Bourbaki

Contents

1 Objects

1.1 Why

We want to talk about things.

1.2 Definition

We use the word **object** with its usual sense in the English language. An object may be tangible, in that we can hold or touch it, or an object may be abstract, in that we can do neither.

1.2.1 Notation

To aid in discussing and denoting objects, let us tend to give them short names. A single Latin letter regularly suffices: for example, a, b or c. To aid our memory, we tend to choose the letter mnemonically.

We denote that the object a and the object b are the same object by $a = b$, read aloud as "a is b." We denote that the object a and b are different by $a \neq b$, read aloud as "a is not b."

2 Sets

2.1 Why

We want to talk about none, one, or several objects considered as a whole, for which we will use the word *set*.

2.2 Definition

A **set** is an abstract object which we think of as several objects considered at once. We say that the set **contains** the objects so considered. We call these the **elements** of the set.

We call the set which contains no objects the **empty set**. We call a set which contains only a single object a **singleton**. A singleton is not the same as the object it contains. Besides these two cases, we think of sets as containing two or more objects.

The objects a set contains may be other sets. This may be subtle at first glance, but becomes familiar with experience.

2.2.1 Notation

Let us tend to denote sets by upper case Latin letters: for example, A, B, and C. To aid our memory, let us tend to use the lower case form of the letter for an element of the set. For example, let A and B be non-empty sets. Let us tend to denote by a an element of A, and likewise, by b an element of B

Let us denote that an object a is an element of a set A by $a \in A$. We read the notation $a \in A$ aloud as "a in A." The $\in$ is a stylized lower case Greek letter: ϵ. It is read aloud "ehp-sih-lawn" and is a mnemonic for "element of". We write $a \notin A$, read aloud as "a not in A," if a is not an element of A.

If we have named the elements of a set, and can list them, let us do so between braces. For example, let a, b, and c be three distinct objects. Denote by $\{a, b, c\}$ the set containing theses three objects and only these three objects. We can further compress notation, and denote this set of three objects by A: so, $A = \{a, b, c\}$. Then $a \in A$, $b \in A$, and $c \in A$. Moreover, if d is an object and $d \in A$, then $d = a$ or $d = b$ or $d = c$.

We denote the empty set by $\varnothing$. Note that $\varnothing \neq \{\varnothing\}$. The left hand side, $\varnothing$, is the empty set. The right hand side, $\{\varnothing\}$, is the singleton whose element is the empty set. We distinguish the set containing one element from the element itself.

3 Set Examples

3.1 Why

We give some examples of objects and sets.

3.2 Examples

For familiar examples, let us start with some tangible objects. Find, or call to mind, a deck of playing cards.

First, consider the set of all the cards. This set contains fifty-two elements. Second, consider the set of cards whose suit is hearts. This set contains thirteen elements: the ace, two, three, four, five, six, seven, eight, nine, ten, jack, queen, and king of hearts. Third, consider the set of twos. This set contains four elements: the two of clubs, the two of spades, the two of hearts, and the two of diamonds.

We can imagine many more sets of cards. If we are holding a deck, each of these can be made tangible: we can touch the elements of the set. But the set itself is always abstract: we can not touch it. It is the idea of the group as distinct from any individual member.

Moreover, the elements of a set need not be tangible. First, consider the set consisting of the suits of the playing card: hearts, diamonds, spades, and clubs. This set has four elements. Each element is a suit.

Second, consider the set consisting of the card types. This set has thirteen elements: ace, two, three, four, five, six, seven, eight, nine, ten, jack, queen, king. The subtlety here is that this set is different than the set of hearts, namely those thirteen cards which are hearts. However these sets are similar: they both have thirteen elements, and there is a natural correspondence between their elements: the ace of hearts with the type ace, the two of hearts with the type two, and so on.

Of course, sets need have nothing to do with playing cards. For example, consider the set of seasons: autumn, winter, spring, and summer. This set has four elements. For another example, consider the set of Latin letters: a, b, c, . . . , x, y, z. This set has twenty-six elements.

4 Set Extension

4.1 Why

When are two sets the same?

4.2 Definition

Two sets are **equal** if and only if they have the same elements.

4.2.1 Notation

Let A and B be two sets. As with any objects, we denote that A and B are equal by $A = B$.

5 Subsets

5.1 Why

We want to speak of sets which contain all the elements of other sets.

5.2 Two Sets

A **subset** of a set A is any set B for which each element of the set B is an element of the set A. In this case, we say that B is a subset of A. Conversely, we say that A is a **superset** of B.

