LANGUAGE Of MATHEMATICS

Ruiqing Li

each proof is of sketch

Contents

Notation Compilation

$I > 1 > 1.1 > 1.1.1 > (A) > (1) > (i) > (a) >$ ①

Definitions that contains the undefined are illegal

so adopt the meta meaning until we literally define it.

Abbreviation for Meta Language	
well defined	
let/assume	
suppose not	
since/for/by	
so/thus/hence	
written/denoted by	
follow/similarly	
trivial/hold	
taken arbitrarily	
closed under/preserve	
specially/moreover	
with/where	
o.w.	otherwise
s.t.	such that
viz.	namely
i.e.	that is
wlog	without loss of generality
iff	if and only if
e.g.	for example
etc	et cetera
qed	proof end
resp.	respectively
wrt.	with respect to
:=	let be
$\equiv$	be
$\neg$	not
$\wedge$	and
$\vee$	or
$\leftrightarrow$ / $\Leftrightarrow$	iff
$\rightarrow$ / $\Rightarrow$	if then
$\forall$	for any
$\exists$	there be such that
$\square$	qed

Abbreviation for Math Language	
$\alpha \wedge \beta$	$\neg(\alpha \rightarrow \neg\beta)$
$\alpha \vee \beta$	$\neg\alpha \rightarrow \beta$
$\alpha \leftrightarrow \beta$	$(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$
$(\exists v)$	$\neg(\forall v)\neg$
$\bigvee_{i=1}^{n} \varphi_i$	$\varphi_1 \vee ... \vee \varphi_n$
$\bigvee_{i=1}^{0} \varphi_i$	$\perp$
$\bar{v}$	$(v_1, ..., v_n)/v_1...v_n$
$\bar{v}P$	$\bigwedge_{i=1}^{n}(v_i P)$
$v_1 P v_2$	$P v_1 v_2$
$v_1 P ... P v_n$	$\bigwedge_{i=1}^{n-1} v_i P v_{i+1}$
$v_1 = v_2$	$v_1 \approx v_2$ (without ambiguity)
xy	$x \cdot y$ (without ambiguity)
$v_1 < v_2$	$v_1 \leq v_2 \wedge \neg v_1 = v_2$ (2-ary relation)
$v \neq w$	$\neg v = w$ (2-ary relation)
$v_1 \geq v_2$	$v_2 \leq v_1$ (2-ary antonymy)
$(\forall P(v))\varphi$	$(\forall v)(P(v) \rightarrow \varphi)$
$(\exists P(v))\varphi$	$(\exists v)(P(v) \wedge \varphi)$
$(\exists^{\geq n} v)\varphi(v)$	$(\exists v_1, ..., v_n) \bigwedge_{i=1}^{n} \varphi(v_i) \wedge \bigwedge_{1 \leq i < j \leq n} v_i \neq v_j$
$(\exists^{=n} v)\varphi(v)$	$(\exists^{\geq n} v)\varphi(v) \wedge \neg(\exists^{\geq n+1} v)\varphi(v)$
$(\exists! v)\varphi(v)$	$(\exists^{=1} v)\varphi(v)$
$\varphi(\bar{v}; \bar{w})$	φ in bounded $\bar{v}$ and free $\bar{w}$
$\varphi(\bar{v})$	φ in free $\bar{v}$
$x \vee y$	$\bigvee\{x, y\}/\sup\{x, y\}$
$X = \{x : P, Q\}$	$(\forall x)(x \in X \leftrightarrow P \wedge Q)$

1 Logic and Mathematical Language

1.1 Language

Definition 1.1.1 (Vocab)

The *vocabulary* is a set $\mathcal{L}$ that consists of ①*parenthese* $(,)$ ②*connectives* $\neg, \to$ ③*quantifier* $\forall$ ④*identity* $\approx$ ⑤*variables* $v_0, ...$ ⑥*predicates* $P^{n_P}, ...$ where $n_P \in \mathbb{N}$ ⑦*functions* $f^{n_f}, ...$ where $n_f \in \mathbb{N}$, written $\mathcal{L} = \{P, ..., f, ...\}$ in brief. Specially, regard P^0 as *truth-value sentence* $\top, \bot$ and f^0 as *constant* $c_0, ...$.

Definition 1.1.2 (Syntax)

The *syntax* of *first-order logic* consists of

(1)$\mathcal{L}$*-expression* is $X_1...X_n$ where $X \in \mathcal{L}$;

(2)$\mathcal{L}$*-term* is in $S_{term} := \bigcap\{X : (\forall v \in \mathcal{L})v \in X, (\forall f \in \mathcal{L}, t \in X)f\bar{t} \in X\}$;

(3)*atomic* $\mathcal{L}$*-formula* is $t_1 \approx t_2$ and $P\bar{t}$ where $P \in \mathcal{L}, t \in S_{term}$;

(4)$\mathcal{L}$*-formula* is in $S_{formula} := \bigcap\{Y : S_{atomic\ formula} \subseteq Y, (\forall v \in \mathcal{L}, \varphi, \psi \in Y)(\neg\varphi, \varphi \to \psi, \forall v\varphi \in Y)\}$;

(5)variable v is *free* in formula φ o.w. *bound*, iff ①v occurs in φ when $\varphi \in S_{atomic\ formula}$ ②v is free in ψ when $\varphi \equiv \neg\psi$ ③v is free in ψ or ϕ when $\varphi \equiv \psi \to \phi$ ④v is free in ψ and $v \neq w$ when $\varphi \equiv (\forall w)\psi$;

(6)$\mathcal{L}$*-sentence* is a formula φ s.t. $(\forall v \in \mathcal{L})v$ is bound in φ.

Definition 1.1.3 (Semantics)

(1)The $\mathcal{L}$*-structure* is a set $\mathcal{M}$ that consists of ①*underlying set* M ②*interpretation function* ${}^{\mathcal{M}}$: $P \mapsto P^{\mathcal{M}} \subseteq M^{n_P}, f \mapsto f^{\mathcal{M}} : M^{n_f} \to M$, written $\mathcal{M} = \{M, P^{\mathcal{M}}, ..., f^{\mathcal{M}}, ...\}$ in brief. Specially, regard M^0 as $\{\Xi\}$ then $\top^{\mathcal{M}}$ as *truth-value* $T := \{\Xi\}$ and $\bot^{\mathcal{M}}$ as $F := \emptyset$.

(2)Interpretation of term $t(\bar{v})$ is a function $t^{\mathcal{M}}$ s.t. for subterm s and $\bar{a} \in M^n$, ①$s^{\mathcal{M}}(\bar{a}) = a_i$ when $s \equiv v_i$ ②$s^{\mathcal{M}}(\bar{a}) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a}), ..., t_n^{\mathcal{M}}(\bar{a}))$ when $s \equiv f\bar{t}$.

(3)For formula $\varphi(\bar{v})$ and $\bar{a} \in M^n$, $\mathcal{M}$ *satisfies* $\varphi(\bar{a})$, written $\mathcal{M} \models \varphi(\bar{a})$, iff

(a)$t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a})$ when $\varphi \equiv t_1 = t_2$

(b)$(t_1^{\mathcal{M}}(\bar{a}), ..., t_n^{\mathcal{M}}(\bar{a})) \in P^{\mathcal{M}}$ when $\varphi \equiv P\bar{t}$

(c)$\mathcal{M} \not\models \psi$ when $\varphi \equiv \neg\psi$

(d)$\mathcal{M} \not\models \psi \vee \mathcal{M} \models \phi$ when $\varphi \equiv \psi \to \phi$

(e)$(\forall b \in M)\mathcal{M} \models \psi(\bar{a}, b)$ when $\varphi \equiv (\forall w)\psi(\bar{v}, w)$.

Definition 1.1.4 (Higher-order and Many-sorted Language)

(1)*First-order language (FOL for short)* has vocab, syntax and semantics above, denoted also by $\mathcal{L}$, while *second-order language (SOL for short)* adds:

(a)for vocab, divide variable into *individual variable* $v_0, ...$ and *predicate variable* $X_1^{n_1}, ...$, replace the rules about variable by individual variable.

(b)for syntax, add recursive rules "$(\forall X \in \mathcal{L})(t \in S_{term})X\bar{t} \in S_{atomic\ formula}$" and "$(\forall X \in \mathcal{L}, \varphi \in$

$S_{formula}$)$\forall X\varphi \in S_{formula}$", add *free predicate variable.*
(c)for semantics, add interpretation of atomic $\mathcal{L}$-formula $X\bar{t}$, add "$(\forall S \in M^{n_X})\mathcal{M} \models \psi(\bar{a}, \bar{R}, S)$ when $\varphi \equiv (\forall Y)\psi(\bar{v}, \bar{X}, Y)$" into $\mathcal{M} \models \varphi(\bar{a}, \bar{R})$.

Informally speaking, 1st-order quantifies only variable that range over individuals, 2th-order also quantifies over sets, 3rd-order also quantifies over sets of sets, etc.

(2)Informally speaking, *many-sorted Language (MSL for short)* divides variables into several parts called *sorts*, then extends vocab, syntax and semantics slightly. Note *many-sorted structure* $\mathcal{M}$ with sorts S contains underlying sets $M_s(s \in S)$ and $^{\mathcal{M}} : P \mapsto P^{\mathcal{M}} \subseteq M_{s_1} \times ... \times M_{s_{n_p}}, f \mapsto f^{\mathcal{M}} : M_{s_1} \times ... \times M_{s_{n_f}} \to M_{s_0}$.

Example 1.1.1

(1)Based on $\mathcal{L} = \{\in\}$ of set theory, the axioms of topology are at least 3rd-order, e.g. "a topology is closed under unions" as "$\forall \mathcal{U}((\forall U \in \mathcal{U} \to U \in \mathcal{T}) \to (\exists V \forall x(x \in V \leftrightarrow \exists U \in \mathcal{U}(x \in U)) \wedge V \in \mathcal{T}))$".

(2)To get lower-order or less-sorted, you might use more expensive vocab, e.g. "axioms of module" is "2-sorted in $\mathcal{L} = \{+_R, \cdot_R, +_M, \cdot\}$ with sorts $\{R, M\}$" but "FOL in $\mathcal{L} = \{+\} \cup R$ where $(R, +_R, \cdot_R)$ is a unitary ring and $r \in R$ is a 1-ary function".

(3)Moreover, we always sacrifice the cheap vocab for a good property in FOL, e.g. "axioms of group" is "①$(\forall x, y, z)(xy)z = x(yz)$ ②$(\exists e)(\forall x)(ex = xe = x \wedge (\exists y)yx = xy = e)$" in $\mathcal{L}' := \{\cdot\}$, but "①$(\forall x, y, z)(xy)z = x(yz)$ ②$(\forall x)ex = xe = x$ ③$(\forall x)xx^{-1} = x^{-1}x = e$" in $\mathcal{L} = \{\cdot, ^{-1}, e\}$. Note every axiom of the latter is universal sentence, there're many advantages in model theory.

Remark 1.1.1

(1)To balance the unique readability against simplicity, for parentheses ①omit outermost ②$\neg$ control nearest ③right-associativity.

(2)Although $\mathcal{L} = \{P, ..., f, ...\}$, actually we omit variable, identity, quantifier etc., e.g. $|\mathcal{L} = \{\in\}| = \aleph_0$.

(3)Usually denote $\{\varphi_i(\bar{v})\}_{i \in I}$ by $\Sigma(\bar{v})$, $\mathcal{M}$ *satisfy* Σ iff $(\forall \varphi \in \Sigma)\mathcal{M} \models \varphi$.

(4)For set X, $\mathcal{L}(X) := \mathcal{L} \cup X$ where $x \in X$ is 0-ary function. For $X \subseteq M$, *expansion* of $\mathcal{M}$ is $\mathcal{M}_X$ that adds interpretion $x^{\mathcal{M}} \mapsto x$ of $\mathcal{L}(X) \setminus \mathcal{L}$; for sublanguage $\mathcal{L}^-$, *reduct* of cM is $\mathcal{M} \upharpoonright \mathcal{L}^-$ that omits interpretation of $\mathcal{L} \setminus \mathcal{L}^-$.

(5)

Remark 1.1.2 (sub, generator and base)

(1)*construction sequence* of t is $(t_1, ..., t_n)$ where

(i)$t_n = t$

(ii)$\forall 1 \leq i \leq n(t_i = v \vee (\exists f \in \mathcal{L}, i_1, ..., i_{n_f} < i)t_i = f(t_{i_1}, ..., t_{i_{n_f}}))$.

(2)So s is *subterm (resp. subformula)* of t iff s is a term (resp. formula) and s is a subsequence of t.

(3)Generally speaking

(i)B is a subtype of A iff in an abstract sense $B \subseteq A$ and B preserves specific type of A.

(ii)There're three methods to describe "type generated by A", i.e. the smallest type containing A, the intersection of types containing A, constructive discription, e.g. substructure *generated* by $A \subseteq M$ is $< A >=< A >_M := \{t^{\mathcal{M}}(\bar{a}) : t(x_1, ..., x_n) \in S_{term}, \bar{a} \in A^n\}$. For type $\mathcal{S}$, usually denote the type generated by A by $\mathcal{S}(A)$. Note the 1st and 2nd are from top to bottom while the 3rd from bottom to top.

(iii)Usually B is generators of A of type $\mathcal{S}$ iff $A = \mathcal{S}(B)$, and B is basis of A of type $\mathcal{S}$ iff B is generators as independent as possible, i.e. $(\forall b \in B)\mathcal{S}(B \setminus \{b\}) \neq A$.

(iv)Usually a free type is a type that has a basis.

◊WARNING: we won't define sub, generator, base of types below unless something interesting happens, e.g. base of topological space has some kind of difference.

1.2 Theory

Definition 1.2.1 (theory)

$\mathcal{L}$-*theory* $T := \{\varphi \in S_{sentence}\}$ whose element is *axiom*.

(1)$\mathcal{M}$ is *model* of T, written $\mathcal{M} \models T$, denote the class of models of T by $Mod(T)$, iff $\mathcal{M}$ *satisfy* T.

$Th(\mathcal{M}) := \{\varphi : \varphi \in S_{sentence} \wedge M \models \varphi\}$ is *complete theory* of $\mathcal{M}$.

(2)φ is

(i)*logical consequence* of T, written $T \models \varphi$, iff $(\forall \mathcal{M} \models T)M \models \varphi$;

(ii)*provable* from T, written $T \vdash \varphi$, iff $\exists$ proof of φ from T.

(4)T is

(i)*satisfiable* iff $(\exists \mathcal{M})\mathcal{M} \models T$;

(ii)*inconsistent* o.w. *consistent*, iff $(\exists \varphi \in S_{sentence})T \vdash \varphi \wedge T \vdash \neg\varphi$;

(iii)*complete* iff $(\forall \varphi)T \vdash \varphi \vee T \nvdash \varphi$;

(iv)*decidable* iff $\exists$ algorithm that when given $\varphi \in S_{sentence}$ as input decides whether $T \models \varphi$.

Remark 1.2.1

(1)$\mathcal{M}, \mathcal{N}$ are *elementarily equivalent*, written $M \equiv N$, iff $Th(\mathcal{M}) = Th(\mathcal{N})$.

(2)*Elementary class* is a class $\mathcal{K}$ of $\mathcal{L}$-structures s.t. $(\exists T)\mathcal{K} = Mod(T)$.

Definition 1.2.2 (recursion)

(1)*basic primitive (function)*

(i)$Z : \mathbb{N} \to \mathbb{N}, x \mapsto 0$

(ii)$S : \mathbb{N} \to \mathbb{N}, x \mapsto x + 1$

(iii)$P_i^n : \mathbb{N}^n \to \mathbb{N}, \bar{x} \mapsto x_i$

denote the set of basic primitive by S_{bp} (2)*primitive recursive (or computable)* element of $\bigcap_{S_{bp}\subseteq X, closed\ under\ composition\ and\ primitive\ recursion} X$ where

(i)(composition)$(\forall h(x_1,...,x_m), g_1(\bar{x}),...,g_m(\bar{x}))f(\bar{x}) = h(g_1(\bar{x}),...,g_m(\bar{x}))$

(ii)*primitive recursion (or computability)* means $(\forall g(\bar{x}), h(\bar{x}, x_{n+1}, x_{n+2}))f(\bar{x},0) = g(\bar{x}) \wedge f(\bar{x}, S(y)) = h(\bar{x}, y, f(\bar{x}, y))$.

(3)μ-*/general recursive*

element of $\bigcap_{S_{bp}\subseteq X, closed\ under\ composition,\ primitive\ recursion\ and\ minimization\ operator} X$ where

minimization operator means $(\forall f: \mathbb{N}^{n+1} \to \mathbb{N})g(\bar{x}) = \mu y(\forall z \leq y(f(\bar{x}, z)\downarrow) \wedge f(\bar{x}, y) = 0)$.

For $P(\bar{x}, z) \subseteq \mathbb{N}^{n+1}$, *minimization quantifier* μ s.t. $(\mu z \leq y)P(\bar{x}, z) = \begin{cases} min\{z : z \leq y \wedge P(\bar{x}, z)\} & if\ well\ defined \\ y+1 & o.w. \end{cases}$.

(4)*(total) recursive* o.w. *partial recursive*

iff f is μ-general recursive and total.

(5)*recursively enumerable (r.e. for short)* of $X \subseteq \mathbb{N}$

iff $X = \emptyset$ or $(\exists$recursive $f)X = \{y : (\exists x)f(x) = y\}$.

Remark 1.2.2

(1)By characteristic function, if we define some property for function or set, then there's natural generalization for the other.

$\Diamond$WARNING: we won't define such property twice below unless something interesting happens.

(2)$\mathcal{L}$ recursive iff ($\exists$ algorithm e)$(\forall\{X_i\}_1^n)$ where X_i is a symbol in $\mathcal{L}$, decides whether $X_1...X_n$ is $\mathcal{L}$-formula.

Remark 1.2.3

(1)(informal)*code*; *Gödel number*; *decode*; *Turing machine*; *procedure*; *halt*; *Turing computable*; *algorithm*; *effective*; *decidable*

(2)(informal)*proof* is a sequence of $\mathcal{L}$-formulas generated by *proof system* which satisfies

(i)length of sequence is finite;

(ii)if $T \vdash \varphi$, then $T \models \varphi$;

(iii)if $|T| < \infty$, then ($\exists$ algorithm e)$(\forall \varphi, \{\varphi_i\}_1^n)$when given $\{\varphi_i\}$ as input, e decides whether $\{\varphi_i\}$ is a proof of φ from T.

Definition 1.2.3 (definition)

$X \subseteq M^n$ is A-*definable* where $A \subseteq M$, denote $\emptyset$-definable by *definable*, iff

$(\exists \varphi(\bar{v}, \bar{w}))(\exists \bar{a} \in A^m)X = \{\bar{x} \in M^n : \mathcal{M} \models \varphi(\bar{x}, \bar{a})\}$.

In this case, $\varphi(\bar{v}, \bar{a})$ *define* X.

Remark 1.2.4

(1)$\mathcal{L}_0$-structure $\mathcal{N}$ is *definably interpreted* in $\mathcal{L}$-structure $\mathcal{M}$ iff

$$(\exists n)(\exists X \subseteq M^n)(\forall P, f \in \mathcal{L}_0)(\exists \mathcal{L} - definable\ P^{\mathcal{X}} \subseteq X^{n_P}, f^{\mathcal{X}} \subseteq X^{n_f+1})(X, P^{\mathcal{X}}, ..., f^{\mathcal{X}}, ...) \cong \mathcal{N}.$$

(2)E.g. interpret finite poset in *Hasse diagram* and interpret category in *commutative diagram*

1.3 Model Theory

2 Set and Category Theory

2.1 various theory

Definition 2.1.1 (equivalence)

$\mathcal{L} = \{\approx\}$, T of *equivalence* consists of

(1)(reflexive) $(\forall x)x \approx x$

(2)(symmetric) $(\forall x, y)x \approx y \to y \approx x$

(3)(transitive) $(\forall x, y, z)x \approx y \wedge y \approx z \to x \approx z$.

Remark 2.1.1

We can compare two equivalence which is *finer* or *coarser*.

◊WARNING: we won't define such property below unless something interesting happens.

Definition 2.1.2 (order)

$\mathcal{L} = \{\leq\}$, T of *partial order* consists of

(1)$(\forall x)x \leq x$

(2)(antisymmetric) $(\forall x, y)(x \leq y \wedge y \leq x \to x = y)$

(3)$(\forall x, y, z)(x \leq y \wedge y \leq z \to x \leq z)$

the underlying set of its model is *partial order set (poset for short)*.

(i)*preorder* (1)(3)

(ii)*linear order (or chain)* (2)-(4)$(\forall x, y)x \leq y \vee y \leq x$

(iii)*well order* (2)-(5)$(\exists x)(\forall y)x \leq y$.

Remark 2.1.2

(1)*Upper closure (resp. lower closure)* of $x \in X$ is $x^{\uparrow X} = \uparrow x := \{y \in X : x \leq y\}$ (resp. $x^{\downarrow X}$ or $\downarrow x$). Upper closure of $A \subseteq X$ is $A^{\uparrow X} = \uparrow A := \bigcup_{a \in A} a^{\uparrow X}$.

(2)*interval* $< x, y >$ where $<$ (resp. $>$) is (or [(resp.) or])

(i)*open interval* $(x, y) := \{z \in X : x < z < y\}$

(ii)*closed interval* $[x, y] := \uparrow x \cap \downarrow y$

(iii)*half-open interval* contains *left-open interval* and *right-open interval*.

$x \in X$ *covers* $y \in X$, written $y \lessdot x$ iff $y < x$ and $\forall z \in X(y \leq z \leq x \to (z = y \vee z = x))$.

(2)$a \in A \subseteq X$ is

maximal (resp. minimal) of A, written $\max A$ (resp. $\min A$) iff $\uparrow a \cap A = \{a\}$

$x \in X \supseteq A$ is

(i)*upper bound (resp. lower bound)* of A, denote the set of upper bounds of A by $ub(A)$ (resp. $lb(A)$), iff $(\forall a \in A)a \leq x$;

(ii)*supremum (or least upper bound) (resp. infimum)* or *join (resp. meet)* of A, written $\sup A$ (resp. $\inf A$) or $\bigvee A$ (resp. $\bigwedge A$), iff $x = \min ub(A)$.

◊WARNING: for $\emptyset$, it not has maximal, might have supremum, always has upper bound. This must be taken into consideration when giving the well defined. (3)Poset X is

(i)*pointed* iff $\min X$ exist, called *bottom* written $\perp$.

(ii)*pointed directed-completed (pointed dcpo for short)* iff any directed subset or emptyset has a supremum in X, denote supremum x of directed subset A by $x = \vee^{\uparrow} A$.

(a)*directed-complete (dcpo for short)* iff any directed subset has a supremum in X;

(b)*finite complete* iff $(\forall x, y \in X) x \vee y, x \wedge y \in X$;

①$\vee$*-semilattice* iff $(\forall x, y \in X) x \vee y$

②$\wedge$*-semilattice* iff $(\forall x, y \in X) x \wedge y$

(c)*Dedekind complete (conditional complete or least-upper-bound property)* iff any non-empty subset with an upper bound has a supremum in X;

(d)*complete* iff any subset has a supremum in X.

(iii)*dense* iff $(\forall x < y)(\exists z) x < z < y$.

(4)$A \subseteq X$ is

(i)*upper set (resp. lower set (or initial segment))* iff $A^{\uparrow X} = A$;

(ii)*(upward) directed (resp. filtered (or downward directed))* iff $A \neq \emptyset$ and $(\forall a, b \in A)(\exists c \in A) a, b \leq c$;

(iii)*ideal (resp. filter)* iff A is a lower set and directed;

(iv)*principal ideal (resp. principal filter)* iff A is ideal and $\max A$ exists;

(v)*cofinal (or frequent) (resp. coinitial)* in X iff $(\forall x \in X)(\exists a \in A) x \leq a$.

◊WARNING: note $\subseteq$ is a partial order relation, so we won't define the same item twice unless something interesting happens. For instance

(i)ideal in order theory coincides with ideal in set theory, where

(a)$(\forall a, b \in A) a \vee b \in A$

(b)$(\forall a \in A)(\forall b = a \wedge b) b \in A$;

(ii)A finite complete poset naturally becomes a lattice, while a pointed dcpo naturally becomes a *complete lattice*.

Theorem 2.1.1

(1)Any subset has a supremum iff any subset has an infimum; any non-empty subset with an upper bound has a supremum iff any non-empty subset with an upper bound has an infimum.

(2)In domain theory, pointed directed-completed iff chain-complete (every chain has a supremum in X).

Definition 2.1.3 (set)

$\mathcal{L} = \{\in\}$, T of *Zermelo-Fraenkel with Choice (ZFC for short)* consists of

(1)(extensionality) $(\forall x)(\forall y)((\forall z)(z \in x \leftrightarrow z \in y) \rightarrow x = y)$

(2)(axiom schema of separation) $(\forall \varphi(x, z, w_1, ..., w_n))(\forall w_1)...(\forall w_n)(\forall z)(\exists y)(\forall x)(x \in y \leftrightarrow (x \in z \wedge$

$\varphi))$

(3)(pairing) $(\forall x)(\forall y)(\exists z)(x \in z \wedge y \in z)$

(4)(union) $(\forall x)(\exists y)(\forall w)((\exists z)(w \in z \wedge z \in x) \to w \in y)$

(5)(power set) $(\forall x)(\exists y)(\forall z)(z \subseteq x \to z \in y)$

(6)(axiom schema of replacement) $(\forall \varphi(x, y, w_1, ..., w_n))(\forall w_1)...(\forall w_n)(\forall z)((\forall x)(x \in z \to (\exists! y)\varphi) \to (\exists u)(\forall x)(x \in z \to (\exists y)(y \in u \wedge \varphi)))$

(7)(infinity) $(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \to \mathcal{S}(y) \in x))$

(8)(regularity) $(\forall x)(x \neq \emptyset \to (\exists y)(y \in x \wedge x \cap y = \emptyset))$

(9)(choice) $(\forall x)((\forall y_1)(\forall y_2)((y_1 \in x \wedge y_2 \in x) \to (y_1 \neq \emptyset \wedge (y_1 = y_2 \vee y_1 \cap y_2 = \emptyset))) \to (\exists z)(\forall y)(y \in x \to (\exists! w) w \in y \cap z))$.

