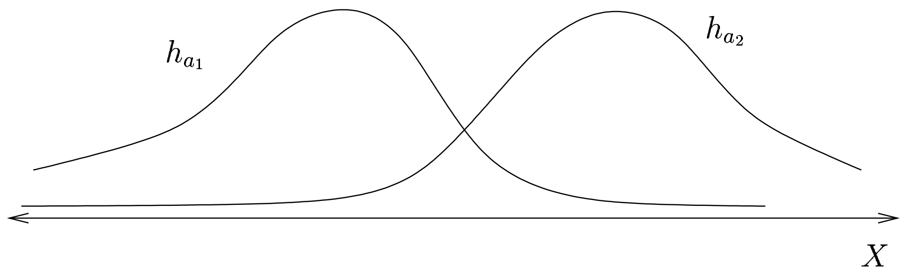
Thus, for each $a \in A$ we have an open neighbourhood W_a of a and a scalar field f_a s.t. $h_a > 0$ on W_a, h_a is C^∞, $\operatorname{supp} h_a \subset$ a coord domain V_a, $0 \le h_a \le 1$.

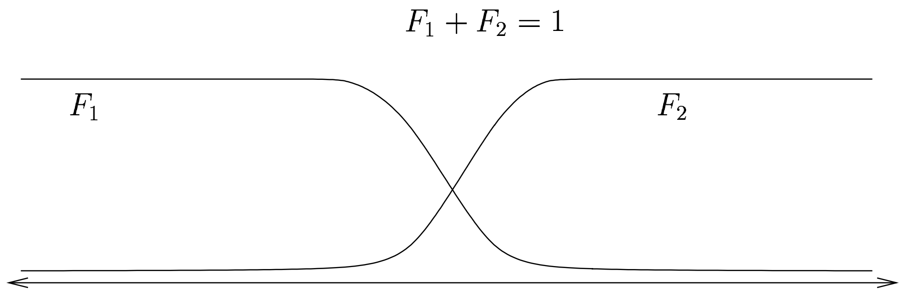


h_a is a bump at a, for each $a \in A$. Since A is compact we can select a finite number of points $a_1, \ldots, a_k$ such that $W_{a_1}, \ldots, W_{a_k}$ cover A. THen put

$$F_i = \left\{ \begin{array}{ll} \frac{h_{a_i}}{h_{a_1} + \cdots + h_{a_k}} & \text{if} h_{a_i} \neq 0 \\ 0 & \text{if} h_{a_i} = 0 \end{array} \right\}$$

to get the required scalar fields $F_1, \ldots, F_k$.





□

7.13 orientation

Definition Let $y^1, \ldots, y^n$ with domain V, $z^1, \ldots, z^n$ with domain W be two coordinate systems on a manifold X. Then they have the *same orientation* if

$$\frac{\partial(y^1, \ldots, y^n)}{\partial(z^1, \ldots, z^n)} = \det \frac{\partial y^i}{\partial z^j} > 0$$

on $V \cap W$.

Since

$$\frac{\partial}{\partial z^j} = \frac{\partial y^i}{\partial z^j} \frac{\partial}{\partial y^i}$$

this means that $\frac{\partial}{\partial z^1}_a, \cdots, \frac{\partial}{\partial z^n}_a$ has the same orientation in $T_a X$ as $\frac{\partial}{\partial y^1}_a, \cdots, \frac{\partial}{\partial y^n}_a$ each $a \in V \cap W$.

We call X *oriented* if a family of mutually compatible coordinate systems is given on X, whose domains cover X and any two of which have the same orientation.

Example $x = r\cos\theta$, $y = r\sin\theta$. Then $\frac{\partial(x,y)}{\partial(r,\theta)} = r > 0$. Therefore x, y and r, θ have the same orientation.

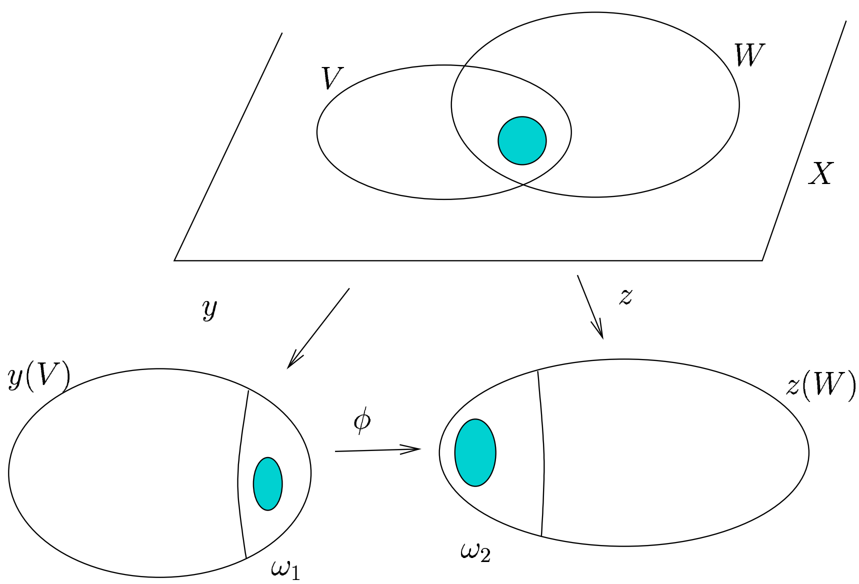
Definition Let ω be a differential n-form with compact support on an oriented n-dimensional Hausdorff manifold X. We define

$$\int_X \omega$$

the *integral of* ω over X as follows:

1. Suppose $\operatorname{supp}\omega \subset V$ where V is the domain of positively oriented coordinates $y^1, \ldots, y^n$ and $\omega = f(y^1, \ldots, y^n)\, dy^1 \wedge \cdots \wedge dy^n$ on V. Then define:

$$\int_x \omega = \int_{y(V)} f(y_1, \ldots, y_n)\, dy_1 \ldots dy_n \qquad \text{dummy } y_i$$



This is independent of the choice of coordinates since if $\operatorname{supp}\omega \subset W$ where W is the domain of positively oriented coordinates $z^1, \ldots, z^n$ then

$$\begin{aligned} \omega &= f(y^1, \ldots, y^n)\, dy^1 \wedge \cdots \wedge dy^n &&= g(z^1, \ldots, z^n)\, dz^1 \wedge \cdots \wedge dz^n \quad \text{(say) on } V \cap W \\ &= y^*[f(x^1, \ldots, x^n)\, dx^1 \wedge \cdots \wedge dx^n &&= z^*[g(x^1, \ldots, x^n)\, dx^1 \wedge \cdots \wedge dx^n \\ &= y^*\omega_1 &&= z^*\omega_2 \end{aligned}$$

(say). So $\omega_1 = \phi^*\omega 2$ where $\phi = z \cdot y^{-1} : y(V \cap W) \longrightarrow z(V \cap W)$. Therefore,

$$\begin{aligned}
\int_{y(V)} f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n &= \int_{y(V)} \omega_1 = \int_{y(V \cap W)} \omega_1 = \int_{z(V \cap W)} \omega_2 \\
&= \int_{z(W)} \omega_2 = \int_{z(W)} g(x_1, \ldots, x_n)\, dx_1 \ldots dx_n
\end{aligned}$$

as required.

2. Choose a partition of unity

$$F_1 + \cdots + F_k = 1$$

on $\operatorname{supp}\omega$. Put $\omega_i = F_i\omega$. Then

$$\omega_1 + \cdots + \omega_k = \omega$$

and $\operatorname{supp}\omega_i \subset \operatorname{supp} F_i \subset V_i$ where V_i is the domain of a coordinates system. Define

$$\int_X \omega = \int_X \omega_1 + \cdots + \int_X \omega_k$$

using (1.). This is independent of the choice of partition of unity since if

$$G_1 + \cdots + G_l = 1 \qquad \text{on } \operatorname{supp}\omega$$

then

$$\begin{aligned}
\sum_{j=1}^{l} \int_X G_j\omega &= \sum_{j=1}^{l} \int_X \sum_{i=1}^{k} F_i G_j \omega \\
&= \sum_{j=1}^{l} \sum_{i=1}^{k} \int_X F_i G_j \omega \\
&\overset{\text{similarly}}{=} \sum_{i=1}^{k} \int_X F_i \omega
\end{aligned}$$

Definition If $A \subset X$ is a Borel set we define

$$\int_A \omega = \int_X \chi_A \omega$$

Example to find the area of the surface X:

$$x^2 + y^2 + z = 2, \qquad z > 0$$

We have:

$$2x\, dx + 2y\, dy + dz = 0$$

constrained to X. Therefore,

$$\begin{aligned} ds^2 &= (dx)^2 + (dy)^2 + (2x\,dx + 2y\,dy)^2 \qquad \text{constrained to } X \\ &= (1+4x^2)(dx)^2 + 8xy\,dx\,dy + (1+4y^2)(dy)^2 \end{aligned}$$

$$g_{ij} = \begin{pmatrix} 1+4x^2 & 4xy \\ 4xy & 1+4y^2 \end{pmatrix}$$

therefore, $\sqrt{g} = \sqrt{\det g_{ij}} = \sqrt{1+4x^2+4y^2}$. So, area element is $\sqrt{1+4x^2+4y^2}\,dx \wedge dy$. Therefore:

$$\begin{aligned} \text{area} &= \int \sqrt{1+4x^2+4y^2}\,dx \wedge dy \\ &= \int \sqrt{1+4r^2}r\,dr \wedge d\theta \\ &= \int_0^{2\pi} \left[\int_0^{\sqrt{2}} r\sqrt{1+4r^2}dr\right] d\theta \\ &= 2\pi \left[\tfrac{2}{3}\,\tfrac{1}{8}(1+4r^2)^{\frac{3}{2}}\right]_0^{\sqrt{2}} \\ &= \tfrac{\pi}{6}[27-1] \\ &= \tfrac{13}{3}\pi \end{aligned}$$