Every set is a subset of itself. So if the set A is the set B, then A is a subset of B and B is a subset of A. Conversely, if A is a subset of B and B is a subset of A, then A is B. To argue that A is B, we argue that membership in A implies membership in B and second, we argue that membership in B implies membership in A.

The **power set** of a set is the set of all subsets of that set. It includes the set itself and the empty set. We call these two sets **improper subsets** of the set. We call all other sets **proper subsets**.

5.2.1 Notation

Let A and B be sets. We denote that A is a subset of B by $A \subset B$. We read the notation $A \subset B$ aloud as "A subset B".

If $A \subset B$ and $B \subset A$, then $A = B$. The converse also holds.

We denote the power set of A by 2^A, read aloud as "two to the A." $A \in 2^A$ and $\varnothing \in 2^A$. However, $A \subset 2^A$ is false.

5.2.2 Examples

Let a, b, c be distinct objects. Let $A = \{a, b, c\}$ and $B = \{a, b\}$. Then $B \subset A$. In other notation, $B \in 2^A$. As always, $\varnothing \in 2^A$ and $A \in 2^A$ as well. In this case, we can list the elements (which are sets) of the power set:

$$2^A = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}.$$

6 Set-Builder Notation

6.1 Why

We specify a subset of a known set via a property.

6.2 Definition

Let A be a nonempty set. We use the notation

$$\{a \in A \mid \text{——} \}$$

to indicate a subset of A that satisfies some property specified after the $|$. We read the symbol $|$ aloud as "such that." We read the whole notation aloud as "a in A such that..."

We call the notation **set-builder notation**. Set-builder notation avoids enumerating elements.

6.3 Example

For example, let L be the set of Latin letters and V the set of Latin vowels. A first notation for V is $\{\mathrm{a}, \mathrm{e}, \mathrm{i}, \mathrm{o}, \mathrm{u}\}$. A second notation for V is $\{l \in L \mid l \text{ is a vowel}\}$. We may prefer the second, in cases when it saves time. This notation is really indispensable for sets which have many members, too many to reasonably write down.

7 Ordered Pairs

7.1 Why

We speak of an ordered pair of objects: one selected from a first set and one selected from a second set.

7.2 Definition

Let A and B be non-empty sets. Let $a \in A$ and $b \in B$. The **ordered pair** of a and b is the set $\{\{a\}, \{a, b\}\}$. The **first element** of $\{\{a\}, \{a, b\}\}$ is a and the **second element** is b.

The **cartesian product** of A and B is the set of all ordered pairs. If $A \neq B$, the ordering causes the cartesian product of A and B to differ from the cartesian product of B with A. If $A = B$, however, the symmetry holds.

7.2.1 Notation

We denote the ordered pair $\{\{a\}, \{a, b\}\}$ by (a, b). We denote the cartesian product of A with B by $A \times B$, read aloud as "A cross B." In this notation, if $A \neq B$, then $A \times B \neq B \times A$.

8 Relations

8.1 Why

How can we relate the elements of two sets?

8.2 Definition

A **relation** between two non-empty sets A and B is a subset of $A \times B$. A relation on a single set C is a subset of $C \times C$.

Let $a \in A$ and $b \in B$. The pair (a, b) may or may not be in a relation on A and B. If $A \neq B$, then (b, a) is not a member of the product $A \times B$, and therefore not in any relation on A and B. If $A = B$, however, it may be that (b, a) is in the relation.

8.2.1 Notation

Let A and B be nonempty sets with $a \in A$ and $b \in B$. Since relations are sets, we can use upper case Latin letters. Let R be a relation on A and B. We denote that $(a, b) \in R$ by aRb, read aloud as "a in relation R to b."

When $A = B$, we tend to use other symbols instead of letters. For example, $\sim$, $=$, $<$, $\leq$, $\prec$, and $\preceq$.

9 Relation Properties

9.1 Why

Often relations are defined over a single set. These have a few simple properties.

9.2 Definitions

- A relation is **reflexive** if every element is related to itself.
- A relation is **symmetric** if two objects are related regardless of their order.
- A relation is **antisymmetric** if two objects are related only in one order, and never both.
- A relation is **transitive** if a first element is related to a second element and the second element is related to the third element, then the first and third element are related

9.2.1 Notation

Let R be a relation on a non-empty set A. R is **reflexive** if

$$(a, a) \in R$$

for all $a \in A$. R is **transitive** if

$$(a, b) \in R \wedge (b, c) \in R \implies (a, c) \in R$$

for all $a, b, c \in A$. R is **symmetric** if

$$(a, b) \in R \implies (b, a) \in R$$

for all $a, b \in A$. R is **anti-symmetric** if

$$(a, b) \in R \implies (b, a) \notin R$$

for all $a, b \in A$.