Remark 2.1.3

For axioms in relatively brief forms above, add details after (2)(5)(9) separately:

(A)By (2), *empty set* $\emptyset := (\forall x)(x \notin \emptyset)$ is well defined. *subset* $z \subseteq x := (\forall w)(w \in z \to w \in x)$.

(B)By (2), *intersection of two* $x \cap y := \{z : z \in x \wedge z \in y\}$ is well defined. By (2)(3), $\{x, y\}$ is well defined. By (2)(4), $\mathcal{U}(x) := \{z : (\exists y)(z \in y \wedge y \in x)\}$ is well defined. *union of two* $x \cup y := \mathcal{U}(x, y)$, *successor function* $\mathcal{S}(x) := x \cup \{x\}$. By (2)(5), *power set* $\mathcal{P}(x) = 2^x := \{z : z \subseteq x\}$ is well defined.

(C)By (2), (6) can be strengthened to $(\forall \varphi(x, y, w_1, ..., w_n))(\forall w_1)...(\forall w_n)(\forall z)((\forall x)(x \in z \to (\exists! y)\varphi) \to (\exists u)(\forall y)(y \in u \leftrightarrow (\exists x)(x \in z \wedge \varphi)))$.

ordered pair $(x, y) := \{\{x\}, \{x, y\}\}$, *index set*, $(x_i)_{i \in I}$. Note it has some subtle difference from "sequence".

Remark 2.1.4

We will adopt *von Neumann-Bernays-Gödel (NBG for short)* set theory along the journey, which is a conservative extension of NFC.

$\mathcal{L} = \{\in, IsSet\}$ where $IsSet$ is a 1-ary predicate, (10)$(\forall x)(IsSet(x) \leftrightarrow (\exists y) x \in y)$ and some subtle differences. And *proper class* is a class that is not a set.

Definition 2.1.4 (set operation and property)

(1)*difference* $x \setminus y := \{z : z \in x, z \notin y\}$

(i)*complement* $x^c := U \setminus x$

(ii)*symmetric difference* $x \triangle y := (x \setminus y) \cup (y \setminus x)$

(2)*union* $\bigcup_{i \in I} X_i := \{x : (\exists i \in I) x \in X_i\}$

disjoint union $\bigsqcup_{i \in I} X_i := \bigcup_{i \in I} \{(x, i) : x \in X_i\}$

(3)*intersection* $\bigcap_{i \in I} X_i := \{x : (\forall i \in I) x \in X_i\}$

(4)*product* $\prod_{i \in I} X_i := \{(x_i)_{i \in I} : x_i \in X_i\}$ or $\{f : I \to \bigcup_{i \in I} X_i | (\forall i \in I) f(i) \in X_i\}$ then denote f by (x_i)

power $X^Y := \{f : Y \to X\}$

(5)*quotient* $X/\sim := \{[x] : x \in X\}$

equivalence class of $x \in X$ about $\sim$ is $[x] := \{y \in X : y \sim x\}$.

(6)$\{X_i\}_{i\in I}$ is

(i)*(pairwise) disjoint* iff $(\forall i, j \in I)X_i \cap X_j = \emptyset$;

(ii)when $I \subseteq \mathbb{N}$, *monotonic* iff (*increasing*) $X_1 \subseteq X_2 \subseteq ...$ or (*decreasing*) $X_1 \supseteq X_2 \supseteq ....$

(a)*limit superior* $\overline{\lim}_{n\to\infty} X_n = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} X_i$;

(b)*limit inferior* $\underline{\lim}_{n\to\infty} X_n = \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} X_i$;

(c)*limit* $lim_{n\to\infty} X_n$ exist iff $\overline{\lim}\, X_n = \underline{\lim}\, X_n$, then let $\lim X_n = \overline{\lim}\, X_n$.

(7)X is/has

(i)*nontrivial* iff $X \neq \emptyset$ (and not proper class or underlying set);

(ii)*disjoint* iff $\bigcap_{x\in X} x = \emptyset$;

(iii)*finite intersection property (FIP for short)* iff $(\forall \{x_i\}_1^n \subseteq X) \bigcap_1^n x_i \neq \emptyset$.

Remark 2.1.5

(1)Specially for $I = \emptyset$, $\bigcup_{i\in I} X_i = \emptyset$, $\bigcap_{i\in I} X_i$ is proper class or underlying set, $\prod_{i\in I} X_i = \{\emptyset\}$.

(2)If we write $(\exists\{X_i\})Y = \bigsqcup_{i\in I} X_i$, usually it is considered as $(\exists disjoint\{X_i\})Y = \bigcup_{i\in I} X_i$.

$\Diamond$WARNING: *strictly increasing* iff $X_1 \subset X_2 \subset ...$, and we won't define monotone etc below unless something interesting happens.

(3)$P \subseteq \mathcal{P}(X)$ is

(i)*cover* of X iff $X \subseteq \bigcup_{p\in P} p$;

(i)*partition* of X iff $\emptyset \notin P$ and $X = \bigsqcup_{p\in P} p$.

(iii)Q is *refinement* of P iff $(\forall p \in P)(\exists\{q_i\} \subseteq Q)p = \bigcup q_i$.

(iv)$P \wedge Q := \{p \cap q : p \in P, q \in Q, p \cap q \neq \emptyset\}$, similarly $P \vee Q$ is the refinest partition of which P, Q are refinement.

(4)To understand limit of set sequence, $\overline{\lim}_{n\to\infty} X_n = \{x : (\exists n_1 < n_2 < ...)(\forall i)x \in X_{n_i}\}$, $\underline{\lim}_{n\to\infty} X_n = \{x : (\exists n_0)(\forall n > n_0)x \in X_n\}$.

(5)$\mathcal{A}_\delta$ is the set of countable intersections of elements in $\mathcal{A}$ while $\mathcal{A}_\delta$ symbolizes countable unions, denote $(\mathcal{A}_\sigma)_\delta$ by $\mathcal{A}_{\sigma\delta}$. Note $GUVAB$ is usually used for open sets while $FWCDK$ for closed, hence G_δ, F_σ etc. is frequently used.

Definition 2.1.5 (family of set)

For set X, $\emptyset \neq \mathcal{F} \subseteq \mathcal{P}(X)$ is

(1)π-*system* Π iff

$(\forall A, B \in \mathcal{F})A \cap B \in \mathcal{F}$;

(2)λ-*system* Λ iff

(i)$\emptyset \in \mathcal{F}$

(ii)$(\forall A \in \mathcal{F})A^c \in \mathcal{F}$

(iii)$(\forall \{A_i\}_1^\infty \subseteq \mathcal{F}) \bigsqcup_1^\infty A_i \in \mathcal{F}$;

(3)*monotone class* $\mathcal{M}$ iff

(i)$(\forall\{A_i\}_1^\infty \subseteq \mathcal{F})((\forall i)A_i \subseteq A_{i+1}) \to \bigcup_1^\infty A_i \in \mathcal{F}$

(ii)$(\forall\{A_i\}_1^\infty \subseteq \mathcal{F})((\forall i)A_i \supseteq A_{i+1}) \to \bigcap_1^\infty A_i \in \mathcal{F}$;

(4)*semiring* iff

(i)$(\forall A, B \in \mathcal{F})A \cap B \in \mathcal{F}$

(ii)$(\forall A, B \in \mathcal{F})(\exists\{C_i\}_1^n \subseteq \mathcal{F})A \setminus B = \bigsqcup_1^n C_i$;

(5)*ring* $\mathcal{R}$ iff

(i)$(\forall A, B \in \mathcal{F})A \cup B \in \mathcal{F}$

(ii)$(\forall A, B \in \mathcal{F})A \setminus B \in \mathcal{F}$;

(6)δ*-ring* iff

(i)$(\forall A, B \in \mathcal{F})A \cup B \in \mathcal{F}$

(ii)$(\forall A, B \in \mathcal{F})A \setminus B \in \mathcal{F}$

(iii)$(\forall\{A_i\}_1^\infty \subseteq \mathcal{F})\bigcap_1^\infty A_i \in \mathcal{F}$;

(7)σ*-ring* $\Sigma\mathcal{R}$ iff

(i)$(\forall\{A_i\}_1^\infty \subseteq \mathcal{F})\bigcup_1^\infty A_i \in \mathcal{F}$

(ii)$(\forall A, B \in \mathcal{F})A \setminus B \in \mathcal{F}$;

(8)*elementary family (or semialgebra)* iff

(i)$\emptyset \in \mathcal{F}$

(ii)$(\forall A, B \in \mathcal{F})A \cap B \in \mathcal{F}$

(iii)$(\forall A \in \mathcal{F})(\exists\{C_i\}_1^n \subseteq \mathcal{F})A^c = \bigsqcup_1^n C_i$;

(9)*algebra (or field)* $\mathcal{A}$ iff

(i)$\emptyset \in \mathcal{F}$

(ii)$(\forall A \in \mathcal{F})A^c \in \mathcal{F}$

(iii)$(\forall A, B \in \mathcal{F})A \cup B \in \mathcal{F}$;

(10)σ*-algebra* Σ iff

(i)$\emptyset \in \mathcal{F}$

(ii)$(\forall A \in \mathcal{F})A^c \in \mathcal{F}$

(iii)$(\forall\{A_i\}_1^\infty \subseteq \mathcal{F})\bigcup_1^\infty A_i \in \mathcal{F}$;

(11)*filter* iff

(i)$(\forall A, B \in \mathcal{F})A \cap B \in \mathcal{F}$

(ii)(upward closure (or isotony)) $(\forall A \in \mathcal{F})(\forall A \subseteq B \subseteq X)B \in \mathcal{F}$;

(12)*ultrafilter (or maximal filter)* iff

(i)$(\forall A, B \in \mathcal{F})A \cap B \in \mathcal{F}$

(ii)$(\forall A \in \mathcal{F})(\forall A \subseteq B \subseteq X)B \in \mathcal{F}$

(iii)(*proper*)$\emptyset \notin \mathcal{F}$

(iv)$(\forall A \subseteq X)A \in \mathcal{F} \vee A^c \in \mathcal{F}$;

(13)*ideal* iff

(i)$(\forall A, B \in \mathcal{F})A \cup B \in \mathcal{F}$

(ii)$(\forall A \in \mathcal{F})(\forall B \subseteq A)B \in \mathcal{F}$;

(14)*prime ideal* iff

(i)$(\forall A, B \in \mathcal{F})A \cup B \in \mathcal{F}$

(ii)$(\forall A \in \mathcal{F})(\forall B \subseteq A)B \in \mathcal{F}$

(iii)(*proper*)$X \notin \mathcal{F}$

(iv)$(\forall A \subseteq X)A \in \mathcal{F} \vee A^c \in \mathcal{F}$;

(15)*topology* $\mathcal{T}$ iff

(i)$\emptyset, X \in \mathcal{F}$

(ii)$(\forall A, B \in \mathcal{F})A \cap B \in \mathcal{F}$

(iii)$(\forall \{A_i\}_{i \in I} \subseteq \mathcal{F}) \bigcup_{i \in I} A_i \in \mathcal{F}$.

Remark 2.1.6

(1)*Space* is a set X with specific structures on it, e.g. *topological space* $(X, \mathcal{T})$, *measurable space* (X, Σ). Then call the element of X *point (pt for short)*.

(2)Note filter could also be defined by nontrivial, downward directed and upward closure. Since under upward closure, $(\forall A, B \in \mathcal{F})A \cap B \in \mathcal{F}$ iff $(\forall A, B \in \mathcal{F})(\exists C \in \mathcal{F})C \subseteq A, B$.

(3)Note for algebra generated by semialgebra, just add all finite disjoint unions. We will use it to induce a measure on an algebra from a measure on a semialgebra. For more construction sequences, see below.

Theorem 2.1.2 (monotone class theorem)

For algebra $\mathcal{A}$ on X, $\mathcal{M}(\mathcal{A}) = \Sigma(\mathcal{A})$.

Proof. By $\mathcal{M}(\mathcal{A}) \subseteq \mathcal{M}(\Sigma(\mathcal{A})) = \Sigma(\mathcal{A})$, suffice to show $\mathcal{M}(\mathcal{A}) = \Sigma(\mathcal{M}(\mathcal{A}))$. Note a class that is both algebra and monotone class is a σ-algebra by $\bigcup_1^\infty E_i = \bigcup_{i=1}^\infty (\bigcup_{j=1}^i E_j)$, suffice to show $\mathcal{M} := \mathcal{M}(\mathcal{A})$ is an algebra. For $E \in \mathcal{C}$, define $\mathcal{M}(E) = \{F \in \mathcal{M} : E \setminus F, F \setminus E, E \cap F \in \mathcal{M}\}$, note $\mathcal{M}(E)$ is a monotone class and $E \in \mathcal{M}(F) \Leftrightarrow F \in \mathcal{M}(E)$. If $E \in \mathcal{A}$, then $\mathcal{A} \subseteq \mathcal{M}(E)$ hence $(\forall E \in \mathcal{A})\mathcal{M} \subseteq \mathcal{M}(E)$. Hence if $F \in \mathcal{M}$, then $(\forall E \in \mathcal{A})F \in \mathcal{M}(E)$, i.e. $(\forall E \in \mathcal{A})E \in \mathcal{M}(F)$ so $\mathcal{A} \subseteq \mathcal{M}(F)$. Therefore $(\forall F \in \mathcal{M})\mathcal{M} \subseteq \mathcal{M}(F)$, i.e. $\mathcal{M}$ is closed under difference and intersection. □

Remark 2.1.7

(1)In the proof above, we use the outer way and a 2-ary version of its common trick (Trick: to prove all elements of a specific type satisfy property P, pick the set of all elements satisfy P, then show that set is the same type and contains all generators). To prove $\mathcal{M}(\mathcal{A})$ is closed under complement, the outer way uses $\mathcal{M}_1 := \{E \in \mathcal{M} : E^c \in \mathcal{M}\}$ and $\mathcal{M} \subseteq \mathcal{M}_1$ by showing $\mathcal{A} \subseteq \mathcal{M}_1$ and $\mathcal{M}_1$ is a monotone class, while the inner way uses transfinite induction to its construction sequence.

(2)For $\mathcal{F} \subseteq \mathcal{P}(X)$,

(i)$\Pi(\mathcal{F}) = \{\bigcap_1^n F_i : (\forall i)F_i \in \mathcal{F}\}$

(ii)$\mathcal{R}(\mathcal{F}) = \{\pi_1 \triangle ... \triangle \pi_n : (\forall i)\pi_i \in \Pi(\mathcal{F})\}$

(iii)$\mathcal{A}(\mathcal{F}) = \mathcal{R}(\mathcal{F} \cup \{X\}) = \mathcal{R}(\mathcal{F}) \cup \{X \setminus R : R \in \mathcal{R}(\mathcal{F})\}$

(iv)Define $\mathcal{F}^* = \{\bigcup_1^\infty (A_i \setminus B_i) : (\forall i) A_i, B_i \in \mathcal{F} \cup \{\emptyset\}\}$, recursively define $\mathcal{F}_0 = \mathcal{F}$ and $\mathcal{F}_\beta = (\bigcup_{\alpha<\beta} \mathcal{F}_\alpha)^*$, then $\Sigma\mathcal{R}(\mathcal{F}) = \bigcup_{\alpha<\omega_1} \mathcal{F}_\alpha$

(v)$\Sigma(\mathcal{F}) = \bigcup_{\alpha<\omega_1} (\mathcal{F} \cup \{X\})_\alpha$.

Btw, with (iv) and transfinite induction, we can get $|\mathcal{B}(\mathbb{R}^n)| = \aleph$.

Definition 2.1.6 (relation and function)

(1)R is a *n-ary relation* on $X_1, ..., X_n$ iff $R \subseteq X_1 \times ... \times X_n$, denote $(x_1, ..., x_n) \in R$ by $R(x_1...x_n)$.

(i)*domain* $dom(R) := \{\bar{x} : (\exists \bar{y}) R\bar{x}\bar{y}\}$

(ii)*range* $ran(R) := \{\bar{y} : (\exists \bar{x}) R\bar{x}\bar{y}\}$

(iii)*image* $R(X) := \{\bar{y} : (\exists \bar{x} \in X) R\bar{x}\bar{y}\}$

(iv)*inverse image* $R^{-1}(Y) := \{\bar{x} : (\exists \bar{y} \in Y) R\bar{x}\bar{y}\}$

(v)*inverse* $R^{-1} := \{(\bar{x}, \bar{y}) | R\bar{y}\bar{x}\}$

(vi)*composition* $S \circ R := \{(\bar{x}, \bar{z}) | (\exists \bar{y}) R\bar{x}\bar{y} \wedge S\bar{y}\bar{z}\}$

(2)Relation f is a *function* iff $(\forall \bar{x})(\exists! \bar{y}) f\bar{x}\bar{y}$, denote $f\bar{x}\bar{y}$ by $f(\bar{x}) = \bar{y}$ and f from $X = dom(f)$ to $Y \supseteq ran(f)$ by $f : X \to Y, \bar{x} \mapsto \bar{y}$.

(i)*restriction* of $f : X \to Y$ to $Z \subseteq X$ is $f|_Z : Z \to Y, z \mapsto f(z)$;

(ii)*extension* of f is g iff restriction of g is f;

(iii)f is

(a)*injective* iff $\forall x_1, x_2 \in X(f(x_1) = f(x_2) \to x_1 = x_2)$;

(b)*surjective* iff $(\forall y \in Y)(\exists x \in X) f(x) = y$;

(c)*bijective* iff f is injective and surjective.

Remark 2.1.8

(1)Especially in recursive theory, for $f : X \to Y$, $dom(f)$ can be a subset of X, then for $x \in X$, $f(x) \downarrow$ (resp. $f(x) \uparrow$) iff $x \in dom(f)$. f is

(i)*total* iff $(\forall x \in X) f(x) \downarrow$

(ii)*partial* iff $(\exists x \in X) f(x) \uparrow$.

(2)For $f : X \to Y$, it naturally induces two set function $f : \mathcal{P}(X) \to \mathcal{P}(Y)$ and $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$, though $f^{-1} : Y \to X$ not always exists.

(3)Two ordered sets X and Y are said to have the same order type iff there exists a order isomorphism between X and Y.

Example 2.1.1

(1)Specific

(i)*Kronecker symbol* $\delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

(ii)*characteristic function* $\chi_A(x) := \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$

(iii)*indexing function* for X is surjective $f : I \to X$, denote X by $\{x_i\}_{i \in I}$

(iv)*identical map* $id_x : X \to X, x \mapsto x$

(v)*constant map* $c_{y_0} : X \to Y, x \mapsto y_0$

(vi)*inclusion map* from $X \subseteq Y$ to Y is $id_Y|_X$, written $\iota : X \hookrightarrow Y$

(vii)*projection map* from $X \supseteq Y$ to Y is $p : Y \to X$ s.t. $p|_X \circ p = p$

(viii)*quotient map (or canonical map)* $\pi : X \to X/\sim, x \mapsto [x]$

(ix)*n-ary operation* on X is $f : X^n \to X$

(2)For $\alpha : M \to N$ where $M \in \mathcal{M}$ and $N \in \mathcal{N}$, α is

(i)*homomorphism* iff

(a)$(\forall f \in \mathcal{L})(\forall a_1, ..., a_{n_f} \in M)\alpha(f^{\mathcal{M}}(a_1, ..., a_{n_f})) = f^{\mathcal{N}}(\alpha(a_1), ..., \alpha(a_{n_f}))$

(b)$((\forall R \in \mathcal{L})(\forall a_1, ..., a_{n_R} \in M)R^{\mathcal{M}}(a_1, ..., a_{n_f}) \leftrightarrow R^{\mathcal{N}}(\alpha(a_1), ..., \alpha(a_{n_f})))$

denote the set of homomorphisms from $\mathcal{M}$ to $\mathcal{N}$ by $Hom(M, N)$ and $End(M) := Hom(M, M)$ whose element is *endomorphism*;

(ii)*embedding (or monomorphism)* iff α is homomorphism and injective;

(iii)*epimorphism* iff α is homomorphism and surjective;

(iv)*isomorphism* iff α is homomorphism and bijective, denote the set of isomorphisms by $Isom(M, N)$ and $Aut(M) := Isom(M, M)$ whose element is *automorphism*.

$\mathcal{M}$ is *isomorphic* to $\mathcal{N}$, written $\mathcal{M} \cong \mathcal{N}$, iff $Isom(M, N) \neq \emptyset$.

Definition 2.1.7 (sub)

(1)$\mathcal{M}$ is *substructure* of $\mathcal{N}$ (or $\mathcal{N}$ is *extension* of $\mathcal{M}$), iff $\emptyset \neq M \subseteq N$ and the inclusion map is embedding.

(2)*up to isomorphism*

We have the following chains of inclusions for continuous functions over a closed, bounded interval of the real line: Continuously differentiable $\subseteq$ Lipschitz continuous $\subseteq$ absolutely continuous $\subseteq$ continuous and bounded variation $\subseteq$ differentiable almost everywhere

Lipschitz continuous that are everywhere differentiable but not continuously differentiable: $f(x) = \begin{cases} x^2 \sim (1/x) & x \neq 0 \\ 0 & o.w. \end{cases}$

(1)For $f : \mathbb{R}^m \to \mathbb{R}^n$ *differentiable* at x_0 iff exist a linear map $J : \mathbb{R}^m \to \mathbb{R}^n$ s.t. $\lim_{h \to 0} \frac{||f(x_0+h) - f(x_0) - J(h)||_{\mathbb{R}^n}}{||h||_{\mathbb{R}^m}} = 0$ (If a function is differentiable at x_0, then all of the partial derivatives exist at x_0, and J is given by Jacobian matrix, $n \times m$) (If all the partial derivatives of a function exist in a neighborhood of a pt x_0 and are continuous at x_0, then the function is differentiable at x_0. However, the existence of the partial derivatives (or even all the directional derivatives) doesn't guarantee differential at pt. E.g. $f(x,y) = \begin{cases} x & y \neq x^2 \\ 0 & o.w. \end{cases}$ not differentiable at $(0,0)$ but all the partial deirvatives and directional derivatives exist. For a continuous example, $f(x,y) = \begin{cases} y^3/(x^2+y^2) & (x,y) \neq (0,0) \\ 0 & o.w. \end{cases}$ not differentiable at $(0,0)$ but all the partial deirvatives and directional derivatives exist.)

For complex-valued $f : \mathbb{C} \to \mathbb{C}$, *differentiable* at $a \in \mathbb{R}$ iff $f'(a) = \lim_{h \to 0, h \in \mathbb{C}} \frac{f(a+h)-f(a)}{h}$ exist. Although it looks similar to real-valued, however more restrictive condition. A function $f : \mathbb{C} \to \mathbb{C}$ that is complex-differentiable at a is automatically differentiable at a when viewed as a function $f : \mathbb{R}^2 \to \mathbb{R}^2$. This is because complex-differentiability implies $\lim_{h \to 0, h \in \mathbb{C}} \frac{|f(a+h)-f(a)-f'(a)h|}{|h|}$ exist. However the converse is wrong, counterexample $f(z) = \frac{z+\overline{z}}{2}$

Any function that is complex-differentiable in a neighborhood of a point is called holomorphic at that point. Such a function is necessarily infinitely differentiable, and in fact analytic.

The key of difference between complex-differentiable and real-differentiable is (???$\mathbb{C}$ has multiplication structure while $\mathbb{R}^2$ not). The complex-differentiable functions locally look not only like a linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$, but like a linear transformation corresponding to multiplication by a complex number, which we can identify with linear transformations of the plane by the obvious multiplication $(a,b) \cdot (c,d) = (ac - bd, ad + bc)$, so the matrix must have the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Such functions have many properties, which gives complex analysis a much different flavor. For example, if we view complex functions on (simply connected subsets of) the complex plane as vector fields on $\mathbb{R}^2$, then complex-differentiable functions will be conservative vector fields. In other words, derivatives are linear maps that approximate the function within o. If f is differentiable at x, then $f(x+t) = f(x) + A(t) + o(t)$ where A is a linear map and $\lim_{t \to 0} \frac{o(t)}{|t|} = 0$, note $|t|$ is the same as a complex or $\mathbb{R}^2$. But in $\mathbb{R}^2$, A can be any linear map, i.e. 2×2 matrix with entries in $\mathbb{R}$. In complex case, A is linear map of 1-dimensional complex vector, i.e. 1×1 matrix with entries in $\mathbb{C}$. if we think $\mathbb{C}$ as $\mathbb{R}^2$, then A is a rotation and scaling, not every 2×2 linear map is of this form.

(2)For a real-valued (pr complex-valued) (?not need? continuous) function f, 1-dimensional *total variation* on $[a,b] \subseteq \mathbb{R}$ is $V_a^b(f) = \sup_{\mathcal{P}} \sum_{i=0}^{n_{\mathcal{P}}-1} |f(x_{i+1}) - f(x_i)|$ where partition $\mathcal{P} = \{x_0, ..., x_{n_{\mathcal{P}}}\}$ n-dimensional Let Ω be an open subset of $\mathbb{R}^n$, $f \in L^1(\Omega)$, *total variation* of f in Ω is $V(f,\Omega) = \sup\{\int_\Omega f(x) \text{div } \varphi(x) dx : \varphi \in C_c^1(\Omega, \mathbb{R}^n), ||\varphi||_{L^\infty(\Omega)} \leq 1\}$ Note note require bounded (1)$C_c^1(\Omega, \mathbb{R}^n)$ the set of continuously differentiable vector functions of compact support contained in Ω (2)$||\,||_{L^\infty(\Omega)}$ is the essential supremum norm (3)div is the divergence operator

A continuous real-valued f on $\mathbb{R}$ is *bounded variation (BV function)* on $[a,b] \subseteq \mathbb{R}$ iff its total variation is finite

For n-dimensional, two equivalent definition.