Chapter 8

Complex Analysis

8.1 Laurent Expansion

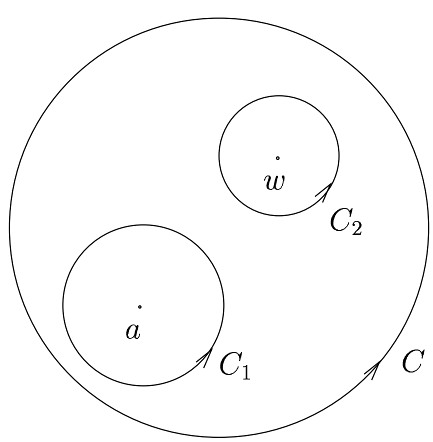
Theorem 8.1.1. *Let f be holomorphic on $V - \{a\}$ where V is open and $a \in V$. Let C be a circle, centre a, radius r s.t. C and it's interior is contained in V. Then $\exists \{c_n\}$ $n = 0, \pm 1, \pm 2, \ldots \in \mathbb{C}$ s.t.*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n \qquad \text{in } 0 < |z-a| < r$$

The RHS is called the Laurent series *of f about a. The coefficient c_n are uniquely determined by:*

$$c_n = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z-a)^{n+1}}\, dz$$

Proof. Let w be inside C. Choose circles C_1, C_2 as shown with centres a, w:



Then

$$\begin{aligned}
f(w) &= \tfrac{1}{2\pi i}\int_{C_2}\tfrac{f(z)}{z-w}\,dz \\
&= \tfrac{1}{2\pi i}\int_{C}\tfrac{f(z)}{z-a}\,d-\tfrac{1}{2\pi i}\int_{C_1}\tfrac{f(z)}{z-w}\,dz \\
&= \tfrac{1}{2\pi i}\int_{C}\tfrac{f(z)}{(z-a)-(w-a)}\,dz+\tfrac{1}{2\pi i}\int_{C_1}\tfrac{f(z)}{(w-a)-(z-a)}\,dz \\
&= \tfrac{1}{2\pi i}\int_{C}\tfrac{f(z)}{(z-a)\left[1-\frac{w-a}{z-a}\right]}\,dz+\tfrac{1}{2\pi i}\int_{C_1}\tfrac{f(z)}{(w-a)\left[1-\frac{z-a}{w-a}\right]}\,dz \\
&= \tfrac{1}{2\pi i}\int_{C}\tfrac{f(z)}{(z-a)}\textstyle\sum_{n=0}^{\infty}\left(\frac{w-a}{z-a}\right)^n\,dz+\tfrac{1}{2\pi i}\int_{C_1}\tfrac{f(z)}{(w-a)}\textstyle\sum_{n=0}^{\infty}\left(\frac{z-a}{w-a}\right)^n\,dz \\
&= \textstyle\sum_{n=0}^{\infty}(w-a)^n\underbrace{\frac{1}{2\pi i}\int_C\frac{f(z)}{(z-a)^{n+1}}\,dz}_{\text{+ve powers }(w-a)}+\textstyle\sum_{n=0}^{\infty}(w-a)^{-n-1}\underbrace{\frac{1}{2\pi i}\int_{C_!}\frac{f(z)}{(z-a)^{-n}}\,dz}_{\text{-ve powers }(w-a)}
\end{aligned}$$

as required. □

Definition If f is holomorphic in $V-\{a\}$ anf

$$f(z)=\sum_{n=-\infty}^{\infty}c_n(z-a)^n=\underbrace{\cdots+\frac{c_{-2}}{(z-a)^2}+\frac{c_{-1}}{z-a}}_{p(z)}+c_0+c_1(z-a)+\cdots$$

is the Laurent series of f at a. $p(z)=\sum_{n=-\infty}^{-1}c_n(z-a)^n$ is called the *principal part* of f at a. $p(z)$ is holomorphic on $\mathbb{C}-\{a\}$.

$f(z)-p(z)$ is holomorphic in V (defining it's value at 0 to be c_0).