(3)*absolutely continuous* I interval of $\mathbb{R}$, $f : I \to \mathbb{R}$ is *absolutely continuous* on I if $\forall \epsilon > 0 \exists \delta > 0$ s.t. $\forall N$, $\forall \{(x_k, y_k) : x_k < y_k \in I\}$ disjoint satisfies $\sum_{k=1}^N (y_k - x_k) < \delta$, then $\sum_{k=1}^N |f(y_k) - f(x_k)| < \epsilon$

absolute continuous $\subseteq$ uniformly continuous

uniformly continuous iff $\forall \epsilon > 0 \exists \delta > 0$ s.t. $\forall |x-y| < \delta$, $|f(x) - f(y)| < \epsilon$

(4)For metric spaces $(X, d_X), (Y, d_Y)$, $f : X \to Y$ is *Lipschitz continuous* if $\exists$ real constant $K \geq 0$ s.t. $\forall x_1, x_2 \in$

X, $d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2)$

α-Hölder continuous if $\exists$ real constant $K \geq 0$ s.t. $\forall x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2)^\alpha$

locally Lipschitz continuous iff for any $x \in X$, exists neighborhood $U \ni x$ s.t. $f|_U$ is Lipschitz continuous

(5)f is *continuously differentiable* if $f'(x)$ exist and is continuous. Continuous functions are said to be of class C^0, continuous differentiable functions C^1, A function of class C^2 if the 1st and 2nd derivative of the function both exist and are continuous. C^∞ *smooth* f^n exist for all n

f is *differentiable* iff the derivative exists at every point in its domain (f is differentiable at x then f must be continuous at x)

Let $G = (g_1, ...g_n)$ be a map from an open set $\Omega \subseteq \mathbb{R}^n$ into $\mathbb{R}^n$ whose components g_j are of class C^1

Denote $((\partial g_i/\partial x_j)(x))$ of linear map partial derivatives at x by $D_x G$. Observe that if G is a linear transformation viz. matrix, then $D_x G = G$ for all x

G is a C^r *diffeomorphism* if G bijection and $G, G^{-1} \in C^r$ (*diffeomorphism* means C^∞ diffeomorphism)

Remark 2.1.9

By Inverse Function Theorem, $f : \Omega \to f(\Omega)$ is C^1 diffeomorphism only need $f \in C^1$, f injective, $J_f = D_x f$ is invertible for all $x \in \Omega$

For C^1 diffeomorphism, only need G is injective and $D_x G$ is invertible for all $x \in \Omega$

a (topological) manifold is a second countable Hausdorff space that is locally homeomorphic to a Euclidean space.

Theorem 2.1.3 (Inverse Function Theorem)

(1)(local) For open $U \subseteq \mathbb{R}^n, V \subseteq \mathbb{R}^m$, $f : U \to V \in C^k$, if $J_f(a)$ is injective for some $a \in U$, then there exists an open nbd $A \subseteq U$ of a and $V \supseteq B \supseteq f(A)$ s.t. $f : A \to B$ is bijective and $f^{-1} : B \to A \in C^k$ $(\det(J_{f^{-1}}(a)) = 1/\det(J_f(a)))$.

(2)(global) For open $U \subseteq \mathbb{R}^n, V \subseteq \mathbb{R}^m$ (more generally, manifold), $f : U \to V \in C^k$, if f is injective on a closed subset $A \subseteq U$ and J_f (the Jacobian matrix, $J_f = \nabla^T f = D_x f = (\frac{\partial f_i}{\partial x_j})_{ij}$) is injective for all $a \in A$, then f is injective in an open nbd A' of A and $f^{-1} : f(A') \to A' \in C^k$ $(D_x(f^{-1}) = [D_{f^{-1}(x)} f]^{-1}$ for all $x \in f(\Omega))$

Lemma 2.1.4 (extend local into global)

For A is a closed subset of a topological manifold X (more generally, a topological space, admitting an exhaustion by compact subsets), topological space Z, if $f : X \to Z$ is a local homeomorphism that is injective on A, then f is injective on some open nbd of A

The mapping $x_0 \mapsto f(x_0)$ is a function, where x_0 is an argument of a function f. At the same time, the mapping $f \mapsto f(x_0)$ is a *functional* where x_0 is a parameter. Provided that f is a linear function from a vector space to the underlying scalar field (the set of all linear function from V to F is also a vector space, called *(algebraic) dual space*, written $Hom(V, F)$ or V^*), the above linear maps are dual to each other, and in functional analysis both are called linear functionals. linear functional $T : V^* \to F$ is *positive* iff $(\forall f \in V^* \wedge f \geq 0) Tf \geq 0$

Riesz: To every positive linear fucntional T on C there corresponds a finite positive Borel measure μ on I s.t. $Tf = \int_I f d\mu$ $(f \in C)$ the converse is obvious

Hilbert space is a real or complex inner space that is also a complete metric space with respect to the distance function induced by the inner product.

Theorem 2.1.5 (For convex)

(1)f *convex* iff $E = \{(x, y) : y \geq f(x), x \in \Omega\}$ convex

Assume supporting hyperplane at $(x_0, f(x_0))$ is $< \eta, x - x_0 > + < \gamma, y - f(x_0) >= 0$, then $\gamma < 0$ and $f(x) \geq f(x_0) - \frac{1}{\gamma}\eta \cdot (x - x_0), \forall x \in \Omega$ (2)for convex A, $x_0 \in \partial A$, $\exists \eta \in \mathbb{R}^n$ $< \eta, x - x_0 >\leq 0$ for all $x \in A$ ($< \eta, x - x_0 >= 0$ is the *supporting hyperplane* at x_0) (3)(hyperplane separation theorem) $A, B \subseteq \mathbb{R}^n$ convex disjoint, then $\exists \eta \in \mathbb{R}^n, c \in \mathbb{R}$ s.t. $< x, \eta >\geq c, < y, \eta >\leq c$ for all $x \in A, y \in B$ (4)

3 Elementary Algebra

3.1 Group, Ring, and Field

3.2 Advanced Linear Algebra

3.3 Commutative Algebra

3.4 Module Theory, Homology Algebra

3.5 Representation Theory

3.6 number theory

Definition 3.6.1 (ordinal and cardinal)

(1)Set X is an *ordinal* iff $(X, \subseteq)$ is strictly well-ordered and $(\forall x \in X)x \subseteq X$, denote the class of all ordinals by Ord.

0, *successor ordinal* iff $(\exists Y \in Ord)X = Y \cup \{Y\}$, *limit ordinal* iff $(\forall Y \in Ord \wedge Y < X)(\exists Z \in Ord)Y < Z < X$.

(2)The *cardinal (or initial ordinal)* of a set X, written $card(X)$ or $|X|$, is the least ordinal number α s.t. there exists a bijection between X and α.

The α-th infinite initial ordinal is written ω_α, and its cardinality is written $\aleph_alpha$.

(3)The *cofinality* of a partial order set P, written $cf(P)$, is the least cardinal of all cofinal subsets of P. An ordinal α is *regular* iff $\alpha = cf(\alpha)$.

(4)*Ordinal-indexed sequence* is a function from ordinal α to set X, specially *(ordinary) sequence* when $\alpha = \omega$. *sequence* in X is $f : \mathbb{Z}_{>0} \to X$, written $\{x_n\}_1^\infty$.

subsequence of f is $f \circ g$ where $g : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ s.t. $(\forall n < m)g(n) < g(m)$.

(2)$\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. (3)For $a, b \in \bar{\mathbb{R}}$,

Remark 3.6.1

(1)ω and $\omega + 1$ have the same cardinal but not the same order type, since there not exists order isomorphic but bijection between them. Bijection could be $f : \omega+1 \to \omega, x \mapsto \begin{cases} 0 & x = \omega \\ x+1 & o.w. \end{cases}$, note order isomorphism preserves the existence of a maximal element so no order isomorphism exists. Moreover, if two ordinals are order-isomorphic then they are equal.

(2)Under the order topology, a limit ordinal is the limit in a topological sense of all smaller ordinals. (3)ω, ω_1(the first uncountable ordinal),...

Example 3.6.1

A discription of ordinal is $0 := \emptyset$, $1 := \{\emptyset\}$, $2 := \{\emptyset, \{\emptyset\}\}$, $3 := \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$

Theorem 3.6.1 (transfinite induction)

If $P(\alpha)$ is true whenever $P(\beta)$ is true for all $\beta < \alpha$, then $P(\alpha)$ is true for all α

Remark 3.6.2

Every ordinal is either zero or a successor or a limit. This distinction is important, because many definitions by transfinite recursion rely upon it. E.g. when defining a function F by transfinite recursion on all ordinals, one defines $F(0)$, then $F(\alpha+1)$ assuming $F(\alpha)$ is defined, then for limit ordinals $F(\beta)$ as the limit of the $F(\alpha)$ for all $\alpha < \beta$.

Definition 3.6.2

(ordinal and cardinal arithmetic) For ordinal, (1)Addition: $\alpha + \beta$ is given by transfinite recursion on β
(i)$\alpha + 0 = \alpha$ (ii)$\alpha + S(\beta) = S(\alpha+\beta)$ for a successor ordinal $S(\beta)$ (iii)$\alpha + \beta = \bigcup_{\delta<\beta}(\alpha+\delta)$ when β is a limit ordinal. (2)Multiplication: $\alpha + \beta$ is given by transfinite recursion on β
(i)$\alpha \cdot 0 = 0$ (ii)$\alpha \cdot S(\beta) = (\alpha\cdot\beta)+\alpha$ for a successor ordinal $S(\beta)$ (iii)$\alpha\cdot\beta = \bigcup_{\delta<\beta}(\alpha\cdot\delta)$ when β is a limit ordinal. (3)Exponentiation: α^β is given by transfinite recursion on β
(i)$\alpha^0 = 1$ (ii)$\alpha^{S(\beta)} = (\alpha^\beta)\cdot\alpha$ for a successor ordinal $S(\beta)$ (iii)$\alpha^\beta = \bigcup_{\delta<\beta}(\alpha^\delta)$ when β is a limit ordinal.
For cardinal, (1)Addition: $|X|+|Y| = |X \cup Y|$ (2)Multiplication: $|X|\cdot|Y| = |X \times Y|$ (3)Exponentiation: $|X|^|Y| = |X^Y|$ where X^Y denote the set of all functions from Y to X

Remark 3.6.3

(1)Note ordinal addition is not commutative, e.g. $1+\omega = \omega \neq \omega+1$, and only left-cancellative.

(2)Every ordinal number α can be uniquely written as *cantor normal form* $\omega^{\beta_1}c_1 + ... + \omega^{\beta_k}c_k$ where $k \in \omega$, $c_i \in \omega \setminus \{0\}$, ordinal $\beta_1 > \beta_2 > ... > \beta_k \geq 0$

real line $(\mathbb{R}, +, \cdot, <)$ where $+, \cdot, <$ satisfy the axioms of a complete archimedeaan ordered field
arithmetic of extend real number: note $0 \cdot \infty = 0$ in measure theory and probability theory

Theorem 3.6.2 (open set decomposition in $\mathbb{R}$)

(1)Let partition $P = \{[a_i, b_i] : 0 \leq i \leq n, a = a_0 < b_0 = a_1 < ... < b_n = b\}$ of $[a,b]$ as *tagged partition* $\{a_i, b : 0 \leq i \leq n\}$. *axis-parallel partition* of $R = \Pi_1^n[a_i,b_i] \subseteq \mathbb{R}^n$ is $P = P_1 \times ... \times P_n = \Pi_1^n\{a_{i,j} : 0 \leq j \leq m_i, a_i = a_{i,0} < ... < a_{i,m_i} = b_i\}$.

(hyper)volume of R $|R| := \Pi_1^n(b_i - a_i)$. *Norm* of P $||P|| := \max_{1\leq i\leq n, 1\leq j\leq m_i}\{a_{i,j} - a_{i,j-1}\}$. Q is *refinement* of P iff $P \subseteq Q$. $P \uplus Q := \Pi_1^n(P_i \cup Q_i)$ where $P = \Pi_1^n P_i, Q = \Pi_1^n Q_i$.

cube in $\mathbb{R}^n$ is a product of n closed intervals whose side length are all equal *n-dimensional interval (or cuboid)* $< a, b >= \Pi_1^n < a_i, b_i >$, also define open, closed, right-open. *right-open 2-adic cuboid* $2^k((j_1, ..., j_n)^T + [0,1)^n)$ where $k, j_1, ..., j_n \in \mathbb{Z}$ for convenience in harmonic analysis.

domain connected and open

$(\forall \mathbb{R}^n \ni U\ open)(\exists \{E_i\}_1^\infty \subseteq \mathbb{R}^n)U = \bigsqcup E_i$ where E could be any h-cube or always pick 2-p h-interval?. Moverover, if $m(U)<\infty$, then $(\exists n\in\mathbb{N})m(U\triangle\bigsqcup_1^n I_i)<\epsilon$.

(1)(For $n>1$, the result not holds) Every open set in $\mathbb{R}$ is a countable disjoint union of open intervals.

(2)$\forall G\in\mathbb{R}^n$ open, then G is a countable disjoint union of right-open binary cuboid (moreover for open, closed, left-open $\forall G\in\bar{\mathbb{R}}^n$ closed cuboid).

Proof. (1)$\forall x\in U$, define $(a_x,b_x)=\bigcup_{x\in(a,b)\subseteq U}(a,b)$, then $U=\bigcup_{x\in U}(a_x,b_x)=\bigsqcup_{countable}(a_n,b_n)$ since (i)if $x\in(a,b)\cap(c,d)$, then $(a,b)=(c,d)=(a_x,b_x)$ (ii)$f:\{(a_x,b_x)\}\to\mathbb{Q},(a_x,b_x)\mapsto y$ for some $y\in(a_x,b_x)$ injective.

(Lou Analysis)Prove for right-open 2-adic cuboid Let the collection of right-open 2-adic cuboids be $\mathcal{D}(\mathbb{R}^n)$. The biggest advantages for it is $(\forall A,B\in\mathcal{D}(\mathbb{R}^n))A\cap B=\emptyset\vee A\subseteq B\vee B\subseteq A$. For any open $V\subseteq\mathbb{R}^n$, for any $x\in V$, let $A(x)$ be the largest set A in $\mathcal{D}(\mathbb{R}^n)$ s.t. $x\in A\subseteq V$, then $A(x)$ exists and is unique. Note for any $x,y\in V$, $A(x),A(y)$ are disjoint or equal, so by there must be rational pt in $A(x)$, get $V=\bigsqcup_\omega A(x)$.

□

Theorem 3.6.3 (Vitali Covering Lemma)

(1)(finite) Let $B_1,\dots,B_n$ be any finite collection of balls contained in an arbitrary metric space. Then there exists a subcollection $B_{j_1},\dots,B_{j_m}$ of these balls which are disjoint and satisfy $B_1\cup\dots\cup B_n\subseteq 3B_{j_1}\cup\dots\cup 3B_{j_m}$

(2)(infinite) Let F be an arbitrary collection of balls in a separable metric space s.t. $R:=\sup\{\text{rad }(B):B\in F\}<\infty$ where rad (B) denotes the radius of the ball B. Then there exists a countable sub-collection $G\subseteq F$ s.t. the balls of G are pairwise disjoint and satisfy $\bigcup_{B\in F}B\subseteq\bigcup_{C\in G}5C$. And moreover, each $B\in F$ intersects some $C\in G$ with $B\subseteq 5C$

Proof. (1)WLOG assume $n>0$. Let B_{j_1} be the ball of largest radius. Once $\{B_{j_i}\}_{i=1}^k$ are chosen, if there is some ball in $B_1,\dots,B_n$ disjoint from $\bigsqcup_{i=1}^k B_{j_i}$, then let $B_{j_{k+1}}$ be such ball with maximal radius, o.w. set $m=k$ and terminate.

$\forall B_i$, there exists the smallest $1\le k\le m$ s.t. $B_i\cap B_{j_k}\ne\emptyset$, then $B_i\subseteq 3B_{j_k}$.

(2) □

4 General Topology

4.1

5 Basic Analysis

5.1 Measure Theory

Sobolev space, Radon measure, Hausdorff measure

5.2 Integration and Differentiation

5.3 Complex Analysis

Holomorphic and meromorphic function, conformal map, linear fractional transformation, Schwarz's lemma

Complex integral: Cauchy's theorem, Cauchy integral formula, residues

Harmonic functions: the mean value property, the reflection principal, Dirichlet's problem

series adn product developments: Laurent series, partial fractions expansions, and canonical products

Special functions: Gamma function, Zeta function, elliptic function

Basics of Riemann surfaces, Riemann mapping theorem, Picard theorem

5.4 Measure

Definition 5.4.1 (measure)

For set X, $\mathcal{M} \subseteq \mathcal{P}(X)$, a *(positive) measure* on $(X, \mathcal{M})$ is $\mu : \mathcal{M} \to [0, \infty]$ s.t.

①$\mu(\emptyset) = 0$

②(countably additive of disjoint) $(\forall \{E_i\}_1^\infty \subseteq \mathcal{M}) \bigsqcup_1^\infty E_i \in \mathcal{M} \to \mu(\bigcup_1^\infty E_i) = \sum_1^\infty \mu(E_i)$.

Then call $(X, \mathcal{M}, \mu)$ *measure space*.

(1)A *content* on $(X, \mathcal{M})$ is (finite additive) (2)A *premeasure* is a measure on a semiring, i.e. $\mathcal{M}$ is a semiring.

(3)A *outer measure* on X is $\mu : \mathcal{P}(X) \to [0, \infty]$ s.t.

①③$(\forall \{E_i\}_1^\infty \subseteq \mathcal{M}, F \in \mathcal{M}) F \subseteq \bigcup_1^\infty E_i \in \mathcal{M} \to \mu(F) \leq \sum_1^\infty \mu(E_i)$.

(4)A *signed measure* on $(X, \mathcal{M})$ is $\mu : \mathcal{M} \to \overline{\mathbb{R}}$ s.t.

①④$(\forall \{E_i\}_1^\infty \subseteq \mathcal{M}) \bigsqcup_1^\infty E_i \in \mathcal{M} \to \mu(\bigcup_1^\infty E_i) = \sum_1^\infty \mu(E_i)$ where $\sum_1^\infty \mu(E_i)$ converges absolutely if $\mu(\bigcup_1^\infty E_j)$ is finite

⑤$\neg \exists E, F \in \mathcal{M}(\mu(E) = \infty \wedge \mu(F) = -\infty)$.

(5)A *complex measure* on $(X, \mathcal{M})$ is $\mu : \mathcal{M} \to \mathbb{C}$ s.t.

⑥$(\forall \{E_i\}_1^\infty \subseteq \mathcal{M}) \bigsqcup_1^\infty E_i \in \mathcal{M} \to \mu(\bigcup_1^\infty E_i) = \sum_1^\infty \mu(E_i)$ where $\sum_1^\infty \mu(E_i)$ converges absolutely.

Remark 5.4.1

(1)Generally speaking, a measure usually is defined on a σ-algebra. We adopt this notion below for convenience. So are signed measure and complex measure. And define *measurable space* is (X, Σ) where Σ is a σ-algebra on X.

(2)We will see the extension of a measure, measure on semialgebra $\Rightarrow$ measure on algebra $\Rightarrow$ outer measure ⇒ measure on *sigma*-algebra ⇒ complete measure. Also we will see the decomposition of a measure, complex measure ⇒ signed measure ⇒ positive measure.

(3)For outer measure, note ③ can be replaced by monotone and countably subadditive.

(4)For signed measure, absolute convergence in ④ is for rearrangement while ⑤ is for avoiding the undefined $-\infty+\infty$, note $0\cdot\infty=0$ in measure theory.

(5)For complex measure, $\mu(\emptyset)=0$ is redundant since $\mu(\emptyset)=\mu(\emptyset)+\mu(\emptyset)\to\mu(\emptyset)=0$ holds in $\mathbb{C}$.

(6)

Definition 5.4.2 (classification of measure and set)

(A)For measure space $(X,\mathcal{M},\mu)$,

(1)μ is

(i)*finite* iff $\mu(X)<\infty$.

(ii)*σ-finite* iff $(\exists\{E_i\}_1^\infty\subseteq\mathcal{M})X=\bigcup E_i\wedge(\forall i)\mu(E_i)<\infty$.

(iii)*semifinite* iff $(\forall E\in\mathcal{M})(\mu(E)=\infty\to(\exists E\supseteq F\in\mathcal{M})0<\mu(F)<\infty)$.

(iv)*complete* iff $(\forall E\in\mathcal{M})(\mu(E)=0\to(\forall E\supseteq F\in\mathcal{P}(X))F\in\mathcal{M})$.

(v)*saturated* iff every local measurable set is measurable.

(2)A is

(i)*(μ-)measurable* o.w. *nonmeasurable* iff $A\in\mathcal{M}$.

(ii)*locally measurable* iff $(\forall E\in\mathcal{M})(\mu(E)<\infty\to A\cap E\in\mathcal{M})$.

(iii)*(μ-)null set* iff $A\in\mathcal{M}\wedge\mu(A)=0$.

(iv)*(μ-)almost everywhere ((μ-)a.e. for short)* iff A^c is a null set.

(B)For topological space $(X,\mathcal{T})$ and measure on σ-algebra $\Sigma\subseteq\mathcal{P}(X)$,

(1)*mu* is

(i)*locally finite* iff $(\forall x\in X)(\exists x\in E\in\Sigma)\mu(E)<\infty$.

(ii)*inner regular (resp. outer regular)* iff every measurable set is inner regular (resp. outer regular).

(iii)*regular* iff μ is inner regular and outer regular.

(2)A is

(i)*inner regular (resp. outer regular)* iff $A\in\Sigma\wedge\mu(A)=\sup\{\mu(F):F\subseteq A,F\in\Sigma,F\textit{ compact}\}$ (resp. $A\in\Sigma\wedge\mu(A)=\sup\{\mu(G):G\supseteq A,F\in\Sigma,F\in\mathcal{T}\}$).

Remark 5.4.2

(1)For *sigma*-finite measure, we have a trick to only consider the "local part" of finite measure. E.g. pick Lebesgue measure $E\subseteq\mathbb{R}$, then we can consider the property of $E\cap[0,1]$ since $\mathbb{R}=\bigcup[i,i+1]$, which can usually be the reason of "wlog assume E is bounded".

(2)For outer measure μ, $A\subseteq X$ is *μ-measurable (or Carathéodory measurable)* iff $(\forall E\subseteq X)\mu(E)=\mu(E\cap A)+\mu(E\cap A^c)$. Note it is equivalent to $(\forall E\subseteq X)\mu(E)\geq\mu(E\cap A)+\mu(E\cap A^c)$.

And after giving the Lebesgue measure definition, measurable usually means Lebesgue measurable.

(2)

For signed measure ν on $(X,\mathcal{A})$, then $E\in\mathcal{A}$ is *positive (resp. negative, null)* for ν if $(\forall F\in\mathcal{A}\wedge F\subseteq E)\nu(F)\geq 0$ (resp. $\nu(F)\leq 0$, $\nu(F)=0$) We shall see "every signed measure can be represented in either of these two forms": (i)$\nu=\mu_1-\mu_2$ where μ_1,μ_2 measure on $\mathcal{A}$ and at least one of them is finite (ii)μ measure on $\mathcal{A}$, $f:X\to[-\infty,\infty]$ measurable function s.t. at least one of $\int f^+d\mu$ and $\int f^-d\mu$ is finite (*extended μ-integrable function*), then $\nu(E)=\int_E fd\mu$ is a signed measure ()

Proposition 5.4.1

For measure space $(X,\mathcal{M},\mu)$,

(1)(monotonic) $(\forall E,F\in\mathcal{M})E\subseteq F\to\mu(E)\leq\mu(F)$.

(2)(countably subadditive) $(\forall\{E_i\}_1^\infty\subseteq\mathcal{M})\mu(\bigcup E_i)\leq\sum\mu(E_i)$.

(3)(continuous from below) $(\forall\{E_i\}_1^\infty\subseteq\mathcal{M})((\forall i)E_i\subseteq E_{i+1}\to\mu(\bigcup E_i)=\lim_{i\to\infty}\mu(E_i))$.

(4)(continuous from above) $(\forall\{E_i\}_1^\infty\subseteq\mathcal{M})(((\forall i)E_i\supseteq E_{i+1}\wedge(\exists N)\mu(E_N)<\infty)\to\mu(\bigcap E_i)=\lim_{i\to\infty}\mu(E_i))$.

(5)(Borel-Cantelli Lemma)$\{E_i\}_1^\infty\subseteq\mathcal{M}$, if $(\exists N)\sum_N^\infty\mu(E_i)<\infty$, then $\mu(\overline{\lim}\,E_n)=0$.

Proof. (5)Note $\overline{\lim}\,E_n=\bigcap_{n=1}^\infty\bigcup_{m=n}^\infty E_m\subseteq\bigcup_{m=n}^\infty E_m$, so $\mu(\overline{\lim}\,E_n)\leq\sum_{m=n}^\infty\mu(E_m)$ by countable subadditivity, then note $\lim_{n\to\infty}\sum_{m=n}^\infty\mu(E_m)=0$ by $(\exists N)\sum_N^\infty\mu(E_i)<\infty$. □

Remark 5.4.3

(1)The condition $(\exists N)\mu(E_N)<\infty$ of continuous from above is necessary, e.g. $(\mathbb{N},\mathcal{P}(\mathbb{N}),|\cdot|)$ and $E_i=\{n:n\geq i\}$.

(2)Signed measure is NOT monotonic and countably subadditive but continuous from below and above.

(3)By continuity from below and above, get $\mu(\underline{\lim}\,E_n)\leq\underline{\lim}\,\mu(E_n)$ and $\overline{\lim}\,\mu(E_n)\leq\mu(\overline{\lim}\,E_n)$ if $(\exists N)\mu(\bigcup_N^\infty E_i)<\infty$.