For any closed curve α in $\mathbb{C}-\{a\}$:

$$\begin{aligned}
\int_\alpha p(z)\,dz &= \int_\alpha\left[\cdots+\tfrac{c_{-2}}{(z-a)^2}+\tfrac{c_{-1}}{z-a}\right]\,dz \\
&= \cdots+0+c_1\int_\alpha\tfrac{dz}{z-a} \\
&= 2\pi i\,\mathrm{Res}(f,a)\,W(\alpha,a)
\end{aligned}$$

where $W(\alpha,a)=\frac{1}{2\pi i}\int_\alpha\frac{dz}{z-a}$ is the *winding number* of α about a. and $\mathrm{Res}(f,a)=c_{-1}$ is the *residue of f at a*

8.2 Residue Theorem

Theorem 8.2.1 (Residue Theorem). *Let f be holomorphic on $V-\{a_1,\ldots,a_n\}$ where $a_1,\ldots,a_n$ are distinct points in a star-shaped (or contractible) open set V. Then for any closed curve α in $V-\{a_1,\ldots,a_n\}$ we have:*

$$\int_\alpha f(z)\,dz = 2\pi i \sum_{j=1}^{n} \mathrm{Res}(f,a_j)\,W(\alpha,a_j)$$

Proof. Let $p_1,\ldots,p_n$ be the principal parts of f at $a_1,\ldots,a_n$ respectively. Then $f-(p_1+\cdots+p_n)$ is holomorphic on V. Therefore, $\int_\alpha [f-(p_1+\cdots+p_n)]\,dz = 0$. So,

$$\int f(z)\,dz = \sum_{j=1}^{n}\int_\alpha p_j(z)\,dz = 2\pi i\sum_{j=1}^{n}\mathrm{Res}(f,a_j)W(\alpha,a_j)$$

as required. □

Definition If f is holomorphic on a neighbourhood of a, excluding possibly a itself, and if the Laurent expansion has at most a finite number of negative powers:

$$\begin{aligned} f(z) &= \textstyle\sum_{r=n}^{\infty} c_r(z-a)^r \\ &= c_n(z-a)^n + c_{n+1}(z-a)^{n+1}+\cdots \qquad c_n \neq 0 \\ &= (z-a)^n\,[c_n + c_{n+1}(z-a)+\cdots] \\ &= = (z-a)^n f_1(z) \qquad f_1 \text{ holomorphic on a nbd of } a \text{ and } f_1(a)\neq 0 \end{aligned}$$

then we say that f has *order* n at a.

e.g. order 2:

$$f(z) = c_2(z-a)^2 + c_3(z-a)^3+\cdots \qquad c_2\neq 0$$

order -3:

$$f(z) = \frac{c_{-2}}{(z-a)^3}+\frac{c_{-2}}{(z-a)^2}+\frac{c_1}{(z-a)}+c_0+c_1(z-a)+\cdots \quad c_{-3}\neq 0$$

If $n>0$ we call a a *zero* of f.
If $n<0$ we call a a *pole* of f.
$n=1$: simple zero; $n=2$: double zero.

$n=-1$: simple pole; $n=-2$: double pole.

If f has order n and g has order m:

$$f(z) = (z-a)^n f_1(z) \quad f_1(a) \neq 0$$

$$g(z) = (z-a)^n g_1(z) \quad g_1(a) \neq 0$$

then:

$$f(z)\,g(z) = (z-a)^{n+m} f_1(z)\,g_1(z) \qquad fg \text{ has order } n+m$$

$$\tfrac{f(z)}{g(z)} = (z-a)^{n-m} \tfrac{f_1(z)}{g_1(z)} \qquad \tfrac{f}{g} \text{ has order } n-m$$

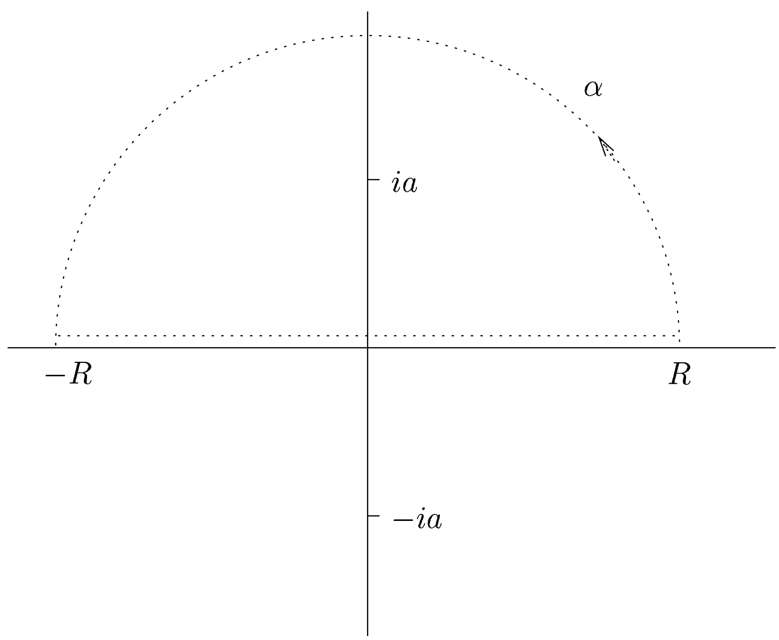
The residue theorem is very useful for evaluating integrals:

Example to evaluate $\int_0^\infty \frac{x \sin x}{x^2+a^2}$, $a>0$ we put

$$f(z) = \tfrac{z\,e^{iz}}{z^2+a^2} \qquad [e^{iz} \text{ is easier to handle than } \sin z]$$

$$= \tfrac{z\,e^{iz}}{(z-ia)(z+ia)}$$

Simple poles at $ia, -ia$.