Lemma 5.4.2 (for extension)

(1)For $\mathcal{B}\subseteq\mathcal{P}(X)$ s.t. $\emptyset,X\in\mathcal{B}$, $\rho:\mathcal{B}\to[0,\infty]$ s.t. $\rho(\emptyset)=0$, then $\rho^*(A):=\inf\{\sum_1^\infty\rho(E_i):E_i\in\mathcal{B},A\subseteq\bigcup_1^\infty E_i\}$ is an outer measure.

(2)For outer measure μ^* on $\mathcal{P}(X)$, $\mathcal{M}$ of μ^*-measurable sets, written $\mathcal{C}$, is a σ-algebra, and $\mu^*|_{\mathcal{M}}$ is a complete measure.

(3)For measure μ on $\mathcal{M}$, let $\overline{\mathcal{M}}:=\{E\cup N:E,F\in\mathcal{M},\mu(F)=0,N\subseteq F\}$, then *completion* $\overline{\mu}:E\cup N\mapsto\mu(E)$ is unique extension of μ to a complete measure on $\overline{\mathcal{M}}$.

Proof. (1)Suffice to show countable subadditivity. $\forall\{A_i\}_1^\infty\subseteq\mathcal{P}(X),\epsilon>0$, note for each i, $\exists\{E_i^k\}_{k=1}^\infty\subseteq\mathcal{B}$ s.t. $A\subseteq\bigcup_{k=1}^\infty E_i^k$ and $\sum_{k=1}^\infty\rho(E_i^k)\leq\rho^*(A_i)+\epsilon 2^{-i}$. Hence $\rho^*(\bigcup A_i)\leq\sum_{i,k}\rho(E_i^k)\leq\sum\rho^*(A_i)+\epsilon$.

(2)Suffice to show countable union and additivity. For disjoint $\{A_i\}_1^\infty\subseteq\mathcal{M}$ and $E\subseteq X$, let $B_n=\cup_1^n A_i$ and $B=\cup_1^\infty A_i$. By induction, $\mu^*(E)=\mu^*(E\cap B_n)+\mu^*(E\cap B_n^c)$ and $\mu^*(E\cap B_n)=\sum_1^n\mu^*(E\cap A_i)$, then

$\mu^*(E) \geq \sum_1^n \mu^*(E\cap A_i) + \mu^*(E\cap B^c)$. Let $n\to\infty$, $\mu^*(E) \geq \sum_1^\infty \mu^*(E\cap A_i) + \mu^*(E\cap B^c) \geq \mu^*(\bigcup_1^\infty (E\cap A_i)) + \mu^*(E\cap B^c) = \mu^*(E\cap B) + \mu^*(E\cap B^c) \geq \mu^*(E)$. Hence $\mu^*(E) = \mu^*(E\cap B) + \mu^*(E\cap B^c)$, and $\sum_1^\infty \mu^*(E\cap A_i) = \mu^*(\bigcup_1^\infty (E\cap A_i))$ then let $E = B$. □

Theorem 5.4.3 (Carathéodory's Extension Theorem)

For premeasure μ_0 on semiring $\mathcal{S}\subseteq\mathcal{P}(X)$, $\exists\mu$ on $\Sigma(\mathcal{S})$ extends μ_0. Specially, if μ_0 is σ-finite, then the extension is unique.

Proof. μ_0 on $\mathcal{S} \Rightarrow \mu_0^*$ on $\mathcal{P}(X) \Rightarrow \mu := \mu_0^*|_{\mathcal{C}}$ on $\mathcal{C} \Rightarrow \mu|_{\Sigma(\mathcal{S})}$. To prove the uniqueness, it suffices to show "if ν is another measure on $\mathcal{C}$ that extends μ_0, then $(\forall E\in\mathcal{M})\nu(E)\leq\mu(E)$ with equality when $\mu(E)<\infty$".
By $E\in\mathcal{C}$, if $E\subseteq\bigcup_1^\infty A_i$ where $A_i\in\mathcal{S}$, then $\nu(E)\leq\sum_1^\infty \nu(A_i) = \sum_1^\infty \mu_0(A_i)$, get $\nu(E)\leq\mu_0^*(E)=\mu(E)$. If $\mu(E)<\infty$, then $\forall\epsilon>0$, there is $A:=\bigcup_1^\infty A_i\supseteq E$ where $A_i\in\mathcal{S}$, $\mu(A)<\mu(E)+\epsilon$, hence $\mu(E)\leq\mu(A) = \lim_{n\to\infty}\mu(\bigcup_1^n A_i) = \lim_{n\to\infty}\nu(\bigcup_1^n A_i) = \nu(A) = \nu(E)+\nu(A\setminus E)\leq\nu(E)+\mu(A\setminus E)\leq\rho(E)+\epsilon$. □

Definition 5.4.3 (product and section)

(1)For σ-algebra Σ_i on X_i where $i\in I$, $X=\prod_{i\in I}X_i$, coordinate map $\pi_i: X\to X_i$, the *product σ-algebra* on X is $\Sigma(\{\pi_i^{-1}(E_i):(\forall i)E_i\in\Sigma_i\})$, written $\bigotimes_{i\in I}\Sigma_i$.

(2)For measure space $(X_i,\mathcal{M}_i,\mu_i)$ where $1\leq i\leq n$,

(i)a *(measurable) rectangle* is $E_1\times...\times E_n$ where $E_i\in\mathcal{M}_i$ is *side* of E. Note the algebra it generates is $\mathcal{A}:=\{\bigsqcup_{j=1}^m E_1^j\times...\times E_n^j : E_i^j\in\mathcal{M}_i\}$ and the σ-algebra it generates is $\bigotimes\mathcal{M}_i$.

(ii)$\nu:\bigsqcup_{j=1}^m E_1^j\times...\times E_n^j\mapsto\sum_{j=1}^m\prod_{i=1}^n\mu_i(E_i^j)$ is a premeasure on $\mathcal{A}$, then induces a measure on $\bigotimes\mathcal{M}_i$, denoted by *product measure* $\mu_1\times...\times\mu_n$.

(3)For $E\subseteq\prod_{i\in I}X_i\times\prod_{j\in J}Y_j$, function f on $\prod_{i\in I}X_i\times\prod_{j\in J}Y_j$, then *$\overline{x}$-section (resp. $\overline{y}$-section)* of E is $E_{\overline{x}}=\{\overline{y}\in\prod_{j\in J}Y_j:(\overline{x},\overline{y})\in E\}$ (resp. $E^{\overline{y}}$), *$\overline{x}$-section (resp. $\overline{y}$-section)* of f is $f_{\overline{x}}(\overline{y})=f(\overline{x},\overline{y})$ (resp. $f^{\overline{y}}$).

Remark 5.4.4

(1)Note the product σ-algebra has associativity $\otimes_1^3\mathcal{M}_j = (\mathcal{M}_1\otimes\mathcal{M}_2)\otimes\mathcal{M}_3$, but the product measure NOT. However, note if μ_i is σ-finite for all i, then ν is σ-finite hence the extension is unique, so $\mu_1\times\mu_2\times\mu_3=(\mu_1\times\mu_2)\times\mu_3$.

Proposition 5.4.4

(1)For $|I|<\omega_1$ and σ-algebra Σ_i on X_i, $\bigotimes_{i\in I}\Sigma_i = \mathcal{M}(\{\prod_{i\in I}E_i : E_i\in\Sigma_i\})$.

(2)If $\Sigma_i=\Sigma(\mathcal{B}_i)$, then $\bigotimes_{i\in I}\Sigma_i=\Sigma(F_1:=\{\pi_i^{-1}(E_i):E_i\in\mathcal{B}_i\})$. Specially, if $|I|<\omega_1$ and $X_i\in\mathcal{B}_i$, then $\bigotimes_{i\in I}\Sigma_i=\Sigma(\{\prod_{i\in I}E_i:E_i\in\mathcal{B}_i\})$.

(3)For metric spaces $\{X_i\}_1^n$, $X=\prod_1^n X_i$ with product metric, then $\bigotimes_1^n\mathcal{B}(X_i)\subseteq\mathcal{B}(X)$. Specially, if X_i is separable for all i, then $\bigotimes_1^n\mathcal{B}(X_i)=\mathcal{B}(X)$.

(4)If $E\in\mathcal{M}\otimes\mathcal{N}$, then $E_x\in\mathcal{N}$ for all $x\in X$ and $E^y\in\mathcal{M}$ for all $y\in Y$. Moreover, if f is $\mathcal{M}\otimes\mathcal{N}$-measurable, then f_x is $\mathcal{N}$-measurable for all $x\in X$ and f^y is $\mathcal{M}$-measurable for all $y\in Y$.

Proof. (1)Note $\pi_i^{-1}(E_i) = \prod_{i\neq j\in I} X_j \times E_i$ and $\prod_{i\in I} E_i = \bigcap_{i\in I}\pi_i^{-1}(E_i)$.
(2)Note $\pi_i^{-1}(E_i) \in \bigotimes_{i\in I}\Sigma_i$ where $E_i \in \mathcal{B}_i$. On the other direction, fix i, $\{E \subseteq X_i : \pi_i^{-1}(E) \in \Sigma(F_1)\}$ is a σ-algebra on X_i that contains $\mathcal{B}_i$ and hence Σ_i, so $\pi_i^{-1}(E) \in \Sigma(F_1)$ where $E \in \Sigma_i$.
(3)Suppose C_i is a countable dense set of X_i, let $\mathcal{B}_i$ be the collection of balls in X_i with rational radius and center in C_i. Note any open set in X_i is a countable union of elements of $\mathcal{B}_i$, any open set in X is a countable union of balls with rational radius and center in $\prod C_i$. So by the above, $\bigotimes_1^n \mathcal{B}(X_i) = \Sigma(\{\prod_1^n E_i : E_i \in \mathcal{B}_i\}) = \mathcal{B}(X)$.
(4)Let $\mathcal{R} = \{E \in X\times Y : (\forall x)E_x \in \mathcal{N}, (\forall y)E^y \in \mathcal{M}\}$, note $\mathcal{R}$ contains all rectangles and is a σ-algebra since $(\bigcup_1^\infty E_j)_x = \bigcup_1^\infty (E_j)_x$ and $(E^c)_x = (E_x)^c$, $\mathcal{R}$. So $\mathcal{R} \supseteq \mathcal{M}\otimes\mathcal{N}$.
Then note $(f_x)^{-1}(B) = (f^{-1}(B))_x$. □

Example 5.4.1 (Lebesgue measure)

(1)*Counting measure* is μ on $(X, \mathcal{P}(X))$ s.t. $\mu : E \mapsto \begin{cases} |E| & |E| < \omega \\ \infty & o.w. \end{cases}$.

(2)*Dirac measure* at $x \in X$ is δ_x on $(X, \mathcal{M})$ s.t. $\delta_x : E \mapsto \begin{cases} 1 & x \in E \\ 0 & o.w. \end{cases}$.

(3)For topological space $(X, \mathcal{T})$, *Borel measure* is any measure on *Borel measurable space* $(X, \mathcal{B}(\mathcal{T}) = \mathcal{B}(X) := \Sigma(\mathcal{T}))$ where the elements of $\mathcal{B}(\mathcal{T})$ are *Borel (measurable) sets*.

(4)For $g : \mathbb{R} \to \mathbb{R}$ increasing and right continuous, then $\exists$ a unique measure on $\mathcal{B}(\mathbb{R})$ s.t. $(\forall a \leq b)\mu_g((a,b]) = g(b) - g(a)$, whose completion is written μ_g on $\overline{\mathcal{B}(\mathbb{R})}$ called *Lebesgue-Stieltjes measure* associated with g.

(5)*Lebesgue measure* m on $\mathcal{L}(\mathbb{R}^n) := \bigotimes_1^n \overline{\mathcal{B}(\mathbb{R})}$ is $\mu_g \times ... \times \mu_g$ where $g(x) = x$.

(6)*Radon measure* is a Borel measure s.t. X is Hausdorff, μ is inner regular and locally finite.

Remark 5.4.5

(1)For $\mathcal{B}(X)$, we should point out the topology on X first, but usually for $X = \mathbb{R}^n$, just pick the topology induced by 2-norm.

(2)For Lebesgue-Stieltjes measure,

(i)Well defined

(a)Premeasure $\mu_0(\bigsqcup_1^n (a_j, b_j]) = \sum_1^n [g(b_j) - g(a_j)]$

①Function: If $\bigsqcup (a_i, b_i] = (a, b]$, then after relabeling the index, get $a = a_1 < b_1 = a_2 < ... < b_n = b$, so $\sum_1^n (g(b_i) - g(a_i)) = g(b) - g(a)$. If h-intervals $\bigsqcup_1^n I_i = \bigsqcup_1^m J_j$, then $\sum_i \mu_0(I_i) = \sum_{i,j}\mu_0(I_i \cap J_j) = \sum_j \mu_0(J_j)$.

②Countable additivity: $\forall\{I_i\}_1^\infty$ h-intervals, wlog assume $\bigsqcup I_i = I = (a, b]$. By $\mu_0(I) = \mu_0(\bigsqcup_1^n I_i) + \mu_0(I \setminus \bigsqcup_1^n I_i) \geq \mu_0(\bigsqcup_1^n I_i) = \sum_1^n \mu_0(I_i)$, then let $n \to \infty$. If a, b is finite hence assume $I_j = (a_j, b_j]$, then $\forall\epsilon > 0\exists\delta, \delta_j > 0$ s.t. $g(a+\delta) - g(a) < \epsilon, g(b_j + \delta_j) - g(b_j) < \epsilon 2^{-j}$. Note $[a+\delta, b] \subseteq \cup_1^\infty (a_j, b_j + \delta_j)$, by Heine-Borel, $\mu_0(I) < \sum_1^\infty \mu_0(I_i) + 2\epsilon$, then let $\epsilon \to 0$. O.w. $a = -\infty$ or $b = \infty$, show $\sum \mu_0(I_i) = \infty$.

(b)note $\{(a, b] : -\infty \leq a \leq b \leq \infty\}$ (strictly speaking, $(a, \infty] \not\subseteq \mathbb{R}$, so it should be replaced by *h-interval* $(a,b],(a,\infty),\emptyset$ where $\infty \le a < b < \infty$) is an algebra, then use Carathéodory Extension and σ-finite premeasure.

(ii)Conversely, if μ on $\mathcal{B}(\mathbb{R})$ is finite on bounded Borel sets, then $F(x) = \begin{cases} \mu((0,x]) & x>0 \\ 0 & x=0 \\ -\mu((-x,0]) & x<0 \end{cases}$ is increasing, right continuous and $\mu = \mu_F$. Specially, if $\mu(\mathbb{R}) < \infty$, then *distribution function* $F(x) = \mu((-\infty,x])$, which differs from above by the constant $\mu((-\infty,0])$.

(iii)Moreover, Lebesgue-Stieltjes measure $\Leftrightarrow$ a complete and regular and σ-finite Borel measure.

(3)For Lebesgue measure, it is the unique Borel measure which is translation invariant, is finite on compact sets and attains 1 on the unit cube. (See bigrudin P50)

As to the condition "finite on compact sets", it could be proved that "For σ-compact LCH X, if μ is a positive Borel measure s.t. this condition holds, then μ is regular".(See bigrudin P48)

By regular, we can get that for $E \subseteq \mathbb{R}^n$, E is measurable iff $(\exists G \in \mathcal{B}(\mathbb{R}^n))(N \in \mathbb{R}^n \wedge m(N)=0)G = E \sqcup N$ iff $(\exists F \in \mathcal{B}(\mathbb{R}^n))(N \in \mathbb{R}^n \wedge m(N)=0)E = F \sqcup N$.

(4)A difference between $\mathcal{B}(\mathbb{R}^n)$ and $\mathcal{L}(\mathbb{R}^n)$, is $\mathcal{B}(\mathbb{R}^{n+m}) = \mathcal{B}(\mathbb{R}^n)\otimes\mathcal{B}(\mathbb{R}^m)$ while $\mathcal{L}(\mathbb{R}^{n+m}) \neq \mathcal{L}(\mathbb{R}^n)\otimes \mathcal{L}(\mathbb{R}^m)$. So $E,F \in \mathcal{B} \Leftrightarrow E\times F \in \mathcal{B}$ while $E,F \in \mathcal{L} \Rightarrow\not\Leftarrow E\times F \in \mathcal{L}$, e.g. E is Lebesgue nonmeasurable and F is a null set.

Proposition 5.4.5 (change-of-variable)

(1)For a complete, regular, σ-finite Borel measure μ on $\mathcal{M} = \Sigma(\mathcal{T})$, the followings are equivalent.

(i)$E \in \mathcal{M}$

(ii)$(\forall \epsilon > 0)(\exists E \subseteq G \in \mathcal{T})m(G \setminus E) < \epsilon$

(iii)$(\forall \epsilon > 0)(\exists E \supseteq F\ \textit{compact})m(E \setminus F) < \epsilon$

(iv)$(\exists E \subseteq G \in G_\delta)m(G \setminus E) = 0$

(v)$(\exists E \supseteq F \in F_\sigma)m(E \setminus F) = 0$.

(2)For Lebesgue measure m,

(i)(translation invariant) under translation $l_a : \mathbb{R}^n \to \mathbb{R}^n$ where $a \in \mathbb{R}^n$,

(a)If $E \in \mathcal{L}(\mathbb{R}^n)$, then $l_a(E) \in \mathcal{L}(\mathbb{R}^n)$ and $m(l_a(E)) = m(E)$.

(b)If $f : \mathbb{R}^n \supseteq E \to \mathbb{R}$ is measurable, then $f \circ l_a : l_a^{-1}(E) \to \mathbb{R}$ is measurable. Moreover, if $f \ge 0$ or $f \in L^1(m)$, then $\int_E f dm = \int_{l_a^{-1}(E)} (f \circ l_a) dm$.

(ii)under linear transformation $T \in GL_n(\mathbb{R})$,

(a)If $E \in \mathcal{L}(\mathbb{R}^n)$, then $T(E) \in \mathcal{L}(\mathbb{R}^n)$ and $m(T(E)) = |\det T| m(E)$.

(b)If $f : \mathbb{R}^n \supseteq E \to \mathbb{R}$ is measurable, then $f \circ T : T^{-1}(E) \to \mathbb{R}$ is measurable. Moreover, if $f \ge 0$ or $f \in L^1(m)$, then $\int_E f dm = |\det T| \int_{T^{-1}(E)} f \circ T dm$.

(iii)(change-of-variables) under C^1 and injective $\varphi : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}^n$ where Ω is open,

(a)If $\Omega \supseteq E \in \mathcal{L}(\mathbb{R}^n)$, then $\varphi(E) \in \mathcal{L}(\mathbb{R}^n)$ and $m(\varphi(E)) = \int_E |\det \varphi_x| dm(x)$.

(b)If $\Omega \supseteq E \in \mathcal{L}(\mathbb{R}^n)$ and $f : \varphi(E) \to \mathbb{R}$ is measurable, then $f \circ \varphi : E \to \mathbb{R}$ is measurable. Moreover, if $f \ge 0$ or $f \in L^1(m)$, then $\int_{\varphi(E)} f dm = \int_\Omega f \circ \varphi(x) |\det \varphi_x| dm(x)$.

Proof. (2)(i)Suffice to show (a) for Borel sets, since $\forall m(N)=0 \exists N \subseteq G \in \mathcal{B}(\mathbb{R}^n)$ s.t. $m(N)=m(G)$. Then measurability holds for continuity, and $m(l_a(E))=m(E)$ follows for cubes for open sets by open-set-decomposition, for Borel sets by regularity.
$\forall E \in \mathcal{B}(\mathbb{R})$, $(\exists F \in \mathcal{B}(\mathbb{R}^n))(\exists N \in \mathbb{R}^n \wedge m(N)=0) f^{-1}(E)=F \sqcup N$, then $(f \circ l_a)^{-1}(E)=l_a^{-1}(F) \sqcup l_a^{-1}(N) \in \mathcal{L}^n$, i.e. $f \circ l_a$ is measurable.
The latter of (b) holds for characteristic function by (a), for simple function by linearity, for nonnegative measurable functions by MCT, for measurable functions by taking positive and negative parts.
(ii)Similarly, suffice to show $m(T(E))=|\det T| m(E)$ for cubes. Note if $m(T(E))=|\det T| m(E), m(S(E))=|\det S| m(E)$ hold for linear transformation T, S, then $m((T \circ S)(E))=|\det T||\det S| m(E)=|\det(T \circ S)| m(E)$. Hence suffice to show for three elementary linear transformation, for S_{ij} interchange the order of integration in x_i and x_j by Fubini, for $M_i(c), A_{ij}(c)$ by linearity and (i).
(iii)Prove for the situation of "φ_x is invertible for all $x \in \Omega$", similarly, suffice to show $m(\varphi(E))=\int_E |\det \varphi_x| dm(x)$ for cube E of side length $2a$ centered at 0. Denote $||\cdot||_\infty$ by $||\cdot||$ below.
By E compact, know,
$||x||=\max_{1 \leq j \leq n} |x_j|$, $||T||=\max_{1 \leq i \leq n} \Sigma_{j=1}^n |T_{i,j}|$, we then have $||Tx|| \leq ||T||\ ||x||$ and $\{x : ||x-a|| \leq h\}$ is the cube of .
Let $Q=\{x : ||x-a|| \leq h\}$ be a cube in Ω, by the mean value theorem, $g_j(x)-g_j(a)=\Sigma_j (x_j-a_j)(\partial g / \partial x_j)(y)$ for some y on the line segment joining x and a s.t. $||G(x)-G(a)|| \leq h \sup_{y \in Q} ||D_y G||$. In other words, $G(Q)$ is contained in a cube of side length $\sup_{y \in Q} ||D_y G||$ times that of Q. So that by Thm2.44, $m(G(Q)) \leq (\sup_{y \in Q} ||D_y G||)^n m(Q)$. If $T \in GL(n, \mathbb{R})$, we can apply this formula with G replaced by $T^{-1} \circ G$ together with Thm2.44 to get $m(G(Q))=|\det T| m(T^{-1}(G(Q))) \leq |\det T| (\sup_{y \in Q} ||T^{-1} D_y G||)^n m(Q)$
Since $D_y G$ is continuous in y, for all $\epsilon>0$, $(\exists \delta>0) ||(D_z G)^{-1} D_y G||^n \leq 1+\epsilon$ if $y, z \in Q$ and $||y-z|| \leq \delta$. Subdivide Q into subcubes $Q_1, \ldots, Q_N$ whose interiors are disjoint, whose side lengths are at most δ, whose centers are $x_1, \ldots, x_N$. Apply above replaced by Q_j and with $T=D_{x_j} G$, obtain $m(G(Q)) \leq \Sigma_1^N m(G(Q_j)) \leq \Sigma_1^N |\det D_{x_j} G| (\sup_{y \in Q_j} ||(D_{x_j} G)^{-1} D_y G||^n) m(Q_j) \leq (1+\epsilon) \Sigma_1^N |\det D_{x_j} G| m(Q_j)$ This last sum is the integral of $\Sigma_1^N |\det D_{x_j} G| \chi_{Q_j}(x)$ which tends uniformly on Q to $|\det D_x G|$ as $\delta \to 0$ since $D_x G$ is continuous. Thus, letting $\delta \to 0, \epsilon \to 0$, find $m(G(Q)) \leq \int_Q |\det D_x G| dx$ □

Theorem 5.4.6 (Littlewood's 1st Principal)
(Every measurable set of finite measure is nearly a finite union of intervals)

For Lebesgue-measurable $E \subseteq \mathbb{R}^n$ with $m(E)<\infty$, $\epsilon>0$, there exists $F=\bigsqcup_1^n I_i$ where I_i is cube s.t. $m(E \triangle F)<\epsilon$.

Definition 5.4.4 (measurable function)

(a)For measurable space $(X, \mathcal{M})$ and $(Y, \mathcal{N})$, $f : X \to Y$ is *$(\mathcal{M}, \mathcal{N}$-)measurable* iff $(\forall E \in \mathcal{N}) f^{-1}(E) \in \mathcal{M}$. Specially,

(1)$f : X \to \mathbb{R}^n$ is *($\mathcal{M}$-)measurable* iff f is $\mathcal{M}, \mathcal{B}(\mathbb{R}^n)$-measurable.

(2)$f : \mathbb{R}^m \to \mathbb{R}^n$ is *Borel measurable* iff f is $\mathcal{B}(\mathbb{R}^m)$-measurable.

(3)$f:\mathbb{R}^m\to\mathbb{R}^n$ is *(Lebesgue) measurable* iff f is $\mathcal{L}(\mathbb{R}^m)$-measurable.

By considering $\mathbb{C}$ as $(\mathbb{R}^2,\mathcal{B}(\mathbb{R}^2))$, generalize the measurable into $\mathbb{C}$; note $\Sigma(\{(a,\infty]:a\in\mathbb{R}\}) = \Sigma(\mathcal{T}(\{[-\infty,a),(a,\infty]:a\in\mathbb{R}\}))$, by considering $\overline{\mathbb{R}}$ as $(\overline{\mathbb{R}},\Sigma(\{(a,\infty]:a\in\mathbb{R}\}))$, generalize the measurable into $\overline{\mathbb{R}}$.

(b)*positive part (resp. negative part)* $f^+:=\max(f,0)$, $f^-:=\max(-f,0)$

$|f|=f^++f^-$ for real-valued, $|f|=\sqrt{(\Re f)^2+(\Im f)^2}$ (note $|f|\le|\Re f|+|\Im f|\le 2|f|$) (c)*Simple function* on $E\in\mathcal{M}$ is $\sum_1^n a_i\chi_{E_i}$ where $a_i\in\mathbb{R}^+\cup\{0\}$, $E_i\in\mathcal{M}$ and $\bigsqcup_1^n E_i=E$.