Choose a closed contour that goes along x-axis $-R$ to R (say) then loops around a pole, upper semi-circle α (say).

$$\int_{-R}^{R} \frac{x\,e^{ix}}{x^2+a^2}\,dx + \int_\alpha \frac{z\,e^{iz}}{z^2+a^2}\,dz = 2\pi i\,\mathrm{Res}(f, ia)$$

1. if $f(z) = \frac{c}{z-ia} + c_0 + c_1(z-ia) + \cdots$ on neighbourhood of ia then $(z-ia)f(z) = c + c_0(z-ia) + c_1(z-ia)^2 + \cdots$ on neighbourhood of ia. Then $\lim_{z \longrightarrow ia}(z-ia)f(z) = c = \mathrm{Res}(f, ia)$. Therefore

$$\begin{aligned} \mathrm{Res}(f, ia) &= c &&= \lim_{z \longrightarrow ia} \tfrac{f(z)}{z-ia} \\ &= \lim_{z \longrightarrow ia} \tfrac{z\, e^{iz}}{z+ia} &&= \tfrac{ia}{2ia} e^{-a} \\ &= \tfrac{1}{2} e^{-a} \end{aligned}$$

2. $\alpha(t) = Re^{it}\ 0 \leq t \leq \pi\ = R(\cos t + i \sin t)$. Therefore

$$\begin{aligned} \left| \int_\alpha \tfrac{z\, e^{iz}}{z^2+a^2}\, dz \right| &= \left| \int_0^\pi \tfrac{Re^{it} e^{iR(\cos t + i \sin t)} iRe^{it}}{R^2 e^{2it} + a^2}\, dt \right| \\ &\leq \tfrac{R^2}{R^2 - a^2} \int_0^\pi e^{-R \sin t}\, dt \end{aligned}$$

which $\longrightarrow 0$ as $R \longrightarrow \infty$ since

$$\lim_{R \to \infty} \int_0^\pi e^{-R \sin t}\, dt \overset{\mathrm{DCT}}{=} \int_0^\pi \lim_{R \to \infty} e^{-R \sin t}\, dt = 0$$

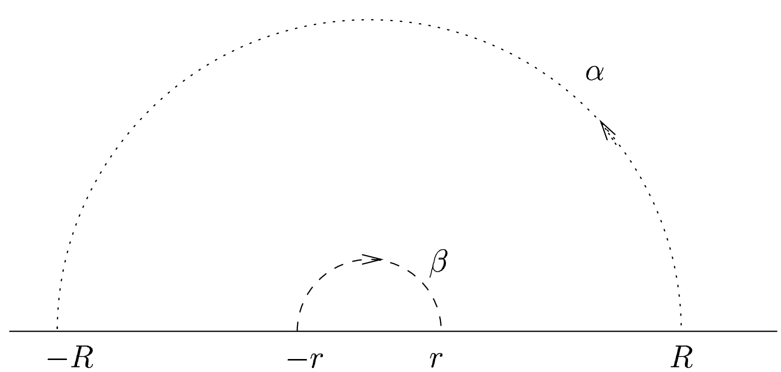
Therefore

$$\int_{-\infty}^{\infty} \frac{x(\cos x + i \sin x)}{x^2 + a^2}\, dx + 0 = 2\pi i \frac{e^{-a}}{2}$$

so

$$\int_{-\infty}^{\infty} \frac{x \sin x}{x^2 + a^2}\, dx = \pi e^{-a}$$

Example to calculate $\int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx$ put $f(z) = \frac{e^{iz}}{z}$ holomorphic except for simple pole at $z = 0$.



Choose a closed contour along x-axis from $-R$ to $-r$, loops around 0 by $\beta(t) = re^{it}$ then r to R then back along $\alpha(t) = Re^{it}$.

$$\int_{-R}^{-r} \frac{e^{ix}}{x}\,dx + \int_\beta \frac{e^{iz}}{z}\,dz + \int_r^R \frac{e^{ix}}{x}\,dx + \int_\alpha \frac{e^{iz}}{z}\,dz = 0$$

1. $\text{Res}(f, 0) = \lim_{z\to 0}(z-0)f(z) = \lim_{z\to 0} e^{iz} = 1$

2. $$\begin{aligned} \left|\int_\alpha \tfrac{e^{iz}}{z}\,dz\right| &= \left|\int_0^\pi \tfrac{e^{iR(\cos t + i\sin t)} iRe^{it}}{Re^{it}}\,dt\right| \\ &\le \int_0^\pi e^{-R\sin t}\,dt \end{aligned}$$

   which $\longrightarrow 0$ as $R \longrightarrow \infty$.