Remark 5.4.6

WARNING: complex measurable function $\Rightarrow$ real measurable function $\Rightarrow$ nonnegative measurable function $\Rightarrow$ simple function $\Rightarrow$ measurable characteristic function $\Rightarrow$ measurable set $\Rightarrow$ Lebesgue $\Rightarrow$ Borel+Null set(controlled by Borel) $\Rightarrow$ Borel $\Rightarrow$ open $\Rightarrow$ interval with rational center and radius. WARNING: we usually define or prove for real-valued, then most of them could be generalized to extended-real-valued, complex-valued and vector-valued.

Proposition 5.4.7

(1)If $\mathcal{N}=\Sigma(\mathcal{B})$, then $f:X\to Y$ is $(\mathcal{M},\mathcal{N})$-measurable iff $(\forall E\in\mathcal{B})f^{-1}(E)\in\mathcal{M}$. Moreover,

(i)If $f:X\to Y$ is continuous wrt. $\mathcal{T}_X,\mathcal{T}_Y$, then f is $\mathcal{B}(X),\mathcal{B}(Y)$-measurable.

(ii)$f:X\to\mathbb{R}$ is $\mathcal{M}$-measurable iff $(\forall a\in\mathbb{R})f^{-1}((a,\infty))\in\mathcal{M}$.

(2)For measurable spaces $(X,\mathcal{M})$, $\{(Y_i,\mathcal{N}_i)\}_{i\in I}$, $(Y=\prod Y_i,\mathcal{N}=\bigotimes\mathcal{N}_i)$, coordinate maps $\pi_i:Y\to Y_i$, then $f:X\to Y$ is $\mathcal{M},\mathcal{N}$-measurable iff $f_i:=\pi_i\circ f$ is $\mathcal{M},\mathcal{N}_i$-measurable for all $i\in I$.

(i)$f:X\to\mathbb{C}$ is $\mathcal{M}$-measurable iff $\Re f,\Im f$ is $\mathcal{M}$-measurable.

(3)For $\overline{\mathbb{R}}$-valued measurable $\{f_i\}_1^\infty$ on $(X,\mathcal{M})$, $\sup f_i(x),\inf f_i(x),\overline{\lim} f_j(x),\underline{\lim} f_j(x)$ is measurable. (Pf. $(\sup f_i)^{-1}((a,\infty])=\bigcup f_i^{-1}((a,\infty]),(\inf f_i)^{-1}((a,\infty])=\bigcap f_i^{-1}((a,\infty])$) Specially, if $f(x)=\lim f(x)$ exists for all $x\in X$, then f measurable.

(i)For measurable f,g, $\max(f,g),\min(f,g),f^+,f^-$ is measurable.

(4)For $E\subseteq X$, χ_E is $\mathcal{M}$-measurable iff $E\in\mathcal{M}$.

Exercise 5.4.1

(1)For $\mathcal{M}$-measurable $f,g:X\to\mathbb{R}$, show that $f+g,fg$ is $\mathcal{M}$-measurable.

(2)For measurable $E_1\subseteq\mathbb{R}^{n_1},E_2\subseteq\mathbb{R}^{n_2}$, *Carathéodory function* $f:E_1\times E_2\to\mathbb{R}^{n_3}$ s.t. f^y is measurable for all $y\in E_2$ and f_x is continuous for all $x\in E_1$, if $g:E_1\to E_2$ is measurable, then $x\mapsto f(x,g(x))$ is measurable.

Proof. (1)$\{f+g>a\}=\bigcup_{r\in\mathbb{Q}}(\{f>r\}\cap\{g>a-r\})$. For fg is measurable, give two methods:
(i)$\{fg>c\}=\bigcup_{a,b\in\mathbb{Q}^+,ab\ge c}(\{f>a\}\cap\{g>b\})\bigcup_{a\in\mathbb{Q}^-,b\in\mathbb{Q}^+,ab\ge c}(\{f<a\}\cap\{g>b\})\bigcup_{a\in\mathbb{Q}^+,b\in\mathbb{Q}^-,ab\ge c}(\{f>a\}\cap\{g<b\})\bigcup_{a,b\in\mathbb{Q}^-,ab\ge c}(\{f<a\}\cap\{g<b\})$.
(ii)$F:X\to\mathbb{R}\times\mathbb{R},x\mapsto(f(x),g(x))$ is $\mathcal{M},\mathcal{B}(\mathbb{R}^2)=\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R})$-measurable by 5.4.7(2), $\varphi:\mathbb{R}\times\mathbb{R}\to\mathbb{R},(x,y)\mapsto$

xy is $\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R})$-measurable by 5.4.7(1)(i).

(2)Pick simple $\varphi_i \to \varphi$ pointwise, assume $\varphi_i = \sum_1^n a_j \chi_{A_j}$, then $x \mapsto f(x, \varphi_i(x))$ is measurable since $(f(\cdot, \varphi_i(\cdot)))^{-1}(E) = \bigcup_1^n ((f^{a_j})^{-1}(E) \times a_j)$. So $f(x, \varphi(x)) = \lim f(x, \varphi_i(x))$ is measurable. □

Theorem 5.4.8 (for completion)

(1)The following holds iff μ complete:

(i)If f is measurable and $f = g$ μ-a.e., then g is measurable.

(ii)If f_n measurable for $n \in \mathbb{N}$ and $f_n \to f$ μ-a.e., then f measurable.

(2)For measure space $(X, \mathcal{M}, \mu)$ and completion $(X, \overline{\mathcal{M}}, \overline{\mu})$, if f is $\overline{\mathcal{M}}$-measurable, then $\exists g$ is $\mathcal{M}$-measurable s.t. $f = g$ $\overline{\mu}$-a.e.

Proof. (2)Pick simple $\varphi_n \to f$ pointwise, let ψ_n be $\mathcal{M}$-measurable simple function with $\psi_n = \varphi_n$ except on a set $E_n \in \overline{\mathcal{M}}$ with $\bar{\mu}(E_n) = 0$. Choose $N \in \mathcal{M}$ s.t. $\mu(N) = 0$ and $N \supseteq \cup_1^\infty E_n$, let $g = \lim \chi_{X \setminus N} \psi_n$, then g is $\mathcal{M}$-measurable and $g = f$ on N^c. □

Example 5.4.2 (Cantor set)

Cantor set C

(1)Construction:

(i)$C = \{\sum_1^\infty a_j 3^{-j} : a_j \in \{0, 2\}\}$ (base 3 expression, easy to see $|C| = \aleph_1$ and every pt is accumulation pt)

(ii)$C = [0,1] \setminus \bigcup_{n=0}^\infty \bigcup_{k=0}^{3^n - 1} (\frac{3k+1}{3^{n+1}}, \frac{3k+2}{3^{n+1}}) = [0,1] \setminus \bigsqcup_{n=0}^\infty \bigsqcup_{k=0}^{2^n - 1} I_k^n$ (easy to see closed)

(iii)remove (easy to see no interior pt and $m(C) = 0$)

(2)Property:

(i)$|C| = \aleph_1, m(C) = 0$

Cor. Since Lebesgue measure is complete, $|\mathcal{L}| \geq |\mathcal{P}(C)| = \aleph_2 > \aleph = |\mathcal{B}_\mathbb{R}|$, so there exists a Lebesgue measurable but not Borel measurable set.

(ii)Stone(compact(bounded closed) totally disconnected Hausdorff), every pt is boundary pt(no interior pt(nowhere dense(meagre))) and accumulation pt(isolated pt(perfect))

Pf. For totally disconnected, $\forall x < y \in C$ note $(\exists x < z < y) z \notin C$, so $x \in [0, z) \cap C, y \in (z, 1] \cap C$ then totally separated hence totally disconnected.

(iii)Generally speaking, a perfect totally disconnected subset on $\mathbb{R}$ is homeomorphic to C.

(3)Generalization:

(i)*Generalized Cantor set* $K = \bigcap_0^\infty K_j$ where $K_0 = [0,1]$, K_j is obtained by removing the open middle-α_j $(0 < \alpha_j < 1)$ from each of the intervals that make up K_{j-1}.

(a)$m(K) = \lim_{n\to\infty} \prod_1^n (1 - \alpha_j)$

(b)Stone(compact(bounded closed) totally disconnected Hausdorff), every pt is boundary pt(no interior pt(nowhere dense(meagre))) and accumulation pt(isolated pt(perfect))

(c)Usually replace middle-α_j by $\alpha 3^{-j}$ $(0 < \alpha \leq 1)$, then $m(K) = 1 - \alpha$.

(ii)*Cantor dust* $\prod_1^n C$

(a)$(\forall x \in [0,2])(\exists a, b \in C)x = a + b$

(b)Compared to Steinhaus, $[-1,1] \subseteq C - C$ while $m(C) = 0$

(iii)*Cantor function* $c : [0,1] \to [0,1], x \mapsto \begin{cases} \sum_1^\infty \frac{a_j}{2^j} & x = \sum_1^\infty \frac{2a_j}{3^j} \in C \\ \sup_{y<x, y\in C} c(y) & o.w. \end{cases}$ is notorious counterexample for continuous but not absolutely continuous function.

Pf. increasing + surjection → continuous. By $m(C) = 0$, $\forall \delta > 0 \exists M \exists x_k < y_k \in [0,1]$ where $1 \leq k \leq M$ s.t. $\sum_1^M (y_k - x_k) < \delta$ and $\sum_1^M (c(y_k) - c(x_k)) = 1$.

Lemma 5.4.9 (cube approximation)

For measurable $E \subseteq \mathbb{R}^n$ with $m(E) > 0$, $0 < \lambda < 1$, there exists a cube I s.t. $\lambda m(I) < m(I \cap E)$. Moreover, there exists $E \subseteq [0,1]$ with $m(E) > 0$ s.t. $\forall$*interval* $I \subseteq [0,1]$, $0 < m(E \cap I) < m(I)$.

Proof. Regular measure + open decomposition + pigeonhole principle.

Let M be the set of closed intervals in $[0,1]$ with rational endpoints, note M is countable so denoted by $\{I_i\}_1^\infty$.

Let X is CTDP mean X is a compact totally disconnected subset of $[0,1]$ that has positive measure.

Claim $\forall I \in M \exists CTDP\ A \subseteq I$ s.t. $(\exists J \in M) J \subseteq I \setminus A$, it holds with generalized Cantor set. Construct $\{A_i\}_1^\infty, \{B_i\}_1^\infty$ as follows:

(i)By Claim, $\exists CTDP\ A_1 \subseteq I_1, M \ni J \subseteq I \setminus A_1$, then $\exists CTDP\ B_1 \subseteq J$. Hence get $A_1, B_1 \subseteq I_1$ are disjoint and CTDP.

(ii)Once $\{A_i\}_1^{n-1}, \{B_i\}_1^{n-1}$ are chosen, note $\bigcup_1^{n-1}(A_i \cup B_i)$ is CTDP, so $\exists M \ni J \subseteq I_n \setminus C_n$. Similarly, $\exists A_n, B_n \subseteq J \subseteq I_n$ are disjoint and CTDP.

Let $E = \bigcup_1^\infty A_i$, then $\forall$*interval* $I \subseteq [0,1]$, $\exists M \ni I_m \subseteq I$. Hence $A_m, B_m \subseteq I$, $0 < m(A_m) \leq m(E \cap I) < m(E \cap I) + m(B_m) \leq m(I)$. ☐

Theorem 5.4.10 (Lebesgue's Density Theorem)

(For Lebesgue measurable $A \subseteq \mathbb{R}^n$, the "boundary" is negligible. The "density" of A is 0 or 1 at a.e. pt in $\mathbb{R}^n$)

$\forall E \subseteq \mathbb{R}$, $m^*(E \triangle \phi(E)) = 0$ where $\phi(E) := \{x \in \mathbb{R} : \lim_{h\to 0} \frac{m^*(E\cap[x-h,x+h])}{2h} = 1\}$.

Proof. Suffice to show $m^*(E \setminus \phi(E)) = 0$ for $\phi(E) \setminus E = E^c \setminus \phi(E)^c \subseteq E^c \setminus \phi(E^c)$. WLOG assume E is bounded. Note $E \setminus \phi(E) = \bigcup_1^\infty A_n$ where $A_n = \{x \in E : \underline{\lim}_{h\to 0} \frac{m^*(E\cap[x-h,x+h])}{2h} < 1 - \frac{1}{n}\}$, $\forall A := A_n \forall \epsilon > 0 \exists A \subseteq G$ open s.t. $m(G) < m^*(A) + \epsilon$.

$\forall a \in A$, pick an open interval I with rational endpoints s.t. $a \in I \subseteq G$ and $m^*(E \cap I) < (1 - \frac{1}{n})m(I)$, which gives a countable cover of A. Note $m^*(A) \leq m(\bigcup_1^\infty I_n)$, so $(\exists N \in \mathbb{N}) m^*(A) - \epsilon < m(\bigcup_{n\leq N} I_n)$. By Vitali Covering Lemma, there exists disjoint subcollection $\{I_{n_j}\}_{j=1}^m$ s.t. $\bigcup_{n\leq N} I_n \subseteq \bigcup_{j\leq m} 3I_{n_j}$, let $X = \bigsqcup_{j\leq m} I_{n_j}$ then $m^*(A) - \epsilon < m(\bigcup_{j\leq m} 3I_{n_j}) \leq 3\sum_{j\leq m} m(I_{n_j}) = 3m(X)$.

Note $m(X) - \epsilon = m(G) - m(G \setminus X) - \epsilon < m^*(A) - m^*(A \setminus X) \leq m^*(A \cap X) \leq \sum_{j\leq m} m^*(A \cap I_{n_j}) \leq \sum_{j\leq m} m^*(E \cap I_{n_j}) \leq \sum_{j\leq m}(1 - \frac{1}{n})m(I_{n_j}) = (1 - \frac{1}{n})m(X)$, so $m(X) < n\epsilon$ then $m^*(A) < \epsilon(1 + 3n)$, let $\epsilon \to 0$

get $m^*(A) = 0$. $\square$

Theorem 5.4.11 (Steinhaus)

For measurable $E \subseteq \mathbb{R}^n$ with $m(E) > 0$, $(\exists \delta > 0) B_\delta(0) \subseteq E - E$

Proof. Give two method:

(A)By Lebesgue's Density Theorem, $\exists e \in E \cap \phi(E)$, pick $\epsilon = \frac{1}{4} > 0$ then $\exists \delta > 0$ s.t. $\frac{m([e-\delta, e+\delta] \setminus E)}{2\delta} < \epsilon$. $\forall x \in [-\frac{\delta}{2}, \frac{\delta}{2}]$, note $m(\{y \in [e - \frac{\delta}{2}, e + \frac{\delta}{2}] : y \notin E\}) < 2\delta\epsilon$ and $m(\{y \in [e - \frac{\delta}{2}, e + \frac{\delta}{2}] : x + y \notin E\}) < 2\delta\epsilon$, hence $m(\{y \in [e - \frac{\delta}{2}, e + \frac{\delta}{2}] : y \in E, x + y \in E\}) \geq \delta - 4\delta\epsilon > 0$. So $[-\frac{\delta}{2}, \frac{\delta}{2}] \subseteq E - E$.

(B)By cube approximation, $\exists I$ with $\frac{3}{4} m(I) < m(I \cap E)$. Suppose not, then $\exists |x| < \frac{m(I)}{2}$ s.t. $E \cap (E + x) = \emptyset$, hence $(I \cap E) \cap ((I \cap E) + x) = \emptyset$. So $\frac{3}{2} m(I) < 2m(I \cap E) = m((I \cap E) \cap ((I \cap E) + x)) \leq m(I \cap (I + x)) < \frac{3}{2} m(I)$, contradiction. $\square$

Theorem 5.4.12 (Vitali)

For $E \subseteq \mathbb{R}$ with $m^*(E) > 0$, there exists Lebesgue nonmeasurable $F \subseteq E$.

Proof. WLOG assume $E \subseteq [0, 1]$. By Choice, let V be the set of representatives of elements of $E/\mathbb{Q}$. Suppose V is Lebesgue measurable, then $\infty \times m(V) = m(\bigsqcup_{r \in \mathbb{Q} \cap [-1,1]} (V + r)) \leq m([-1, 2]) \leq 3 \Rightarrow m(V) = 0$, hence $m^*(E) \leq m(\bigsqcup_{r \in \mathbb{Q} \cap [-1,1]} (V + r))$, contradiction. $\square$

5.5 Integration

Definition 5.5.1 (modes of convergence)

For measurable $E, \{f_n\}_1^\infty, f$, the mode of $f_n \to f$ on E is

(1)*a.e. (pointwise)* iff $\mu(\{x \in E : (\exists \epsilon > 0)(\forall N \in \mathbb{N})(\exists n \geq N) |f_n(x) - f(x)| \geq \epsilon\}) = 0$.

(2)*uniformly* iff $(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall n \geq N)(\forall x \in E) |f_n(x) - f(x)| < \epsilon$.

(a)*compact uniformly* iff $(\forall E \supseteq K \text{ compact}) f_n|_K \to f|_K$ uniformly.

(b)*almost uniformly* iff $(\forall \epsilon > 0)(\exists E \supseteq E_\epsilon \in \mathcal{M} \wedge \mu(E) < \epsilon) f_n|_{E \setminus E_\epsilon} \to f|_{E \setminus E_\epsilon}$ uniformly.

(3)*in L^p* iff $(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall n \geq N)(\int_E |f_n - f|^p d\mu)^{\frac{1}{p}} < \epsilon$.

(4)*in measure* iff $(\forall \delta > 0)(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall n \geq N) \mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) < \epsilon$.

Remark 5.5.1

(1)WARNING: There's an annoying point that in all this definition above, we need f_n, f to be finite a.e. (probably only a.e. pointwise not need, so if we write $f_n \to f$ (except a.e.), it imply the condition finite a.e.), then these definitions are on $X \setminus E \bigcup_1^\infty E_n$ where f_n is finite on $X \setminus E_n$ and f is finite on $X \setminus E$.

(2)Note "a.e.", "in L^p" and "in measure" are relevant to the measure μ we pick, usually it is Lebesgue measure m.

Note that convergence of almost uniform NOT means that the sequence converges uniformly a.e. as might be inferred from the name.

(3)It will be useful to keep in mind the following examples on $\mathbb{R}$:

(i)$f_n = n^{-1}\chi_{[0,n]}$ (ii)$f_n = \chi_{[n,n+1]}$ (iii)$f_n = n\chi_{[0,1/n]}$ (iv)$f_n = \chi_{[\frac{j}{2^k},\frac{j+1}{2^k}]}$ where $n = 2^k + j$ with $0 \le j < 2^k$

$f \to 0$ a.e. (i)(ii)(iii); $f \to 0$ uniformly (i); $f \to 0$ in L^1 (iv); $f \to 0$ in measure (i)(iii)(iv). Note (iv) NOT converges for any $x \in [0,1]$.

(4)To interchange $\forall$ and a.e., it need the underset of $\forall$ is countable.

Theorem 5.5.1 (inequality)

(1)(Linearity) For $f, g \in L^1$, $c \in \mathbb{R}$, then $\int(f + cg) = \int f + c\int g$.

(2)(Triangle) For $f \in L^1$, then $|\int f| \le \int |f|$, equality holds when exists α s.t. $\alpha f = |f|$ a.e.

(3)(Monotone) For $f, g \in L^1, f \le g$, then $\int f \le \int g$.

(4)(Chebyshev) For $A > 0$, $\mu(\{|f_n - f| \ge A\}) \le \frac{1}{A}\int |f_n - f| d\mu$.

(5)(Hölder) For conjugate exponents p, q, $f \in L^p$, $g \in L^q$, then $fg \in L^r$ and $||fg||_1 \le ||f||_p||g||_q$, equality holds when exists $\alpha, \beta \in [0, \infty)$ not all zero s.t. $\alpha|f|^p = \beta|g|^q$ a.e.

(weighted) Specially, for conjugate exponents p, q, $f \in L^p$, $g \in L^q$, nonnegative $w \in L^1$, then $|\int fgw| \le (\int |f|^p w)^{\frac{1}{p}}(\int |g|^q w)^{\frac{1}{q}}$.

Moreover, for conjugate exponents p, q, measurable f, then $||f||_p = \sup_{g \in L^q, ||g||_q = 1}\int fg$.

(6)(Minkowski) For $p \in [0, \infty]$, $\{f_n\}_1^m \subseteq L^p$, $||\sum_1^m f_n||_p \le \sum_1^m ||f_n||_p$, equality holds when exists $\alpha_n \in [0, \infty)$ not all zero s.t. $\alpha_n f_n^p$ equals a.e.

(7)(Jensen) For convex Ω, convex $\varphi : \Omega \to \mathbb{R}$, $m(E) < \infty$, $L^1 \ni f : E \to \Omega$, then $\varphi(\frac{1}{m(E)}\int_E f) \le \frac{1}{m(E)}\int_E \varphi \circ f$.

Proof. (5)Let $A = ||f||_p, B = ||g||_q$, assume $AB \neq 0$ o.w. $f = 0$ a.e. or $g = 0$ a.e. By Young, $\int \frac{|fg|}{AB} \le \int(\frac{1}{p}\frac{|f|^p}{A^p} + \frac{1}{q}\frac{|g|^q}{B^q}) = 1$.

(6)Let $F = |\sum_1^m f_n|^{p-1}$ and q be the conjugate exponent of p, then $\int |\sum_1^m f_n|^p \le \int F\sum_1^m |f_k| \le \sum_1^m(\int F^q)^{\frac{1}{q}}(\int |f_k|^p)^{\frac{1}{p}} = (\int |\sum_1^m f_n|^p)^{1-\frac{1}{p}}\sum_1^m(\int |f_n|^p)^{\frac{1}{p}}$.

(7)Note $y_0 := \frac{1}{m(E)}\int_E f \in \Omega$, o.w. by convex Ω, $\exists \eta \in \Omega$ s.t. $(\forall y \in \Omega) < y - y_0, \eta >< 0$ since $\{< y - y_0, \eta >: y \in \Omega\}$ is open, then $< y_0, \eta >= \frac{1}{m(E)}\int_E < f(x), \eta > dm(x) << y_0, \eta >$.

By convex φ, $\exists \gamma \in \Omega$ s.t. $(\forall y \in \Omega)\varphi(y) \ge \varphi(y_0)+ < \gamma, y - y_0 >$, then $\frac{1}{m(E)}\int_E \varphi \circ f \ge \frac{1}{m(E)}\int_E(\varphi(y_0)+ < \gamma, f(x) - y_0 >)dm(x) \ge \varphi(y_0)$. □

Remark 5.5.2

$f \in L^1$, then $||f||_1$ is absolutely continuous, i.e. $(\forall \epsilon > 0)(\exists \delta > 0)(\forall E \supseteq \omega \in \mathcal{M} \wedge \mu(\omega) < \delta)\int_\omega |f|d\mu < \epsilon$. Specially $\lim_{M\to\infty}\int_{|f|\ge M}|f|d\mu = 0$.

Pf. Exists increasing simple function $\varphi_n \to |f|$ pointwise, then $\overline{\lim}_{E\supseteq\mu\in\mathcal{M},\mu(\omega)\to 0^+}\int_\omega |f| \le \overline{\lim}_{E\supseteq\mu\in\mathcal{M},\mu(\omega)\to 0^+}\int_\omega(|f| - \varphi_n) + \overline{\lim}_{E\supseteq\mu\in\mathcal{M},\mu(\omega)\to 0^+}\int_\omega \varphi_n \le \int_E(|f| - \varphi_n) = \int_E |f| - \int_E \varphi_n$, let $n \to \infty$ and by MCT.

Theorem 5.5.2 (Convergence)

For measure space $(X, \mathcal{M}, \mu)$

(1)(Monotone Convergence Theorem (MCT for short)) For nonnegative measurable $\{f_n\}_1^\infty$ and $f_n \le f_{n+1}, \forall n \ge 1$, a.e. $x \in X$, then $\int \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int f_n$.

(for series) For nonnegative measurable $\{f_n\}_1^\infty$, then $\int \sum_{n=1}^\infty f_n = \sum_{n=1}^\infty \int f_n$.

(2)(Fatou) For nonnegative measurable $\{f_n\}_1^\infty$, $\int(\underline{\lim}_{n\to\infty} f_n) \le \underline{\lim}_{n\to\infty} \int f_n$.

(3)(Dominated convergence Theorem (DCT for short)) For measurable $\{f_n\}_1^\infty$, $g \in L^1$ s.t. $|f_n| \le g, \forall n \ge 1$, a.e. $x \in X$, if $f_n \to f$ a.e., then $f \in L^1$ and $\int f = \lim_{n\to\infty} \int f_n$.

(for series) For $\{f_j\} \subseteq L^1$ s.t. $\sum_1^\infty \int |f_j| < \infty$, then $\sum_1^\infty f_j$ is absolutely convergent a.e. and $\int \sum_1^\infty f_j = \sum_1^\infty \int f_j$.

(corollary of DCT and Riesz) For measurable $\{f_n\}_1^\infty$, $g \in L^1$ s.t. $|f_n| \le g, \forall n \ge 1$, a.e. $x \in X$, if $f_n \to f$ in measure, then $f \in L^1$ and $\int f = \lim_{n\to\infty} \int f_n$.

(4)(Egorov; Littlewood's 3rd Principal)

(Every convergent sequence of functions is nearly uniformly convergent)

$f_n \to f$ a.e. on E + $\mu(E) < \infty$ + $f < \infty$ a.e. $\Rightarrow f_n \to f$ almost uniformly on E.

(5)$f_n \to f$ uniformly $\Rightarrow f_n \to f$ almost uniformly $\Rightarrow f_n \to f$ a.e. + $f_n \to f$ in measure.

(6)(Riesz) $f_n \to f$ in measure $\Rightarrow (\exists \{f_{n_k}\} \subseteq \{f_n\}) f_{n_k} \to f, k \to \infty$ a.e.

(corollary of Egorov and Riesz) For $\mu(E) < \infty$, measurable $f_n, f < \infty$ a.e., then $f_n \to f$ in measure iff $(\forall \{f_{n,k}\} \subseteq \{f_n\})(\exists \{f_{n,k,i} \subseteq \{f_{n,k}\}\}) f_{n,k,i} \to f, i \to \infty$ a.e.