3. $\frac{e^{iz}}{z} = \frac{1}{z} + g(z)$, g holomorphic in neighbourhood of 0. Therefore

   $$\int_\beta \frac{e^{iz}}{z}\,dz = \int_\beta \frac{dz}{z} + \int_\beta g(z)\,dz$$

   Let $\sup|g(z)| = M$ (say on a closed ball centre 0 redius $\delta > 0$ (say).

   $$\left|\int_\beta g(z)\,dz\right| \le M\pi r \longrightarrow 0 \text{ as } r \longrightarrow 0$$

   if $r \le \delta$

   $$\int_\beta \frac{dz}{z} = -\int_0^\pi \frac{ire^{it}}{re^{it}}\,dt = -\int_0^\pi i\,dt = -i\pi \quad \forall r$$

   therefore

   $$\lim_{r\longrightarrow 0}\int_\beta \frac{e^{iz}}{z}\,dz = -i\pi$$

   therefore

   $$\int_{-\infty}^0 \frac{e^{ix}}{x}\,dx - i\pi + \int_0^\infty \frac{e^{ix}}{x}\,dx + 0 = 0$$

   so

   $$\int_{-\infty}^\infty \frac{\cos x + i\sin x}{x}\,dx = i\pi$$

   and

   $$\int_{-\infty}^\infty \frac{\sin x}{x}\,dx = \pi$$

8.3 Uniqueness of analytic continuation

Theorem 8.3.1 (Uniqueness of analytic continuation). *Let f, g be holomorphic on a connected open set V. Let $a \in V$ and let $\{z_k\}$ be a sequence in V $(\neq a)$ converging to a s.t.*

$$f(z_k) = g(z_k) \qquad \forall k$$

Then $f = g$ on V.

Proof. Put $F = f - g$, so $F(z_k) = 0$. To show $F = 0$ on V.

1. F is holomorphic at a. Therefore $F(z) = b_0 + b_1(z-a) + b_2(z-a)^2 + \cdots$ on $|z-a| < R$ (say). $F(a) = \lim F(z_k) = 0$. Therefore $b_0 = 0$.

   Suppose we know that $b_0, b_1, \ldots, b_{m-1}$ are all zero.

   $$\begin{aligned} F(z) &= (z-a)^m[b_m + b_{m+1}(z-a) + b_{m+2}(z-a)^2 + \cdots] \\ &= (z-a)^m F_m(z) \end{aligned}$$

   (say). $F(z_k) = 0$, so $F_m(z_k) = 0$ and $F_m(a) = 0$. Therefore $b_m = 0$. $b_r = 0 \forall r$ and $F = 0$ on an open ball centre a.

2. Put $V = W \cup W'$ where $W = \{z \in V : F = 0 \text{ on an open ball centre } z\}$ and $W' = V - W$. Then W is open, and $a \in W$ by 1. Therefore W is non-empty. Suppose $c \in W'$ and c is a non-interior point of W'. Then each integer $r > 0 \exists\, w_r \in W$ s.t. $|w_r - c| < \frac{1}{r}$, $F(w_r) = 0$ and $\lim w_r = c$. Therefore $F = 0$ on an open ball centre c by 1., and $c \in W$. Therefore W' is empty.

□

Chapter 9

General Change of Variable in a multiple integral

9.1 Preliminary result

Theorem 9.1.1. *Let* $\mathbb{R}^n \supset V \xrightarrow{\phi} \mathbb{R}^n$ *be* C^1, V *open,* $a \in V$, $\det \phi'(a) \neq 0$. *Then*

$$\lim \frac{m(\phi B)}{m(B)} = |\det \phi'(a)|$$

where the limit is taken over cubes B *containing* a *with radius* $B \longrightarrow 0$.

Proof. Let $\|\cdot\|$ be the sup norm on $\mathbb{R}^n$.

$$\|(\alpha_1, \ldots, \alpha_n)\| = \max(|\alpha_1|, \ldots, |\alpha_n|)$$

so a ball, radius r is a cube, with side $2r$.

Let $0 < \epsilon < 1$. Put $T(x) = [\phi'(x)]^{-1}$. Fix a closed cube J containing a and put $k = \sup_{x \in J} \|T(x)\|$. Choose $\delta > 0$ s.t. $\|\phi'(x) - \phi'(y)\| \leq \frac{\epsilon}{k}$ all $\|x - y\| \leq 2\delta$; $x, y \in J$.

If B is a cube $\subset J$ containing a of radius $\leq \delta$. and centre c (say). Consider:

$T(c)\phi$ has derivative $T(c)\phi'(x)$ which equals 1 at $x = c$ and

$$\begin{aligned} \|T(c)\phi'(x) - 1\| &= \|T(c)\phi'(x) - T(c)\phi'(c)\| \\ &\leq \|T(c\| \, \|\phi'(x) - \phi'(c)\| \\ &\leq k \, \tfrac{\epsilon}{k} \\ &= \epsilon \end{aligned}$$

therefore $(1-\epsilon)B_1 \subset T(c)\phi B \subset (1+\epsilon)B_1$, where B_1 is a translate of B to new centre $T(c)\phi(x)$. Therefore

$$(1-\epsilon)^n m(B) \leq |\det T(c)| m(\phi B) \leq (1+\epsilon)^n m(B)$$

$$\implies (1-\epsilon)^n \leq |\det T(c)| \frac{m(\phi B)}{m(B)} \leq (1+\epsilon)^n$$

$$\implies \lim |\det T(c)| \frac{m(\phi B)}{m(B)} = 1$$

$$\implies |\det T(a)| \lim \frac{m(\phi B)}{m(B)} = 1$$

Therefore:

$$\lim \frac{m(\phi B)}{m(B)} = \frac{1}{|\det T(a)|} = |\det \phi'(a)|$$

as required □

Theorem 9.1.2. *Let f be a continuous real valued function on an open neighbourhood of a in $\mathbb{R}^n$. THen*

$$\lim \frac{\int_B f(x)\,dx}{m(B)} = f(a)$$

where the limit is taken over cubes B containing a with radius $B \longrightarrow 0$.

Proof. Let $\epsilon > 0$. Choose $\delta > 0$ s.t. $|f(x) - f(a)| < \epsilon \,\forall\, \|x - a\| < \delta$. Then each cube B containing a of radius $\leq \frac{\delta}{2}$ we have:

$$\left| \frac{\int_B f(x)\,dx}{m(B)} - f(a) \right| = \frac{\left|\int_B [f(x) - f(a)]\,dx\right|}{m(B)} \leq \frac{\epsilon m(B)}{m(B)} = \epsilon$$

hence result. □

Recall that the σ-algebra generated by the topology of $\mathbb{R}^n$ is called the collection of Borel Sets in $\mathbb{R}^n$.

Theorem 9.1.3. *Let A be a Borel set in $\mathbb{R}^r$, B be a Borel set in $\mathbb{R}^s$. Then $A \times B$ is a Borel set in $\mathbb{R}^{r+s}$.*

Proof. For fixed V open in $\mathbb{R}^r$, the sets:

$$\{V \times W : W \text{ open in } \mathbb{R}^s\} \tag{9.1}$$

are all open in $\mathbb{R}^{r+s}$. Therefore the σ-algebra generated by Eqn(9.1):

$$\{V \times B : B \text{ Borel in } \mathbb{R}^s\}$$

consists of Borel sets in $\mathbb{R}^{r+s}$. Hence for fixed B Borel in $\mathbb{R}^s$, the set:

$$\{V \times B : V \text{ open in } \mathbb{R}^r \tag{9.2}$$

are all Borel sets in $\mathbb{R}^{r+s}$. Therefore the σ-algebra generated by Eqn(9.2):

$$\{A \times B : A \text{ Borel in } \mathbb{R}^r\}$$

consists of Borel sets in $\mathbb{R}^{r+s}$. Therefore $A \times B$ is Borel for each A Borel in $\mathbb{R}^r$, B Borel in $\mathbb{R}^s$. □

Theorem 9.1.4. *Let E be a lebesgue measurable subset of $\mathbb{R}^n$. THen there is a Borel set B containing E such that*

$$B - E = B \cap E'$$

has measure zero, and hence $m(B) = m(E)$.

Proof. We already know this is true for $n = 1$. So we use induction on n. Assume true for $n - 1$. Let $E \subset \mathbb{R}^n$, E measurable.

1. Let $m(E) < \infty$. Let k be an integer > 0.

$$E \subset \mathbb{R}^{n-1} \times \mathbb{R}$$

choose a sequence of rectangles

$$A_1 \times B - 1, A_2 \times B_2, \ldots$$

covering E with $A_i \subset \mathbb{R}^{n-1}$ measurable and $B_i \subset \mathbb{R}$ measurable, and such that

$$m(E) \leq \sum_{i=1}^{\infty} m(A_i)m(B_i) \leq m(E) + \frac{1}{k}$$

By the induction hypothesis choose Borel sets C_i, D_i s.t.

$$A_i \subset C_i,\ B_i \subset D_i,\ m(A_i) = m(C_i),\ m(B_i) = m(D_i)$$

Put $B_k = \bigcup_{i=1}^{\infty} C_i \times D_i$. Then $E \subset B_k$, B_k is Borel and $m(E) \leq m(B_k) \leq \sum_{i=1}^{\infty} m(C_i)m(d_i) \leq m(E) + \frac{1}{k}$.

Put $B = \bigcap_{k=1}^{\infty} B_k$. Then $E \subset B$, B is Borel and $m(E) = m(B)$. Therefore $m(B \cap E') = m(B) - m(E) = 0$ as required.

2. Let $m(E) = \infty$. Put $E_k = \{x \in E : k \le |x| < k+1\}$. $E = \bigcup_{k=0}^{\infty} E_k$ is a countable disjoint union, and $m(E_k) < \infty$. For each integer k choose by 1. Borel B_k s.t. $E_k \subset B_k$ and $m(B_k \cap E'_k) = 0$

   Put $B = \bigcup B_k$. Then $E \subset B$ and $m(B \cap E') = m(B) - m(E)$ as required. ???????????