(7)$f_n \to f$ in $L^1 \Rightarrow f_n \to f$ in measure.

(8)If sequence $\{f_n\}$ of measurable functions is *Cauchy in measure*, i.e. $(\forall \epsilon > 0) \lim_{m\to\infty, n\to\infty} \mu(\{x : |f_n(x) - f_m(x)| \ge \epsilon\}) = 0$, then $\exists f$ s.t. $f_n \to f$ in measure. Moreover, if $f_n \to g$ in measure, then $g = f$ a.e.

(Vitali Convergence Theorem)

Proof. (1)WLOG assume $f_n \le f_{n+1}, \forall n \ge 1 \forall x \in X$ since $m(\bigcup_{n=1}^\infty \{f_n > f_{n+1}\}) = 0$, (Trick: for measurable, integrable and integral etc., there's no difference of a.e. and pointwise) and $f(x) = \lim_{n\to\infty} f_n(x)$ exists in $\overline{\mathbb{R}}$. For each $n \ge 1$, pick an increasing sequence $\{\varphi_{n,k}\}_1^\infty$ of simple functions s.t. $\varphi_{n,k} \to \min\{f_n, n\}$ pointwise. (Trick: use cutoff function to approximate ∞)

Let $\phi_n = \max_{1\le i\le n}\{\varphi_{i,n}\}$, then $\{\phi_n\}_1^\infty$ is an increasing sequence of simple functions and $\phi_n \to f$ pointwise. Note if MCT holds for simple function, then $f \ge f_n \ge \phi_n$, so $\int f = \lim \int f \ge \lim \int f_n \ge \lim \int \phi_n = \int f$. Suffice to show $(\forall f \ge g \textit{ simple}) \lim \int \phi_n \ge \int g$, fix $0 < c < 1$ and let $E_n := \{\phi_n \ge cg\}$, hence $\{E_n\}$ is increasing and $X = \bigcup E_n$. So $\int \phi_n \ge \int_{E_n} \phi_n \ge c \int_{E_n} g$, then let $n \to \infty$ and $c \to 1^-$.

(2)By MCT, $\int(\underline{\lim} f_n) = \lim_{n\to\infty} \int(\inf_{k\ge n} f_k) \le \lim_{n\to\infty} \inf_{k\ge n} \int f_k = \underline{\lim} \int f_n$.

(3)Trivial for f and $f \in L^1$. By Fatou, $\int g + \int f \le \underline{\lim} \int(g + f_n) = \int g + \underline{\lim} \int f_n$ and $\int g - \int f \le \underline{\lim} \int(g - f_n) = \int g - \overline{\lim} \int f_n$.

(4)By $f_n \to f$ a.e., $\mu(\bigcup_{i=1}^\infty \bigcap_{N=1}^\infty \bigcup_{k=N}^\infty E\{|f_k - f| \ge \frac{1}{i}\}) = 0$. So $\lim_{N\to\infty} \bigcup_{k=N}^\infty E\{|f_k - f| \ge \frac{1}{i}\} = 0$, then $\forall \epsilon > 0 \forall i \ge 1 \exists N_i \ge 1$ s.t. $\mu(\bigcup_{k=N_i}^\infty E\{|f_k - f| \ge \frac{1}{i}\}) < \epsilon 2^{-i}$, let $F = E \setminus \bigcup_{i=1}^\infty \bigcup_{k=N_i}^\infty E\{|f_k - f| \ge \frac{1}{i}\}$.

(6)By $f_k \to f$ in measure, then pick subsequence $\{f_{k_j}\}$ s.t. $m(E\{|f_{k_j} - f| \ge \frac{1}{j}\}) \le \frac{1}{2^j}$. Let $E_N = \bigcup_{j=N}^\infty E\{|f_{k_j} -$

$f| \geq \frac{1}{j}\}$, $F = \bigcap_{N=1}^{\infty} E_N$, note $\{E_N\}$ monotone decreasing and $m(E_N) \leq \frac{1}{2^{N-1}}$ so $m(F) = 0$.
$\forall x \in E \setminus F \exists N \geq 1$ s.t. $x \notin E_N$ i.e. $(\forall j \geq N)|f_{k_j}(x) - f(x)| < \frac{1}{j}$, get $f_{k_j} \to f$ pointwise on $E \setminus F$.
(8)Pick subsequence $\{f_{n_j}\}$ s.t. $\mu(E_j := \{|f_{n_j} - g_{n_{j+1}}| \geq 2^{-j}\}) \leq 2^{-j}$. Let $F_k = \bigcup_{j=k}^{\infty} E_j$ and $F = \bigcap_1^{\infty} F_k$, then $\mu(F_k) \leq 2^{1-k}, \mu(F) = 0$ and $\{f_{n_j}\}$ is pointwise Cauchy on F_k^c. Let $f(x) = \begin{cases} \lim f_{n_j}(x) & x \in F^c \\ 0 & x \in F \end{cases}$, then $g_j \to f$ in measure.
Hence $f_n \to f$ in measure, since $\{|f_n - f| \geq \epsilon\} \subseteq \{|f_n - f_{n_j}| \geq \frac{1}{2}\epsilon\} \cup \{|f_{n_j} - f| \geq \frac{1}{2}\epsilon\}$. And $f = g$ a.e., since $\{|g - f| \geq \epsilon\} \subseteq \{|f - f_n| \geq \frac{1}{2}\epsilon\} \cup \{|f_n - g| \geq \frac{1}{2}\epsilon\}$.

□

Theorem 5.5.3 (Function Approximation)

(1)For nonnegative measurable f, there exists an increasing sequence $\{\varphi_n\}_1^{\infty}$ of simple function s.t. $\varphi \to f$ pointwise and $\varphi_n \to f$ uniformly on any set on which f is bounded.

(2)$f \in L^1(\mu), \epsilon > 0$, there is an integrable simple function $\varphi = \Sigma a_j \chi_{E_j}$ s.t. $\int |f - \varphi| < \epsilon$. If μ is a Lebesgue-Stieltjes measure on $\mathbb{R}$, then i could be finite; moreover, exist continuous g with compact support $||f - g||_{L^1} < \epsilon$

Theorem 5.5.4

If $f \in L^1(m)$ and $\epsilon > 0$, then $\exists$ simple $\varphi = \Sigma_1^N a_j \chi_{R_j}$ where R_j is a product of intervals s.t. $\int |f - \varphi| < \epsilon$ and $\exists$ continuous g that vanishes outside a bounded set s.t. $\int |f - g| < \epsilon$

Proof. In thm above, approximate f by simple functions, then use above (iii) to approximate the latter by function φ of the desired form. Finally, approximate φ by continuous function by applying an obvious generalization of thm above □

[Lusin; Littlewood's 2nd Principal] (Every measurable function is nearly continuous)

For measurable $E \subseteq \mathbb{R}^n$, $f : E \to \mathbb{R}$, then f is measurable iff $(\forall \epsilon > 0)(\exists E \supseteq F \textit{ closed}) m(E \setminus F) < \epsilon$ and $f|_F$ is continuous.

$\Rightarrow$: WLOG assume E is bounded o.w. consider $E \cap \{k - 1 \leq |x| < k\}$ and pick $\epsilon 2^{-k}$. By Simple Function Approximation, $\exists \varphi_n \to f$ pointwise where $\{\varphi_n\}_1^{\infty}$ is an increasing sequence of simple functions. $\forall \varphi_n = \sum_{i=1}^{N_n} a_{n_i} \chi_{E_{n_i}} \exists E_{n_i} \supseteq F_{n_i}$ *closed* s.t. $m(E_{n_i} \setminus F_{n_i}) < \epsilon 2^{-i-n}$, let $F_n = \bigcup_{i=1}^{N_n} F_{n_i}$ then $m(E \setminus F_n) < \epsilon 2^{-n}$.
By Egorov, $(\exists E \supseteq F_0 \textit{ closed}) m(E \setminus F_0) < \epsilon$ and $\varphi_n \to f$ uniformly on F_0. Let $F = \bigcap_{n=0}^{\infty} F_n$ then $E \supseteq F$ is closed, $m(E \setminus F) < 2\epsilon$, $\varphi_n|_F$ is continuous hence $f|_F$ is continuous.
$\Leftarrow$: $\forall n \in \mathbb{N} \exists E \supseteq F_n$ *closed* s.t. $m(E \setminus F_n) \leq \frac{1}{n}$ and $f|_{F_n}$ is continuous. Let $H = \bigcup_{n=1}^{\infty} F_n$ then H is an F_σ set hence measurable and $m(E \setminus H) = 0$.
$\forall a \in \mathbb{R}$, $\{f > a\} = \bigcup_{n=1}^{\infty} \{x \in F_n : f(x) > a\} \cup \{x \in E \setminus H : f(x) > a\}$. Note $\{x \in F_n : f(x) > a\}$ is relatively open wrt. F_n hence measurable, $\{x \in E \setminus H : f(x) > a\}$ is a null set hence measurable.

(3)(Uniform limit theorem)For topological space X, metric space Y, $f_n : X \to Y$ converge uniformly to $f : X \to Y$, if any f_n is continuous, then f is continuous.

Pf. $d(f(x),f(y)) \le d(f(x),f_N(x)) + d(f_N(x),f_N(y)) + d(f_N(y),f(y))$.

(Dini) For compact topological space X, increase sequence $\{f_n\}_1^\infty$ of continuous functions on X, $f_n \to f$ pointwise where f is continuous, then $f_n \to f$ uniformly.

Pf. $\forall \epsilon > 0$, note $f - f_n$ is continuous so $E_n := \{f - f_n < \epsilon\}$ is open. Note $\{E_n\}_1^\infty$ is increasing and $X = \bigcup_1^\infty E_n$, by compact, $\exists N \in \mathbb{N}$ s.t. $X = E_N$. Hence $\forall n \ge N \forall x \in X(|f(x) - f_n(x)| < \epsilon)$.

Proof. (1)For $n \ge 0$, let $E_n^k = f^{-1}([k2^{-n},(k+1)2^{-n}))$, $F_n = f^{-1}([2^n,\infty])$ and $\varphi_n = \sum_{k=0}^{2^{2n}-1} k2^{-n}\chi_{E_n^k} + 2^n\chi_{F_n}$.
(2) □

Theorem 5.5.5 (Fubini-Tonelli)

For σ-finite measure spaces $(X,\mathcal{M},\mu),(Y,\mathcal{N},\nu)$,

(1)If $E \in \mathcal{M}\otimes\mathcal{N}$, then $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable, and $\mu\times\nu(E) = \int \nu(E_x)d\mu(x) = \int \mu(E^y)d\nu(y)$.

(2)(Tonelli)If $f \in L^+(X\times Y)$, then $f_x, f^y \in L^+$, $g(x) := \int f_x d\nu, h(y) := \int f^y d\mu \in L^+$, $\int f d(\mu\times\nu) = \int(\int f(x,y)d\nu(y))d\mu(x) = \int(\int f(x,y)d\mu(x))d\nu(y)$.

(3)(Fubini)If $f \in L^1(\mu\times\nu)$, then $f_x, f^y \in L^1$ a.e., a.e.-defined $g(x) = \int f_x d\mu, h(y) = \int f^y d\nu \in L^1$, $\int f d(\mu\times\nu) = \int(\int f(x,y)d\nu(y))d\mu(x) = \int(\int f(x,y)d\mu(x))d\nu(y)$.

For complete σ-finite measure spaces $(X,\mathcal{M},\mu),(Y,\mathcal{N},\nu)$, completion $(X\times Y,\mathcal{L},\lambda)$ of $(X\times Y,\mathcal{M}\otimes\mathcal{N},\mu\times\nu)$,

(4)(Tonelli)If $f \in L^+(\lambda)$, then $f_x, f^y \in L^+$ a.e., a.e.-defined $x \mapsto \int f_x d\nu, y \mapsto \int f^y d\mu \in L^+$, $\int f d\lambda = \iint f(x,y)d\mu(x)d\nu(y) = \iint f(x,y)d\nu(y)d\mu(x)$.

(5)(Fubini)If $f \in L^1(\lambda)$, then $f_x, f^y \in L^1$, a.e.-defined $x \mapsto \int f_x d\nu, y \mapsto \int f^y d\mu \in L^1$, $\int f d\lambda = \iint f(x,y)d\mu(x)d\nu(y) = \iint f(x,y)d\nu(y)d\mu(x)$.

Proof. (1)WLOG assume μ and ν are finite, because $X\times Y = \bigcup_1^\infty(X_j\times Y_j)$ where $\{X_j\times Y_j\}$ is an increasing sequence of rectangles of finite measure then consider $E\cap(X_j\times Y_j)$ and use MCT. Let $\mathcal{C}$ be the set of $E \in \mathcal{M}\otimes\mathcal{N}$ for which conclusions hold.
For $E = A\times B$, then $\nu(E_x) = \chi_A(x)\nu(B)$, $\mu(E^y) = \mu(A)\chi_B(y)$, $\mu\times\nu(E) = \mu(A)\nu(B)$, so $E \in \mathcal{C}$. Suffice to show $\mathcal{C}$ is a monotone class.
For increasing sequence $\{E_n\} \subseteq \mathcal{C}$, $E = \bigcup_1^\infty E_n$, then measurable and increasing $f_n(y) = \mu((E_n)^y) \to f(y) = \mu(E^y)$ pointwise, hence E^y is measurable and by MCT $\int \mu(E^y)d\nu(y) = \lim\int\mu((E_n)^y)d\nu(y) = \lim\mu\times\nu(E_n) = \mu\times\nu(E)$, so $E \in \mathcal{C}$. Similarly, decreasing sequence follows by DCT.
(2)Note it holds for simple function, then Tonelli follows by simple function approximation and MCT.
(3)Note if $f \in L^+(X\times Y)$ and $\int f d(\mu\times\nu) < \infty$, then $g < \infty, h < \infty$ and $f_x, f^y \in L^1$ a.e. So for $f \in L^1(\mu\times\nu)$, Fubini follows by Tonelli.
(4)By Lemma "For $E \in \mathcal{M}\times\mathcal{N}$ with $\mu\times\nu(E) = 0$, then $\nu(E_x) = \mu(E^y) = 0$ a.e." and "For $\mathcal{L}$-measurable f with $f = 0$ λ-a.e., then $f_x, f^y \in L^1$ a.e., $\int f_x d\nu = \int f^y d\mu = 0$ a.e.". □

Remark 5.5.3

(1)Usually omit the brackets if condition holds, $\int(\int f(x,y)d\mu(x))d\nu(y) = \iint f(x,y)d\mu(x)d\nu(y) = \iint f d\mu d\nu$.

(2)The condition "σ-finite" "$f \in L^+$" or "$f \in L^1$" is necessary.

(i)To see f_x, f^y is measurable for all x, y, $\iint f d\mu d\nu, \iint f d\nu d\mu$ exist but NOT equal, f is nonnegative but NOT measurable, pick $X = Y = \omega_1$, $\mathcal{M} = \mathcal{N}$ is σ-algebra of countable or cocountable sets, $\mu = \nu : A \mapsto \begin{cases} 0 & A \text{ countable} \\ 1 & A \text{ cocountable} \end{cases}$, $E = \{(x,y) : y < x\}$, $f = \chi_E$.

(ii)To see f_x, f^y is measurable for all x, y, $\iint f d\mu d\nu, \iint f d\nu d\mu$ exist but NOT equal, f is measurable and $\int |f| d(\mu \times \nu) = \infty$, pick $X = Y = \mathbb{N}$, $\mathcal{M} = \mathcal{N} = \mathcal{P}(\mathbb{N})$, $\mu = \nu$ is counting measure, $f : (m,n) \mapsto \begin{cases} 1 & m = n \\ -1 & m = n+1 \\ 0 & o.w. \end{cases}$.

(3)Trick: to reverse the order of integration in a double integral $\iint f d\mu d\nu$, first verify $\int |f| d(\mu \times \nu) < \infty$ by Tonelli to evaluate it as an iterated integral, then apply Fubini to get $\iint f d\mu d\nu = \iint f d\nu d\mu$.

Definition 5.5.2 (Integral)

(1)(Riemann integral) For bounded $E \subseteq \mathbb{R}^n$, bounded $f : E \to \mathbb{R}$, f is *Riemann integrable* on E iff $(\exists E \subseteq R \text{ cube}) \inf_{P \text{ partition of } R} U(f,P) = \sup_{P \text{ partition of } R} L(f,P)$ where

(i)$f : R \to \mathbb{R}$ is extended by $R \setminus E \to \{0\}$

(ii)$U(f;P) = \sum_{r \in P} \sup_{x \in r} f(x)|r|$ is *upper Darboux sum* and $\inf U(f,P)$ is *upper Riemann integral.*

(iii)$L(f;P) = \sum_{r \in P} \inf_{x \in r} f(x)|r|$ is *lower Darboux sum* and $\sup L(f,P)$ is *lower Riemann integral.*

Then this value is *Riemann integral* of f on E, written $\int_E f$.

(2)(Integral) For measure space $(X, \mathcal{M}, \mu)$ where μ is a positive measure,

(i)the *integral* of nonnegative simple function $\varphi = \sum_1^n a_i \chi_{X_i}$ on X is $\int \varphi d\mu = \sum_1^n a_i \mu(X_i)$.

(ii)the *integral* of nonnegative simple function $\varphi = \sum_1^n a_i \chi_{X_i}$ on $E \in X$ is $\int_E \varphi = \int \varphi \chi_E$, where $\varphi \chi_E = \sum_1^n a_i \chi_{X_i \cap E}$ is a nonnegative simple function. Hence we only define the integral on X below.

(iii)the *integral* of nonnegative measurable function $f : X \to [0, \infty]$ is $\int f = \sup\{\int \varphi : 0 \le \varphi \le f, \varphi \text{ simple}\}$, denote the set of nonnegative measurable functions on X by $L^+(X;\mu)$ (L^+ for short).

(iv)for real-valued measurable function $f : X \to [-\infty, \infty]$, f is

(a)*integrable* iff f^+, f^- are integrable, then $\int f = \int f^+ - \int f^-$ is the *integral* of f.

(b)*extended integrable* iff at least one of f^+, f^- is integrable, then extend integral to $\pm\infty$.

(v)for complex-valued measurable function $f : X \to \mathbb{C}$, f is *integrable* iff $(\Re f)^+, (\Re f)^-, (\Im f)^+, (\Im f)^-$ are integrable, then $\int f = \int (\Re f)^+ - \int (\Re f)^- + i \int (\Im f)^+ - i \int (\Im f)^-$ is the *integral* of f.

(vi)for vector-valued measurable function $f = (f_i)_{i \in I}^T$, f is *integrable* iff f_i is integrable for all $i \in I$, then $\int f = (\int f_i)_{i \in I}^T$ is the *integral* of f.

(3)(p-seminorm) For $0 < p < \infty$, complex-valued measurable function $f : X \to \mathbb{C}^n$, (i)*p-seminorm* of f is $||f||_p = (\int |f|^p d\mu)^{\frac{1}{p}}$.

(ii)*essential supremum* of f is $\text{esssup}\, f = \inf_{\mu(N)=0} \sup_{x\in X\setminus N} f(x)$.

(iii)*essential inferior* of f is $\text{essinf}\, f = \sup_{\mu(N)=0} \inf_{x\in X\setminus N} f(x)$.

(iv)$||f||_\infty = \text{esssup}\,|f|$.

f is *p-integrable* iff $||f||_p < \infty$, then FOR $p \geq 1$ denote the set of all p-integrable functions on X by $L^p(X;\mathbb{C}^n;\mu)$ ($L^p(X;\mu)$ or L^p for short). f is *essential bounded* iff $||f||_\infty < \infty$, then denote the set of all essential bounded functions on X by $L^\infty(X;\mathbb{C}^n;\mu)$.

Remark 5.5.4

(1)The choice of cube R NOT change the value of Darboux sum, so we can replace "$\exists E \subseteq R$ *cube*" by $\forall E \subseteq R$ *cube*.

Btw, it could be proved that "f is Riemann integrable on R iff $(\forall\epsilon > 0)(\exists P\ \text{partition of}\ R)U(f,P) - L(f,P) < \epsilon$". Then it is a trick that let $\omega_F = \sup_{x\in F} f(x) - \inf_{x\in F} f(x)$ and get "f is Riemann integrable iff $\lim_{||P||\to 0^+} \sum_{F\in P} \omega_F = 0$", which would be an easy way to show Lebesgue Criterion.

(2)For bounded $E \subseteq \mathbb{R}^n$, E is *Jordan measurable* iff χ_E is Riemann integrable, then denote $\int_E \chi_E$ by *(Jordan) content (or Peano content)* $|E|$. Similarly, *outer (Jordan) content* is $\inf U(\chi_E, P)$ while *inner (Jordan) content* is $\sup L(\chi_E, P)$.

Then we could verify it is a content. Although we could define Jordan content first, then define the integral viz. the restriction of Riemann integral for Jordan measurable space, but it is not worthy of an arduous work.

(3)For $\int \varphi d\mu$, written $\int \varphi$ for short and $\int \varphi(x) d\mu(x)$ for long. Specially for Lebesgue measure, we always denote $\int \varphi(x) d\mu(x)$ by $\int \varphi(x) dx$ for Lebesgue-Stieltjes integral.

(4)$L^p(\mu)$ is a Banach space, note integral is a linear functional on it.

To see Bolzano-Weierstrass, Heine-Borel, Accumulation Point Theorem NOT hold, Pick $L^p(\mathbb{R})$, $f_n = \chi_{[n,n+1]}$, bounded closed $F = \{f_n\}$, open cover $B_n := \{g \in L^p : ||g - f_n||_p < \frac{1}{2}\}$.

(5)There's a problem about Riemann integral. It can be NOT integrable after cutoff, while Lebesgue integral always holds.

To see $f \in C([0,1])$, $g(x) = \begin{cases} f(x) & f(x) \geq 1 \\ 0 & o.w. \end{cases}$ is NOT Riemann integrable but Lebesgue integrable on $[0,1]$, pick $f = \begin{cases} 1 & x \in K \\ 1 - d(x,K) & o.w. \end{cases}$ where K is a generalized Cantor set.

Theorem 5.5.6 (Lebesgue Criterion)

For bounded $E \subseteq \mathbb{R}^n$, bounded $f : E \to \mathbb{R}$,

(i)If f is Riemann integrable, then f is Lebesgue integrable and two integrals are equal.

(ii)f is Riemann integrable iff the set of discontinuities of f on E has Lebesgue measure zero.

Proof. Note $\exists\{P_k\}_1^\infty$ is a sequence of successive refinement of partitions of rectangle $R \supseteq E$, so $\lim U(f, P_k)$

equals the upper Riemann integral while $\lim L(f, P_k)$ equals the lower Riemann integral. By $L(f, P_k), U(f, P_k)$ can be represented as Lebesgue integrals of simple functions l_k and u_k, then $l_k \to l$ and $u_k \to u$ pointwise and $l \le f \le u$. By Dominated Convergence Theorem, $\int ldm = \lim L(f, P_k), \lim U(f, P_k) = \int udm$.
For (i), if f is Riemann integrable, then $\int l = \lim L(f, P_k) = \lim U(f, P_k) = \int u$ so $\int f = \int l = \int u$.
For (ii), choose P_k s.t. $||P_k|| < \frac{1}{k}$, note f is continuous at x iff $l(x) = u(x)$, so f is Riemann integrable iff $\int u = \int l$ iff $u = l$ a.e. iff the set of discontinuities of f on E has Lebesgue measure zero. □

Remark 5.5.5

(1)Pick χ_E, then get Corollary

(i)For Jordan measurable $E \subseteq \mathbb{R}^n$, E is Lebesgue measurable and $|E| = m(E)$.

(ii)For bounded $E \subset \mathbb{R}^n$, E is Jordan measurable iff $m(\partial E) = 0$.

Actually for bounded $E \subseteq \mathbb{R}^n$, the inner content of E is the Lebesgue measure of $\text{Int}(E)$, and the outer content of E is the Lebesgue measure of $\overline{E}$.

(2)To see there exists a Jordan nonmeasurable bounded domain, pick $K \subseteq [0,1]$ is generalized Cantor set with positive measure, then $U = ((0,1) \times (-1,1) \setminus (K \times [0,1]))$ is bounded domain and $\partial U = [0,1] \times [-1,1] \setminus U$ has positive measure.

To see there exists an Riemann integrable function on a Jordan nonmeasurable set that is not zero at every pt, pick χ_C on C where C is the Cantor set.

To see there exists a positive function on $[0,1]$ with lower Riemann integral 0, pick *Riemann function* $R : \mathbb{R} \to \mathbb{R}, x \mapsto \begin{cases} 1/p & x = q/p, p \in \mathbb{Z}^+, q \in \mathbb{Z}, gcd(p,q) = 1 \\ 1 & x = 0 \\ 0 & x \notin \mathbb{Q} \end{cases}$ then replace $x \mapsto 0 \quad x \notin \mathbb{Q}$ by $x \mapsto 1 \quad x \notin \mathbb{Q}$.

To see there exist a Riemann integrable function f on $[0,2]$ and a continuous bijective function $g : [0,2] \to [0,2]$ s.t. $f \circ g$ is not Riemann integrable, pick $f = \chi_{2C}$ where C is the Cantor set, $h(x) = x + C(x) : [0,1] \to [0,2]$ where $C(x)$ is the Cantor function, $g = 2h^{-1}$ since h is strictly increasing and continuous. The set of continuities of f is $2C$, then the set of continuities of $f \circ g$ is $g^{-1}(2C) = h(C)$, note $m(h(C)) = \mu_h(h^{-1}(h(C))) = \mu_h(C) = m(C) + \mu_C(C) = 1$. (Use the property of Lebesgue-Stieltjes measure "For continuous increasing function G on $[a,b]$, if $E \subseteq [G(a), G(b)]$ is a Borel set, then $m(E) = \mu_G(G^{-1}(E))$".)

(3)Deeply speaking, the essential difference between Jordan content and Lebesgue measure is that Jordan content is a 1-step approximation while Lebesgue measure is a 2-step approximation. The first approximates from the outside by open sets and from the inside by compact sets, and the second approximates the open sets from the inside and the compact sets from the outside by finite unions of cubes.

Proposition 5.5.7

(1)For $f, g \in L^1$, then $(\forall E \in \mathcal{M}) \int_E f = \int_E g$ iff $\int |f - g| = 0$ iff $f = g$ a.e.

(2)(1st Mean Value Theorem for integral) For f bounded, nonnegative $g \in L^1$, then $\exists \inf_{x \in E} f(x) \le$

$\eta \leq \sup_{x\in E} f(x)$ s.t. $\int_E fg = \eta \int_E g$. Specially, if E is connected and f is continuous, then $\exists \xi \in E$ s.t. $\int_E fg = f(\xi) \int_E g$.

Proof. (1)Suffice to show $\int |f-g| = 0 \Rightarrow (\forall E \in \mathcal{M}) \int_E f = \int_E g \Rightarrow f = g$ a.e. First $|\int_E f - \int_E g| \leq \int \chi_E |f-g| \leq \int |f-g| = 0$. Second suppose not, then at least one of f^+, f^- be nonzero on a set of positive measure. WLOG $m(\{f^+ > 0\}) > 0$, then $\int_{\{f^+>0\}} f - \int_{\{f^+>0\}} g > 0$ contradiction.
(2)If $\int_E g = 0$, then $g = 0$ a.e. so $\int_E fg = 0$. Assume $\int_E g > 0$, then $m(\{g > 0\}) > 0$ so let $\eta = \frac{\int_E fg}{\int_E g}$. Know $f(E)$ is connected, if $\eta \notin f(E)$, then wlog assume $\eta = \inf f(E)$. So $m(\{fg > \eta g\}) = m(\{g > 0\})$ get $\int_E fg > \eta \int_E g$ contradiction. □

Remark 5.5.6

(1)Integral makes no difference if alter functions on null sets. So redefine $L^1(\mu)$ be the set of equivalent classed of a.e.-defined integrable functions on X. It has two advantages:

(i)For completion $\overline{\mu}$, there exists a natural bijection between $L^1(\mu)$ and $L^1(\overline{\mu})$, so we can identify them.

(ii)Then positivity holds. L^1 is a metric space with $d(f,g) = \int |f-g|$ while seminorm becomes norm.

Exercise 5.5.1

(1)$m(E) < \infty$, $f \in L^\infty$, then $\lim_{p\to\infty} ||f||_p = M := ||f||_\infty$. Moreover, if $M > 0$, then $\lim_{p\to\infty} \frac{\int_E |f|^{p+1}}{\int_E |f|^p} = M$.

(2)$\{f_n\}_1^\infty \subseteq L^+$, $f_n \to f$ a.e. and $\int f = \lim \int f_n < \infty$, then $(\forall E \in \mathcal{M}) \int_E f = \lim \int_E f_n$. However, it NOT holds if $\int f = \lim \int f_n = \infty$.

Proof. (1)$\overline{\lim}_{p\to\infty} (\int_E |f|^p)^{\frac{1}{p}} \leq \overline{\lim}_{p\to\infty} M(m(E))^{\frac{1}{p}} = M$, wlog assume $M > 0$, $\underline{\lim}_{p\to\infty} (\int_E |f|^p)^{\frac{1}{p}} \geq \underline{\lim}_{p\to\infty} (\int_{E\{|f|>M-\epsilon\}} |f|^p$
$(M-\epsilon) \underline{\lim}_{p\to\infty} m(E\{|f| > M-\epsilon\})^{\frac{1}{p}} = M - \epsilon$ for any $\epsilon > 0$.
Similarly $\overline{\lim}_{p\to\infty} \frac{\int_E |f|^{p+1}}{\int_E |f|^p} \leq M$, note $\overline{\lim}_{p\to\infty} \frac{\int_{E\{|f|<M-\epsilon\}} |f|^p}{\int_E |f|^p} \leq \overline{\lim}_{p\to\infty} \frac{\int_{E\{|f|<M-\epsilon\}} |f|^p}{\int_{E\{|f|>M-\epsilon/2\}} |f|^p} \leq \frac{m(E\{|f|\leq M-\epsilon\})}{m(E\{|f|\geq M-\epsilon/2\})} \overline{\lim}_{p\to\infty} \frac{(M-\epsilon)^p}{(M-\epsilon/2)}$
0 for any $\epsilon > 0$, then $\underline{\lim}_{p\to\infty} \frac{\int_E |f|^{p+1}}{\int_E |f|^p} \geq (M-\epsilon) \underline{\lim}_{p\to\infty} \frac{\int_{E\{|f|\geq M-\epsilon\}} |f|^p}{\int_E |f|^p} = M - \epsilon$.
(2)By Fatou, $\int_E f = \int \underline{\lim} f_n \chi_E \leq \underline{\lim} \int f_n \chi_E \leq \lim \int_E f_n$, similarly $\int_{E^c} f \leq \lim \int_{E^c} f_n$, then consider the sum.
Counterexample, $X = \mathbb{R}, \mu = m, E_n = (-\infty, 0) \cup [n, n+1], f_n = \chi_{E_n}, E = [0, \infty)$. □

Definition 5.5.3

gamma function $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ $(\Re\ z > 0)$ Well-defined: $\int_0^1 |t^{z-1} e^{-t}| dt \leq \int_0^1 t^{\Re\ z-1} dt < \infty$ and $\int_1^\infty |t^{z-1} e^{-t}| dt \leq \int_1^\infty e^{-t/2} < \infty$ Prop: (1)$\Gamma(z+1) = z\Gamma(z)$ by integration by parts $\int_\epsilon^N t^z e^{-t} dt = -t^z e^{-t}|_\epsilon^N + z \int_\epsilon^N t^{z-1} e^{-t} dt$ then $\epsilon \to 0, N \to \infty$ $\Gamma(z+1) = z\Gamma(z)$ can extend Γ to almost(except for singularities at the nonpositive integers) the entire complex plane (2)$\Gamma(n+1) = n!$ (Many of the applications of the gamma function involve the fact that it provides an extension of the factorial function to nonintegers)

Construction of this surface measure is motivated by a familiar fact from plane geometry: define the surface measure of a subset of the unit sphere in terms of the Lebesgue measure of the corresponding sector of the unit ball.

Definition 5.5.4

denote $\{x \in \mathbb{R}^n : |x| = 1\}$ by S^{n-1}. If $x \in \mathbb{R}^n \setminus \{0\}$, *polar coordinate* of x are $r = |x| \in (0,\infty)$, $x' = \frac{x}{|x|} \in S^{n-1}$. The map $\Phi(x) = (r, x')$ is a continuous bijection from $\mathbb{R}^n \setminus \{0\}$ to $(0,\infty) \times S^{n-1}$ whose (continuous) inverse is $\Phi^{-1}(r, x') = rx'$. Denote the Borel measure on $(0,\infty) \times S^{n-1}$ by m_* induced by Φ from Lebesgue measure on $\mathbb{R}^n$, that is $m_*(E) = m(\Phi^{-1}(E))$. Moreover define measure $\rho = \rho_n$ on $(0,\infty)$ by $\rho(E) = \int_E r^{n-1} dr$

Theorem 5.5.8

Exists a unique Borel measure $\sigma = \sigma_{n-1}$ on S^{n-1} s.t. $m_* = \rho \times \sigma$. if f is Borel measurable on $\mathbb{R}^n$ and $f \geq 0$ or $f \in L^1(m)$, then $\int_{\mathbb{R}^n} f(x)dx = \int_0^\infty \int_{S^{n-1}} f(rx')r^{n-1} d\sigma(x')dr$

Proof. When f is a characteristic function of a set, is merely a restatement of $m_* = \rho \times \sigma$ and it follows for general f by the usual linearity and approximation arguments. Hence we need only to construct σ. If E is a Borel set in S^{n-1}, for $a > 0$ let $E_a = \Phi^{-1}((0,a] \times E) = \{rx' : 0 < r \leq a, x' \in E\}$. If it holds when $f = \chi_{E_1}$ we must have $m(E_1) = \int_0^1 \int_E r^{n-1} d\sigma(x')dr = \sigma(E) \int_0^1 r^{n-1} dr = \sigma(E)/n$

Therefore define $\sigma(E)$ to be $n \cdot m(E_1)$ since the map $E \mapsto E_1$ takes Borel sets to Borel sets and commutes with unions, intersections and complements. It is clear that σ is a Borel measure on S^{n-1}. Also, since E_a is the image of E_1 under the map $x \mapsto ax$, it follows from Thm2.44 that $m(E_a) = a^n m(E_1)$ hence if $0 < a < b$, $m_*((a,b] \times E) = m(E_b \setminus E_a) = \frac{b^n - a^n}{n}\sigma(E) = \sigma(E) \int_a^b r^{n-1} dr = \rho \times \sigma((a,b] \times E)$ Fix $E \in \mathcal{B}_{S^{n-1}}$ and let $\mathcal{A}_E$ be the collection of finite disjoint unions of sets of the form $(a,b] \times E$. By proposition1.7, $\mathcal{A}_E$ is an algebra on $(0,\infty) \times E$ that the σ-algebra $\mathcal{M}_E = \{A \times E : A \in \mathcal{B}_{(0,\infty)}\}$. By the preceding calculation we have $m_* = \rho \times \sigma$ on $\mathcal{A}_E$ and hence by the uniqueness assertion of Thm1.14 $m_* = \rho \times \sigma$ on $\mathcal{M}_E$. But $\bigcup\{\mathcal{M}_E : E \in \mathcal{B}_{S^{n-1}}\}$ is precisely the set of Borel rectangles in $(0,\infty) \times S^{n-1}$, so another applications of the uniqueness theorem shows that $m_* = \rho \times \sigma$ on all Borel sets □

Remark 5.5.7

Of course, it can be extended to Lebesgue measurable functions by considering the completion of the measure σ

Corollary 5.5.9

If f is a measurable function on $\mathbb{R}^n$, nonnegative or integrable, s.t. $f(x) = g(|x|)$ for some function g on $(0,\infty)$, then $\int f(x)dx = \sigma(S^{n-1}) \int_0^\infty g(r)r^{n-1} dr$

Corollary 5.5.10

Let c and C denote positive constants, let $B = \{x \in \mathbb{R}^n : |x| < c\}$. Suppose that f is a measurable function on $\mathbb{R}^n$ (1)If $|f(x)| \leq C|x|^{-a}$ on B for some $a < n$, then $f \in L^1(B)$. However, if $|f(x)| \geq C|x|^{-n}$ on B, then $f \notin L^1(B)$ (2)If $|f(x)| \leq C|x|^{-a}$ on B^c for some (??)$a > n$, then $f \in L^1(B^c)$. However, if $|f(x)| \geq C|x|^{-n}$ on B, then $f \notin L^1(B^c)$

Pf. apply above to $|x|^{-a}\chi_B$ and $|x|^{-a}\chi_{B^c}$

Proposition 5.5.11

(1)If $a > 0$, then $\int_{\mathbb{R}^n} \exp(-a|x|^2)dx = (\frac{\pi}{a})^{n/2}$ (2)$\sigma(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$ (3)If $B^n = \{x \in \mathbb{R}^n : |x| < 1\}$, then $m(B^n) = \frac{\pi^{n/2}}{\Gamma(\frac{1}{2}n+1)}$ (4)$\Gamma(n) = (n-1)!$, $\Gamma(n+\frac{1}{2}) = (n-\frac{1}{2})(n-\frac{3}{2})(\frac{1}{2})\sqrt{\pi}$

Proof. (1)Denote the integral on the left by I_n. For $n = 2$, by Corollary2.51, $I_2 = 2\pi\int_0^\infty re^{-ar^2}dr = -(\frac{\pi}{a})e^{-ar^2}|_0^\infty = \frac{\pi}{a}$ Since $\exp(-a|x|^2) = \Pi_1^n \exp(-ax_j^2)$, Tonelli implies $I_n = (I_1)^n$. In particular, $I_1 = (I_2)^{1/2}$ so $I_n = (I_2)^{n/2} =$ (2)by above and substitution $s = r^2$, $\pi^{n/2} = \int_{\mathbb{R}^n} e^{-|x|^2}dx = \sigma(S^{n-1})\int_0^\infty r^{n-1}e^{-r^2}dr = \frac{\sigma(S^{n-1})}{2}\int_0^\infty s^{n/2-1}e^{-s}ds = \frac{\sigma(S^{n-1})}{2}\Gamma(\frac{n}{2})$ (3)$m(B^n) = n^{-1}\sigma(S^{n-1})$ by definition of σ, and $\frac{1}{2}n\Gamma(\frac{1}{2}n) = \Gamma(\frac{1}{2}n+1)$ by the functional equation for the gamma function. (4)By the functional equation, $\Gamma(n+\frac{1}{2}) = (n-\frac{1}{2})...(\frac{1}{2})\Gamma(\frac{1}{2})$, then $s = r^2$, $\Gamma(\frac{1}{2}) = \int_0^\infty s^{-1/2}e^{-s}ds = 2\int_0^\infty e^{-r^2}dr = \int_{-\infty}^\infty e^{-r^2}dr = \sqrt{\pi}$ □

5.6 Differentiation

Theorem 5.6.1 (The Hahn Decomposition Theorem)

ν signed measure on $(X, \mathcal{A})$, then $\exists$ positive set P and negative set N s.t. $P \sqcup N = X$. If P',N' is another such pair, then $P \triangle P'$ is null for ν

Proof. WLOG assume ν does not obtain the value $+\infty$. $m = \sup\{\nu(E) : E \textit{ positive set}\}$. So exists $\{P_j\}$ of positive sets with $\nu(P_j) \to m$. Let $P = \bigcup_1^\infty P_j$, note positive set closec under countable union, P is a positive set and $\nu(P) = m$.

Claim $N = X \setminus P$ is a negative set. First, N cannot contain any nonnull positive set, o.w. $E \subseteq N$ is positive with $\nu(E) > 0$, then consider $E \cup P$.

Second, by above, if $A \subseteq N, \nu(A) > 0$, there exist $B \subseteq A$ with $\nu(B) > \nu(A)$ with $\nu(A \setminus B) < 0$.

If N is not negative, then, we can specify a sequence of subsets $\{A_j\}$ of N and a sequence $\{n_j\}$ of positive integers as follows: $n_i = \inf\{n : \exists B \subseteq A_{i-1}, \nu(B) > \nu(A_{i-1}) + \frac{1}{n}\}$ specially $n_1 = \inf\{n : \exists B \subseteq N, \nu(B) > \frac{1}{n}\}$, and A_i is such a set.

Let $A = \bigcap_1^\infty A_j$. Then $\infty > \nu(A) = \lim_{j\to\infty}\nu(A_j) > \sum_1^\infty \frac{1}{n_j}$, so $n_j \to \infty$ as $j \to \infty$.But once again, there exists $B \subseteq A$ with $\nu(B) > \nu(A) + \frac{1}{n}$ for some $n \in \mathbb{N}$. For j sufficiently large we have $n < n_j$, and $B \subseteq A_{j-1}$, which contradicts the construction of n_j and A_j.

Finally, if P', N' is another pair, we observe that $P \setminus P' \subseteq P$ and $P \setminus P' \subseteq N'$ so it's both positive and negative hence null.

□

Remark 5.6.1

(1)*Hahn decomposition*, it is usually not unique (ν-null sets can be transferred), but it lead to canonical representation as the difference of two positive measures (2)μ, ν signed measure, are *mutually singular* or ν is *singular w.r.t.* μ, written $\mu \perp \nu$ iff $(\exists E, F \in \mathcal{A}) E \cap F = \emptyset, E \cup F = X, E\ null\ for\ \mu, F\ null\ for\ \nu$ (Informally speaking, mutual singularity means that μ,ν live on disjoint sets)

Definition 5.6.1 (The Jordan Decomposition Theorem)

If ν signed measure, then exists unique positive measures ν^+ and ν^- s.t. $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$

Proof. Let $X = P \cup N$ be a Hahn decomposition, define $\nu^+(E) = \nu(E \cap P)$ and $\nu^-(E) = -\nu(E \cap N)$. Then clearly $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$. If $\nu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$, then assume $E, F \in \mathcal{M}$ s.t. $E \cap F = \emptyset$, $E \cup F = X$ and $\mu^+(F) = \mu^-(E) = 0$, then $X = E \cup F$ is another Hahn decomposition, so $P \triangle E$ ν-null, so $(\forall A \in \mathcal{M}) \mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) = \nu^+(A)$. □

Remark 5.6.2

(1)ν^+, ν^- called *positive (negative) variation* of ν, and $\nu = \nu^+ - \nu^-$ is *Jordan decomposition* of ν. *total variation* of ν is $|\nu| := \nu^+ + \nu^-$ Prop. $E \in \mathcal{A}$ ν-null iff $|\nu|(E) = 0$ $\nu \perp \mu$ iff $|\nu| \perp \mu$ iff $\nu^+ \perp \mu$ and $\nu^- \perp \mu$ (2)observe ν is of the form $\nu(E) = \int_E f d\mu$ where $\mu = |\nu|$ and $f = \chi_P - \chi_N$ where $X = P \cup N$ is a Hahn decomposition. (3)$L^1(\nu) = L^1(\nu^+) \cap L^1(\nu^-) = L^1(|\mu|)$ and $\int f d\nu = \int f d\nu^+ - \int f d\nu^-$ ν is *finite (resp. σ-finite)* iff $|\nu|$ is finite (resp. σ-finite). (4)absolutely continuous: ν signed measure, μ measure, then $\nu \ll \mu$ iff $(\forall \mu(E) = 0) \nu(E) = 0$ (extended) ν, μ signed measure, then $\nu \ll \mu$ iff $\nu \ll |\mu|$ Prop. $\nu \ll \mu$ iff $\nu_1, \nu_2 \ll \mu$ iff $|\nu| \ll \mu$ $\nu \ll \mu$, $\nu \perp \mu$, then $\nu = 0$ Pf. $X = E \sqcup F$, $\mu(E) = |\nu|(F) = 0$ and then $|\nu|(E) = 0$

Theorem 5.6.2

ν finite (σ-finite has counterexample) signed measure, $\nu \ll \mu$ iff $(\forall \epsilon > 0)(\exists \delta > 0)(\forall \mu(E) < \delta)|\nu(E)| < \epsilon$

Proof. Since $\nu \ll \mu$ iff $|\nu| \ll \mu$ and $|\nu(E)| \leq |\nu|(E)$, it suffices to assume $\nu = |\nu|$ is positive. The left side is trivial, on the other hand, suppose not. Then $\exists \epsilon > 0$ s.t. $(\forall n \in \mathbb{N})(\exists E_n \in \mathcal{M} \wedge \mu(E_n) < 2^{-n})\nu(E_n) \geq \epsilon$. Let $F_k = \bigcup_k^\infty E_n$ and $F = \bigcap_1^\infty F_k$. Then $\mu(F_k) < 2^{1-k}$ so $\mu(F) = 0$. But $\nu(F_k) \geq \epsilon$ and by ν finite $\nu(F) = \lim \nu(F_k) \geq \epsilon$ Thus it is false that $\nu \ll \mu$ □

If μ is a measure and f is an extended μ-integrable function, the signed measure ν defined by $\nu = \int_E f d\mu$ is absolutely continuous wrt. μ. It is finite iff $f \in L^1(\mu)$. For complex-valued, also hold, so get:

Corollary 5.6.3

If $f \in L^1(\mu)$, then $(\forall \epsilon > 0)(\exists \delta > 0)|\int_E f d\mu| < \epsilon$ whenever $\mu(E) < \delta$

We shall use $d\nu = f d\mu$ to express the relationship $\nu(E) = \int_E f d\mu$

Lemma 5.6.4

ν, μ finite measures on $(X.\mathcal{M})$, either $\nu \perp \mu$ or $\exists \epsilon > 0, E \in \mathcal{M}$ s.t. $\mu(E) > 0$ and $\nu \geq \epsilon\mu$ on E

Proof. Let $X = P_n \sqcup N_n$ be a Hahn decomposition for $\nu - n^{-1}\mu$, $P = \bigcup_1^\infty P_n, N = \bigcap_1^\infty N_n$. Then N is a negative set for $\nu - n^{-1}\mu$, get $0 \leq \nu(N) \leq n^{-1}\mu(N)$ for all n so $\nu(N) = 0$. If $\mu(P) = 0$ then $\nu \perp \mu$, o.w. $\mu(P_n) > 0$ for some n, and P_n is a positive set for $\nu - n^{-1}\mu$ □

Theorem 5.6.5 (Lebesgue-Radon-Nikodym decomposition theorem)

ν σ-finite signed measure and μ σ-finite positive measure on $(X, \mathcal{M})$. Then $\exists$ unique σ-finite signed measures λ, ρ on $(X, \mathcal{M})$ s.t. $\lambda \perp \mu$, $\rho \ll \mu$ and $\nu = \lambda + \rho$. Moreover, there is an extended μ-integrable function $f : X \to \mathbb{R}$ s.t. $d\rho = fd\mu$ and any two such functions are equal μ-a.e.

Proof. 1 If ν and μ are finite positive measure. Let $\mathcal{F} = \{f : X \to [0, \infty] : (\forall E \in \mathcal{M}) \int_E fd\mu \leq \nu(E)\}$. $\mathcal{F}$ is nonempty since $0 \in \mathcal{F}$. If $f, g \in \mathcal{F}$, then $h = \max(f, g) \in \mathcal{F}$ since $\int_E hd\mu = \int_{E\cap A} fd\mu + \int_{E\setminus A} gd\mu \leq \nu(E \cap A) + \nu(E \setminus A) = \nu(E)$ where $A = \{x : f(x) > g(x)\}$. Let $a = \sup\{\int fd\mu : f \in \mathcal{F}\}$, noting $a \leq \nu(X) < \infty$ and choose a sequence $\{f_n\} \subseteq \mathcal{F}$ s.t. $\int f_n d\mu \to a$. Let $g_n = \max(f_1, ..., f_n)$ and $f = \sup f_n$, then $g_n \in \mathcal{F}$, g_n increases pointwise to f, so $\int g_n d\mu = a$ hence by MCT $\int fd\mu = a$. Claim λ s.t. $d\lambda = d\nu - fd\mu$ is singular wrt. μ. If not, then by lemma, $\exists E \in \mathcal{M}, \epsilon > 0$ s.t. $\mu(E) > 0$ and $\lambda \geq \epsilon\mu$ on E. Then $\epsilon\chi_E d\mu \leq d\lambda$, i.e. $(f + \epsilon\chi_E)d\mu \leq d\nu$ so $f + \epsilon\chi_E \in \mathcal{F}$, contradiction. Suffice to show uniqueness. If also $d\nu = d\lambda' + f'd\mu$ then $d\lambda - d\lambda' = (f' - f)d\mu$. Note $\lambda - \lambda' \perp \mu$ and $(f' - f)d\mu \ll d\mu$ hence $d\lambda - d\lambda' = (f' - f)d\mu = 0$. So $\lambda = \lambda'$ and $f = f'$ μ-a.e. 2 if both σ-finite, then $X = \bigsqcup_1^\infty A_j$ where $\nu(A_j), \mu(A_j) < \infty$ by taking intersection, let $\mu_j(E) = \mu(E \cap A_j)$ and $\nu_j(E) = \nu(E \cap A_j)$, use above. 3 ν signed measure, then ν^+ and ν^- □

Remark 5.6.3

Lebesgue decomposition $\nu = \lambda + \rho$ where $\lambda \perp \mu$ and $\rho \ll \mu$. If $\nu \ll \mu$, then $d\nu = fd\mu$ for some f, this result is *Radan-Nikodym Theorem*, f is *Radan-Nikodym derivative* of ν wrt. μ, denote f(the class of functions equal to f μ-a.e.) by $\frac{d\nu}{d\mu}$

Proposition 5.6.6

ν σ-finite signed measure and μ, λ σ-finite measure on $(X, \mathcal{M})$ s.t. $\nu \ll \mu$ and $\mu \ll \lambda$. (1)If $g \in L^1(\nu)$, then $g(\frac{d\nu}{d\mu}) \in L^1(\mu)$ and $\int gd\nu = \int g\frac{d\nu}{d\mu}d\mu$ (2)$\nu \ll \lambda$ and $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$ λ-a.e.

Proof. By considering ν^+ and ν^- separately, assume $\nu \geq 0$. $\int gd\nu = \int g(\frac{d\nu}{d\mu})d\mu$ is true when $g = \chi_E$ by definition, then is true for simple functions by linearity then for nonnegative measurable functions by MCT and finally for functions in $L^1(\nu)$ by linearity. replace ν, μ by μ, λ and let $g = \chi_E\frac{d\nu}{d\mu}$, obtain $\nu(E) = \int_E \frac{d\nu}{d\mu}d\mu = \int_E \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}d\lambda$ for all $E \in \mathcal{M}$. □

Corollary 5.6.7

(1)$\mu \ll \lambda, \lambda \ll \mu$, then $\frac{d\lambda}{d\mu}\frac{d\mu}{d\lambda} = 1$ a.e. (wrt. λ or μ) (2)(simple but important observation) If $\mu_1, ..., \mu_n$ are measures on $(X, \mathcal{M})$, then $\exists$ measure μ s.t. $\mu_j \ll \mu$ for all j, namely $\mu = \sum_1^n \mu_j$.

For complex measure, let ν_r, ν_i be the real and imaginary parts of ν (note ν_r, ν_i are signed measure that don't assume the values $\pm\infty$ hence finite) $\nu \perp \mu$ iff $\nu_a \perp \mu_b$ where $a, b \in \{r, i\}$ $\nu \ll \lambda$ iff $\nu_r \ll \lambda$ and $\nu_i \ll \lambda$

Theorem 5.6.8 (Lebesgue-Radan-Nikodym Theorem)

ν complex measure, μ σ-finite positive measure on $(X, \mathcal{M})$, then $\exists$ complex measure λ and $f \in L^1(\mu)$ s.t. $\lambda \perp \mu$ and $d\nu = d\lambda + f d\mu$. If also $\lambda' \perp \mu$ and $d\nu = d\lambda' + f' d\mu$, then $\lambda = \lambda'$ and $f = f'$ μ-a.e.