□

9.2 General change of variable in a multiple integral

Theorem 9.2.1 (General change of variable in a multiple integral). *Let $\mathbb{R}^n \supset V \xrightarrow{\phi} W \subset \mathbb{R}^n$ be a C^1 diffeomorphism of open V onto open W. Let f be integrable on E. Then*

$$\int_W f(x)\,dx = \int_V f(\phi(x))|\det \phi'(x)|\,dx \qquad \textit{(Lebesgue Integrals)} \tag{9.3}$$

Proof. 1. we have $f = f^+ - f^-$ with $f^+, f^- \ge 0$. Therefore, it is sufficient to prove Eqn(9.3) for $f \ge 0$.

2. if $f \ge 0$ then $\exists$ a monotone increasing sequence of non-negative simple functions $\{f_n\}$ such that $f = \lim f_n$ so, using the monotone convergence theorem, it is sufficient to prove Eqn(9.3) for f simple.

3. if f is simple then $f = \sum_{i=1}^{k} a_i \chi_{E_i}$ with $\{E_i\}$ Lebesgue measurable, so it is sufficient to prove Eqn(9.3) with $f = \chi_E$ with E Lebesgue measurable.

4. if E Lebesgue measurable $\exists$ Borel B s.t. $E \subset B$ and $Z = B - E$ has measure zero. Therefore, $\chi_E = \chi_B - \chi_Z$, and it is sufficient to prove Eqn(9.3) for $f = \chi_B$, B Borel, and for $f = \chi_Z$, Z measure zero.

5. If $f = \chi_E$ with E Borel then $E = \phi F$ with F Borel and Eqn(9.3) reduces to

$$\int_W \chi_{\phi F}(x)\,dx = \int_V \chi_F(x)|\det \phi'(x)|\,dx$$

   i.e.

$$m(\phi F) = \int_F |\det \phi'(x)|\,dx \tag{9.4}$$

   If Eqn(9.4) holds for each rectangle $F \subset V$ then Eqn(9.4) holds for $F \in$ ring R of finite disjoint unions of rectangles $\subset V$.

Therefore Eqn(9.4) holds for $F \in$ monotone class generated by R.

Therefore Eqn(9.4) holds for $F \in \sigma$-algebra generated by R (monotone class lemma). So Eqn(9.4) holds for any Borel set $F \subset V$

6. to show Eqn(9.4) holds for any rectangle $F \subset V$: for each rectangle $F \subset V$ put

$$\lambda(F) = m(\phi F) - \int_F |\det \phi'(x)|\, dx$$

Then, λ is additive:

$$\lambda(\bigcup B_j) = \sum \lambda(B_j)$$

for a disjoint union.

Must prove $\lambda(F) = 0$ for each rectangle F. Suppose B is a rectangle for which $\lambda(B) \neq 0$.

Suppose B is a cube. $|\lambda(B)| > 0$. Therefore, $\exists\, \epsilon > 0$ s.t. $|\lambda(B)| \geq \epsilon\, m(B)$.

Divide B into disjoint subcubes, each of $\frac{1}{2}$ the edge of B. For one such, B_1 say,

$$|almbda(B_1)| \geq \epsilon\, m(B_1)$$

Sivide again

$$|\lambda(B_2)| \geq \epsilon\, m(B_2)$$

continuing we get a decreasing sequence of cubes $\{B_k\}$ converging to a say, with

$$\lim_{h \longrightarrow \infty} \frac{|\lambda(B_k)|}{m(B_k)} \geq \epsilon > 0$$

But

$$\begin{aligned} \lim_{h\to\infty} \tfrac{\lambda(B_k)}{m(B_k)} &= \lim_{h\to\infty} \tfrac{m(\phi B_k) - \int_{B_k} |\det \phi'(x)|\, dx}{m(B_k)} \\ &= |\det \phi'(a)| - |\det \phi'(a)| \\ &= 0 \end{aligned}$$

a contradiction. Therefore $\lambda(B) = 0$ for all cubes B. Therefore, $m(\phi B) = \int_B |\det \phi'(x)|\, dx$ for all cubes B as required.

7. Now suppose Z has measure zero, and choose a Borel set B s.t.

$$Z \subset B \subset V$$

and s.t. B has measure zero.

Then

$$m(\phi Z) \le m(\phi B) = \int_B |\det \phi'(x)|\, dx = 0$$

since B has measure zero. Therefore ϕZ has measure zero.

Therefore under a C^1 diffeomorphism we have:

$$Z \text{ of measure zero } \implies \phi Z \text{ of measure zero}$$

therefore $f = \chi_Z$ with Z of measure zero $\implies$ $f \cdots \phi = \chi_{\phi^{-1}Z}$ with $\phi^{-1}Z$ of measure zero.

Therefore, Eqn(9.3) holds for $f = \chi_Z$ since then both sides are zero. This completes the proof. □