Definition 5.6.2

total variation of complex measure ν is the positive measure $|\nu|$ determined by "if $d\nu = f d\mu$ where μ positive measure, then $d|\nu| = |f| d\mu$". Well-defined: 1 we can take $\mu = |\nu_r| + |\nu_i|$ and use Radan to get $d\nu = f d\mu$ exist 2 if $d\nu = f_1 d\mu_1 = f_2 d\mu_2$, let $\rho = \mu_1 + \mu_2$, then $f_1 \frac{d\mu_1}{d\rho} d\rho = d\nu = f_2 \frac{d\mu_2}{d\rho} d\rho$ so that $f_1 \frac{d\mu_1}{d\rho} = f_2 \frac{d\mu_2}{d\rho}$ ρ-a.e. Since $\frac{d\mu_j}{d\rho}$ is nonnegative, $|f_1| \frac{d\mu_1}{d\rho} = |f_1 \frac{d\mu_1}{d\rho}| = |f_2 \frac{d\mu_2}{d\rho}| = |f_2| \frac{d\mu_2}{d\rho}$ ρ-a.e. Thus $|f_1| d\mu_1 = |f_1| \frac{d\mu_1}{d\rho} d\rho = |f_2| d\mu_2$, hence the definition of $|\nu|$ is independent of the choice of μ and f 3 This definition agrees with the previous definition when signed measure, for in that case $d\nu = (\chi_P - \chi_N) d|\nu|$ where $X = P \sqcup N$ is a Hahn decomposition and $|\chi_P - \chi_N| = 1$

Proposition 5.6.9

(1)$|\nu(E)| \ll |\nu|(E)$ for all $E \in \mathcal{M}$ (2)$\nu \ll |\nu|$ and $\frac{d\nu}{d|\nu|}$ has absolute value 1 $|\nu|$-a.e. (3)$L^1(\nu) = L^1(|\nu|)$ and if $f \in L^1(\nu)$, then $|\int f d\nu| \leq \int |f| d|\nu|$ (4)$|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$

Proof. Suppose $d\nu = f d\mu$ as in the define of $|\nu|$. Then $|\nu(E)| = |\int_E f d\mu| \leq \int_E |f| d\mu = |\nu|(E)$. If $g = \frac{d\nu}{d|\nu|}$, then $f d\mu = d\nu = g d|\nu| = g|f| d\mu$, so $g|f| = f$ μ-a.e. hence $|\nu|$-a.e., note $|f| > 0$ $|\nu|$-a.e. so $|g| = 1$ $|\nu|$-a.e. □

Take $(X, \mathcal{M}) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ $\mu = m$ below *pointwise derivative* of ν wrt. m, let $B(x, r)$ be the open ball of radius r about x in $\mathbb{R}^n$, then $F(x) = \lim_{r \to 0} \frac{\nu(B(x,r))}{m(B(x,r))}$ exists For $c \in \mathbb{R}$, $cB := B(x, cr)$

Lemma 5.6.10 (Vitali Covering Lemma)

(1)(finite version)Metric space, $\{B_j\}_1^n$, then exist $\{B_{j_i}\}_{i=1}^m$ disjoint and $\bigcup_{j=1}^n B_j \subseteq \bigcup_{i=1}^m 3B_{j_i}$ (2)(infinite version)Separable metric space, $\mathcal{F} = \{B_j\}_{j \in J}$ s.t. $R := \sup\{\text{rad } (B) : B \in \mathcal{F}\} < \infty$, then exist a countable sub-collection $\mathcal{G} \subseteq \mathcal{F}$ s.t. disjoint and $\bigcup_{B \in \mathcal{F}} B \subseteq \bigcup_{C \in \mathcal{G}} 5C$. Moreover, each $B \in \mathcal{F}$ intersects some $C \in \mathcal{G}$ with $B \subseteq 5C$

Proof. (1)Assume $n > 0$, let B_{j_1} be the ball of maximal radius. Once $B_{j_1}, ..., B_{j_k}$ are chosen, if there is some ball in $\{B_j\}$ that is disjoint from $\bigcup_{i=1}^k B_{j_i}$, then let $B_{j_{k+1}}$ be such ball with maximal radius, o.w. set $m = k$ and terminate. $(\forall B_j)(\exists B_{j_i})$ has the minimal i s.t. $B_j \cap B_{j_i} \neq \emptyset$, then $B_j \subseteq 3B_{j_i}$ (2)Let $\mathcal{F}_n = \{B \in \mathcal{F} : 2^{-n-1}R < rad(B) \leq 2^{-n}R\}, n \geq 0$. First let $\mathcal{H}_0 = \mathcal{F}_0$ and $\mathcal{G}_0$ be a maximal disjoint subcollection of $\mathcal{H}_0$

(exist by Zorn). Once $\mathcal{G}_0, ..., \mathcal{G}_n$ are chosen, let $\mathcal{H}_{n+1} = \{B \in \mathcal{F}_{n+1} : B \cap \bigcup_0^n \mathcal{G}_i = \emptyset\}$ and $\mathcal{G}_{n+1}$ be a maximal disjoint subcollection of $\mathcal{H}_{n+1}$. $G := \bigcup_0^\infty G_i$, note it is countable since the metric space is separable. Moreover, $\forall B \in \mathcal{F}, \exists C \in \mathcal{G}$ s.t. $B \subseteq C$(pick the minimal $C \cap B \neq \emptyset$) □

Definition 5.6.3

measurable $f : \mathbb{R}^n \to \mathbb{C}$ is *locally integrable* (wrt. m) if $\int_K |f(x)|dx < \infty$ for every bounded measurable set $K \subseteq \mathbb{R}^n$, denote the space of locally integrable functions by L^1_{loc}. If $f \in L^1_{loc}, x \in \mathbb{R}^n, r > 0$, define $A_r f(x) = \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y)dy$

Theorem 5.6.11

$\mu(X) < \infty$, $f \in L^1(\mu)$, S is a closed set in $\mathbb{C}$ and $A_E(f) = \frac{1}{\mu(E)} \int_E f d\mu \in S$ for every $E \in \mathcal{M}$ with $\mu(E) > 0$, then $f(x) \in S$ a.e. $x \in X$

Proof. Pick $\Delta = \overline{B_r(\alpha)} \subseteq S^c$, it suffices to prove $\mu(E) = 0$ where $E = f^{-1}(\Delta)$ If $\mu(E) > 0$, then $|A_E(f) - \alpha| \leq \frac{1}{\mu(E)} \int_E |f - \alpha| d\mu \leq r$ □

Lemma 5.6.12

If $f \in L^1_{loc}$, $A_r f(x)$ is jointly continuous in r and x

Proof. Note $m(B(x,r)) = cr^n$ where $c = m(B(0,1))$ and $m(S(x,r)) = 0$ where $S(x,r) = \{y : |y - x| = r\}$. $\chi_{B(x,r)} \to \chi_{B(x_0,r_0)}$ pointwise on $\mathbb{R}^n \setminus S(x_0, r_0)$, as $r \to r_0$ and $x \to x_0$. Hence $\chi_{B(x,r)} \to \chi_{B(x_0,r_0)}$ a.e., and $\chi_{B(x,r)} \leq \chi_{B(x_0,r_0+1)}$ if $r < r_0 + \frac{1}{2}$ and $|x - x_0| < \frac{1}{2}$. By DCT, $\int_{B(x,r)} f(y)dy$ is continuous in r and x hence so is $A_r f(x) = c^{-1} r^{-n} \int_{B(x,r)} f(y)dy$ □

Definition 5.6.4

if $f \in L^1_{loc}$, *Hardy-Littlewood maximal function* $Hf(x) = \sup_{r>0} A_r|f|(x) = \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y)|dy$. Hf is measurable for $(Hf)^{-1}((a,\infty)) = \bigcup_{r>0} (A_r|f|)^{-1}((a,\infty))$

Lemma 5.6.13

$\mathcal{F}$ collection of open balls in $\mathbb{R}^n$, let $U = \bigcup_{B \in \mathcal{F}} B$. If $c < m(U)$, then exist disjoint $B_1, ..., B_k \in \mathcal{F}$ s.t. $\sum_1^k m(B_j) > 3^{-n} c$

Proof. Note exist compact $K \subseteq U$ with $m(K) > c$, and finitely many of balls $A_1, ..., A_m$ cover K. Then by Vitali. □

Theorem 5.6.14 (The Maixmal Theorem)

$\exists C > 0$ s.t. $\forall f \in L^1, \alpha > 0$, $m(\{x : Hf(x) > \alpha\}) \leq \frac{C}{\alpha} \int |f(x)|dx$

Proof. Let $E_\alpha = \{x : Hf(x) > \alpha\}$. For each $x \in E_\alpha$, choose $r_x > 0$ s.t. $A_{r_x}|f|(x) > \alpha$. $B(x, r_x)$ cover E_α, so by lemma above, if $c < m(E_\alpha)$, then exist $x_1, ..., x_k \in E_\alpha$ s.t. $B_j = B(x_j, r_{x_j})$ are disjoint and $\sum_1^k m(B_j) > 3^{-n} c$.

But then $c < 3^n \sum_1^k m(B_j) \leq \frac{3^n}{\alpha} \sum_1^k \int_{B_j} |f(y)|dy \leq \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f(y)|dy$, let $c \to m(E_\alpha)$ □

notion: $\limsup_{r\to R} f(r) = \lim_{\epsilon\to 0} \sup_{0<|r-R|<\epsilon} f(r) = \inf_{\epsilon\to 0} \sup_{0<|r-R|<\epsilon} f(r)$, note $\lim_{r\to R} f(r) = c$ iff $\limsup_{r\to R} |f(r) - c| = 0$

Theorem 5.6.15

If $f \in L^1_{loc}$, then $\lim_{r\to 0} A_r f(x) = f(x)$, a.e. $x \in \mathbb{R}^n$ (equivalently, $\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|dy = 0$ a.e. x)

Proof. Suffice to show for $N \in \mathbb{N}$, $A_r f(x) \to f(x)$ a.e. x with $|x| \leq N$. But for $|x| \leq N$, $r \leq 1$, $A_r f(x)$ depend only on the value $f(y)$ for $|y| \leq N+1$, so by replacing f with $f\chi_{B(0,N+1)}$, assume $f \in L^1$. Given $\epsilon > 0$, exist a continuous integrable function g s.t. $\int |g(y) - f(y)|dy < \epsilon$. Note $(\forall x \in \mathbb{R}^n)(\forall \delta > 0)(\exists r > 0)(\forall |y - x| < r)|g(y) - g(x)| < \delta$. Hence $|A_r g(x) - g(x)| = \frac{1}{m(B(x,r))} |\int_{B(x,r)} (g(y) - g(x))dy| < \delta$. Therefore $A_r g(x) \to g(x)$, as $r \to 0$ for every x, so $\limsup_{r\to 0} |A_r f(x) - f(x)| = \limsup_{r\to 0} |A_r(f-g)(x) + (A_r g - g)(x) + (g-f)(x)| \leq H(f-g)(x) + 0 + |f-g|(x)$. Hence if $E_\alpha = \{x : \limsup_{r\to 0} |A_r f(x) - f(x)| > \alpha\}, F_\alpha = \{x : |f-g|(x) > \alpha\}$, then $E_\alpha \subseteq F_{\frac{\alpha}{2}} \cup \{x : H(f-g)(x) > \frac{\alpha}{2}\}$. But $\frac{\alpha}{2} m(F_{\frac{\alpha}{2}}) \leq \int_{F_{\frac{\alpha}{2}}} |f(x) - g(x)|dx < \epsilon$. By maximal theorem, $m(E_\alpha) \leq \frac{2\epsilon}{\alpha} + \frac{2C\epsilon}{\alpha}$. Since ϵ if arbitrary, $m(E_\alpha) = 0$ for all $\alpha > 0$. But $\lim_{r\to 0} A_r f(x) = f(x)$ for all $x \notin \bigcup_1^\infty E_{\frac{1}{n}}$. □

Actually, something stronger is true. *Lebesgue set* of f is $L_f = \{x : \lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|dy = 0\}$

Theorem 5.6.16

If $f \in L^1_{loc}$, then $m((L_f)^c) = 0$

Proof. Apply above to $g_c(x) = |f(x) - c|$, get $\lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - c|dy = |f(x) - c|$. Let D be a countable dense subset of $\mathbb{C}$, let $E = \bigcup_{c\in D} E_c$, then $m(E) = 0$. If $x \notin E$, $\forall \epsilon > 0$, choose $c \in D$ with $|f(x) - c| < \epsilon$ s.t. $|f(y) - f(x)| < |f(y) - c| + \epsilon$. Then $\limsup_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|dy \leq |f(x) - c| + \epsilon < 2\epsilon$. □

Definition 5.6.5

A family $\{E_r\_r > 0\}$ of Borel subsets of $\mathbb{R}^n$ is *shrink nicely* to $x \in \mathbb{R}^n$ if $(\forall r) E_r \subseteq B(x,r)$ and $(\exists \alpha)(\forall r) m(E_r) > \alpha m(B(x,r))$

A Borel measure ν on $\mathbb{R}^n$ is *regular* iff $\nu(K) < \infty$ for any compact K and $\nu(E) = \inf\{\nu(U) : E \subseteq U \textit{ open}\}$ for every $E \in \mathcal{B}_{\mathbb{R}^n}$ (Actually the latter condition implies the former; note the former implies σ-finite) A signed or complex Borel measure ν is regular if $|\nu|$ regular E.g., if $f \in L^+(\mathbb{R}^n)$, then fdm regular iff $f \in L^1_{loc}$

Theorem 5.6.17 (Lebesgue Differentiation Theorem)

$f \in L^1_{loc}$, for every $x \in L_f$ (in particular for a.e. x), $\lim_{r\to 0} \frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|dy = 0$ and $\lim_{r\to 0} \frac{1}{m(E_r)} \int_{E_r} f(y)dy = f(x)$ for every family $\{E_r\}_{r>0}$ that shrinks nicely to x.

Proof. For some $\alpha > 0$, $\frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|dy \leq \frac{1}{\alpha m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|dy$ □

Theorem 5.6.18

ν regualr signed or complex Borel measure on $\mathbb{R}^n$, $d\nu = d\lambda + f dm$ be its Lebesgue-Radon-Nikodym representation. Then for m-a.e. $x \in \mathbb{R}^n$, $\lim_{r\to 0} \frac{\nu(E_r)}{m(E_r)} = f(x)$ for every family $\{E_r\}_{r>0}$ that shrinks nicely to x

Proof. Note $d|\nu| = d|\lambda| + |f|d\mu$, so the regularity of ν implies the regularity of λ and fdm. Get $f \in L^1_{loc}$, by Lebesgue Differentiation Theorem, it suffices to show if λ is regular and $\lambda \perp m$, then for m-a.e. x, $\frac{\lambda(E_r)}{m(E_r)} \to 0$ as $r \to 0$ when E_r shrinks nicely to x. It also suffices to take $E_r = B(x,r)$ and assume λ is positive since for some $\alpha > 0$, we have $|\frac{\lambda(E_r)}{m(E_r)}| \le \frac{|\lambda|(E_r)}{m(E_r)} \le \frac{|\lambda|(B(x,r))}{m(E_r)} \le \frac{|\lambda|(B(x,r))}{\alpha m(B(x,r))}$. Assume $\lambda \ge 0$, then let A be a Borel set s.t. $\lambda(A) = m(A^c) = 0$, let $F_k = \{x \in A : \limsup_{r\to 0} \frac{\lambda(B(x,r))}{m(B(x,r))} > \frac{1}{k}\}$, it suffice to show $(\forall k) m(F_k) = 0$. Similar to the proof of maximal theorem, given $\epsilon > 0$, $\exists A \subseteq U_\epsilon$ open s.t. $\lambda(U_\epsilon) < \epsilon$. Each $x \in F_k$ is the center of a ball $B_x \subseteq U_\epsilon$ s.t. $\lambda(B_x) > k^{-1} m(B_x)$. By Lemma, if $V_\epsilon = \bigcup_{x\in F_k} B_x$ and $c < m(V_\epsilon)$, there exist $x_1, ..., x_J$ s.t. $B_{x_1}, ..., B_{x_J}$ are disjoint and $c < 3^n \sum_1^J m(B_{x_j}) \le 3^n k \sum_1^J \lambda(B_{x_j}) \le 3^n k \lambda(V_\epsilon) \le 3^n k \lambda(U_\epsilon) \le 3^n k\epsilon$, get $m(V_\epsilon) \le 3^n k\epsilon$. Since $F_k \subseteq V_\epsilon$, so $m(F_k) = 0$ □

Theorem 5.6.19

weak-1(can't use 1 to control)-1 Estimate

6 Algebraic Topology

6.1 Homology

Fundamental groups, covering spaces, higher homotopy groups, fibrations and the long exact sequence of a fibration

singular homology and cohomology, relative homology, CW complexes and the homology of CW complexes

Mayer-Vietoris sequence, universal coefficient theorem, Kunneth formula, Poincare duality, Lefschetz fixed point formula, Hopf index theorem, Cech cohomology adn deRham cohomology, equivalence between singular, Cech and de Rham cohomology

6.2 Homotopy

7 Functional Analysis

7.1 Functional Analysis

Hilbert space, Hahn-Banah Theorem, open mapping theorem, uniform boundedness theorem, closed graph theorem

Basic properties of compact operators, Riesz-Fredholm Theory, spectrum of compact operators

Fourier series, Fourier transform, convolution

7.2 Harmonic Analysis

8 Differential Manifolds

8.1 Manifold

Smooth manifold, inverse function theorem, implicit function theorem, submanifolds, Sard's Theorem, embedding theorem, transversality, degree theory, integration on manifolds

real and comlex vector bundles, tangent and cotangent bundles, basic operations on bundles such as dual bundle, tensor products, exterior products, direct sums, pull-back bundles

differential forms, exterior product, exterior derivative, deRham cohomology, behavior under pull-back

Matrices on vector bundles

Riemann metrics, geodesic, existence and uniqueness of geodesics

associated vector bundles: relation between principal bundles and vector bundles covariant derivative for a vector bundle and connection on a principal bundle, and their relation

curvature, flat connection, parallel transport

Levi-Civita connection and properties of the Riemann curvature tensor, manifolds of constant curvature

Jacobi fields, second variation of geodesics

Manifolds of nonpositive curvature, manifolds of positive curvature

8.2 Lie Group and Lie Algebra

Basics of matrix Lie groups over $\mathbb{R}$ and $\mathbb{C}$: definition of $Gl(n), SU(n), SO(n), U(n)$, their manifold structures, Lie algebras, right and left invariant vector fields and differential forms, the exponential map

principal Lie group bundle for matrix groups

8.3 symplectic geometry

8.4 Riemann Geometry

9 Algebraic Geometry

9.1 Algebraic Curves and Surfaces

9.2 Prerequisite

Algebraic variety

9.3 Birational Geometry

9.4 Hodge Theory

9.5 Moduli Space

10 Complex Geometry

11 Dynamical System

11.1 Differential Equation

Existence and uniqueness theorems for solutions of ODE; explicit solutions of simple equations; self-adjoint boundary value problems on finite intervals; critical points, phase space, stability analysis

First order partial differential equations, linear and quasi-linear PDE

Phase plane analysis, Burgers equation, Hamilton-Jacobi equation

Potential equations: Green functions and existence of solutions of Dirichlet problem, harmonic functions, maximal principal and applications, existence of solutions of Neumann's problem

Heat equation, Dirichlet problem, fundamnetal solutions

Wave equations: initial condition and boundary condition, well-posedness, Sturm-Liouville eigenvalue problem, energy functional method, uniqueness and stability of solutions

Distributions, Sobolev embedding theorem

11.2 Ergodic Theory

11.3 Stability, Control and Chaos Theory

12 Number Theory

12.1 Analytic Number Theory

12.2 Algebraic Number Theory

12.3 Arithmetic Geometry

13 Probability Theory

13.1 Random Variable

Sample space, Probability space Random Variables (discrete, constant, multivariate; independent, identically-distributed, uncorrelated) Probability Distribution (continuous, cumulative, discrete, joint; normal/Guassian, binomial, bernoulli, exponential) Probability density function Probability Distribution function Probability mass function characteristic function, generating function, various modes of convergence of random variable

Moment Expectation, Expected value, variance, central moment, factorial moment, coefficient of variation, correlation, covariance, cumulance

Conditioning Bayes' Theorem, conditional expectation given a sigma field, prior Probability

Limit theorem law of large numbers central limit theorem large deviations theory law of total covariance/cumulance/expectatio

13.2 Stochastic Process

Markov chain, Guass-Markov process, random graph, random matrix, Stochastic calculus, Martingales, Basic properties of Possion processes, basic properties of Brownian motion

13.3 Distribution Theory

families of continuous distributions: normal, chi-sq, t, F, gamma, beta; families of discrete distributions: multinomial, Possion, negative binomial;

13.4 Statistics

Basic statistics: sample mean, variance, median and quantiles

Testing: Neyman-Pearson paradigm, null and alternative hypotheses, simple and composite hypotheses, type I and II errors, power, most powerful test, likelihood radio test, Neyman-Pearson Theorem, generalized likelihood ratio test

Estimation: parameter estimation, method of moments, maximum likelihood estimation, criteria for evaluation of estimators, Fisher information and its use, confidence interval

Bayesian Statistics: Prior, posterior, conjugate priors, Bayesian estimators

Large sample properties: consistency, asymptotic normality, chi-sq approximation to likelihood radio statistics

MLE: maximum likelihood estimate, MAP: maximum a posteriori, linear regression Bayesian estimation, conjugate priordistribution, posterior probability latent variable, EM: expectation maximization algorithm, mixture model MC: Markov Chain, Markov process Monte Carlo method, MCMC: Markov Chain Monte Carlo, Gibbs sampling

14 Combinatorics Theory

14.1 Graph Theory

Algebraic graph theory, Ramsey theory, Van der Waerden's theorem, Hales-Jewett theorem, Umbral calculus, binomial type polynomial sequences

14.2 Matroid Theory

14.3 Enumerative Combinatorics

14.4 Algebraic Combinatorics

14.5 Geometric Combinatorics

14.6 Analytic Combinatorics

15 Computation Theory

15.1 Interpolation and Approximation

Trigonometric interpolation and approximation, fast Fourier transform; approximation by rational function; polynomial and spline interpolations and approximation; least-squares approximation

15.2 Numerical Solution of Differential Equation

ODE: Single step methods and multi-step methods, stability, accuracy and convergence; absolute stability, long time behavior; numerical methods for stiff ODE'S PDE: finite difference method, finite element method and spectral method; stability, accuracy and convergence, Lax equivalence theorem

15.3 Linear and Nonlinear Programming

Linear Systems and Eigenvalue Problems: Classical and modern iterative method for linear systems and eigenvalue problem, condition number and singular value decomposition, iterative methods for large sparse system of linear equations

Nonlinear Equation Solvers: Convergence of iterative methods (bisection, Newton's Method, quasi-Newton's methods and fixed-point methods) for both scalar equations and systems, finding roots of polynomial

Simplex method, interior method, penalty method, Newton's method, homotopy method and fixed point method, dynamic programming

15.4 Mathematical Modeling, Simulation, and Applied Analysis

Scaling behavior and asymptotic analysis, stationary phase analysis, boundary layer analysis, qualitative and quantitative analysis of mathematical models, Monte-Carlo method

15.5 Computability and Complexity

16 Mathematical Physics

16.1 Classical Mechanics

Lagrangian formalism: inciple of least action, Euler-Lagrangian equation, Noether Theorem, Kepler problem, rigid body

Hamiltonian formalism: Hamilton's equation, Poisson bracket, Liouville's Theorem, canonical transformation, Hamilton-Jacobi theory

16.2 Electrodynamics

Electrostatics and magnetostatics: fields, potentials, charges, electric and magnetic fields in matter

Electrodynamics: Coulomb's law, Lorentz force law, Ohm's law, Faraday's law, Guass's law, Maxwell's equation, conservation laws, electromagnetic waves, radiation

Basic Methods: the method of images, separation of variables, multipole expansion

16.3 Thermodynamics and Statistical Physics

Fundamental principal of thermodynamics, thermodynamic potentials and process, phase equilibrium and phase transitions, partition function, entropy

Probability theory, the microcanonical, canonical and grand-canonical ensembles, The Boltzmann, Bose and Fermi Statistical distributions

Examples: ideal gas model, paramagnet, ideal quantum gases, degenerate Fermi systems; photons and phonons; Bose-Einstein condensation

16.4 Quantum Mechanics

Fundamental concepts: Hilbert space, states, observables, wave functions, Schrodinger equation, Schrodinger and Heisenberg pictures, canonical quantization, density matrix

Examples: harmonic oscillator, hydrogen atom model, potential well problems

Symmetry in quantum mechanics, angular momentum, spin, identical particles, and atomic structure

Perturbation theory, scattering, approximation method

16.5 General Relativity

Differential geometry: metric, vector, tensor, differential forms, manifold, connections, curvature, geodesic, tetrads, Lie derivative, isometries and Killing vector

Gravitation: the principle of equivalence, Einstein's equation, Hilbert-Einstein action

Exact solution: Minkowski, de Sitter, anti-de Sitter spacetimes, and black hole solution

Causal structure

16.6 Quantum Field Theory

Classical field theory: Lagrangian and Hamiltonian formalism, Noether theorem

Quantization: canonical quantization and path integrals

Fermions: representations of Poincare group, Dirac equation

S-matrix: LSZ reduction, Feymann propagator, Feymann rules, normal ordering

Wick's theorem, the optical theorem, locality

Renormalization: regularization and cutoff, counter terms, renormalization group