10 Functions

10.1 Why

We want a notion for a correspondence between two sets.

10.2 Definition

A **functional** relation on two sets relates each element of the first set with a unique element of the second set. A **function** is a functional relation.

The **domain** of the function is the first set and **codomain** of the function is the second set. The function **maps** elements **from** the domain **to** the codomain. We call the codomain element associated with the domain element the **result** of **applying** the function to the domain element.

10.2.1 Notation

Let A and B be sets. If A is the domain and B the codomain, we denote the set of functions from A to B by $A \to B$, read aloud as “A to B”.

We denote functions by lower case latin letters, especially

f, g, and h. The letter f is a mnemonic for function; g and h follow f in the Latin alphabet. We denote that $f \in A \to B$ by $f : A \to B$, read aloud as “f from A to B”.

Let $f : A \to B$. For each element $a \in A$, we denote the result of applying f to a by $f(a)$, read aloud “f of a.” We sometimes drop the parentheses, and write the result as f_a, read aloud as “f sub a.”

Let $g : A \times B \to C$. We often write $g(a, b)$ or g_{ab} instead of $g((a, b))$. We read $g(a, b)$ aloud as “g of a and b”. We read g_{ab} aloud as “g sub a b.”

11 Identity Functions

11.1 Why

What is an example of a function?

11.2 Definition

Consider the relation on a non-empy set which includes ordered pairs of elements from the set if the first element is the second element. In other words, two elements are related if and only if they are the same object.

This relation is functional: Each element corresponds to only one element: itself. We call this functional relation the **identity function** of the set.

11.2.1 Notation

Let A be a non-empty set. If $f : A \to A$ satisfies $f(a) = a$ for each $a \in A$, then f is the identity function.

We denote the identity function on A by $\mathbf{id}_A$. So $\mathbf{id}_A : A \to A$ satisfies $\mathbf{id}_A(a) = a$ for each $a \in A$.

12 Function Properties

12.1 Why

TODO

12.2 Definition

Let $f : A \to B$. The **image** of a set $C \subset A$ is the set $\{f(c) \in B \mid c \in C\}$. The **range** of f is the image of the domain. The **inverse image** of a set $D \subset B$ is the set $\{a \in A \mid f(a) \in B\}$.

The range need not equal the codomain; though it, like every other image, is a subset of the codomain. The function maps to domain **on** to the codomain if the range and codomain are equal; in this case we call the function **onto**. This language suggests that every element of the codomain is used by f. It means that for each element b of the codomain, we can find an element a of the domain so that $f(a) = b$.

An element of the codomain may be the result of several elements of the domain. This overlapping, using an element of the codomain more than once, is a regular occurrence. If a function is a unique correspondence in that every domain element has a different result, we call it **one-to-one**. This language is meant to suggest that each element of the domain corresponds to one and exactly one element of the codomain, and vice versa.

12.2.1 Notation

Let $f : A \to B$. We denote the image of $C \subset A$ by $f(C)$, read aloud as "f of C." This notation is overloaded: for $c \in C$, $f(c) \in A$, whereas $f(C) \subset A$. Read aloud, the two are indistinguishable, so we must be careful to specify whether we mean an element c or a set C. The property that f is onto can be written succintly as $f(A) = B$. We denote the inverse image of $D \subset B$ by $f^{-1}(D)$, read aloud as "f inverse D."

13 Natural Numbers

13.1 Why

We want to count, forever.

13.2 Definition

We define the set of **natural numbers** implicitly. There is an element of the set which we call **one**. For each element of the set, there is a unique corresponding element, called the **successor** of the former, which is also in the set. We call the elements **numbers** and call the set itself the **naturals**.

The **successor function** is the correspondence between elements of the natural numbers and their successors. Its domain and codomain is the set of natural numbers. It is a one-to-one correspondence. It is not onto, however, since the element one has no successor.

To recap, we begin with a distinguished element, called one, in the set. There is a second element, the successor of one, in the set. We call the successor of one **two**. We call the successor of two **three**. And so on using the English language in the usual manner. We are saying, in the language of sets, that the essence of counting is starting with one and adding one repeatedly.

13.2.1 Notation

We denote the set of natural numbers by N, a mnemonic for natural. We often denote elements of N by n, a mnemonic for number, or m, a letter close to n. We denote the element called one by 1.

14 Natural Induction

14.1 Why

We want to count, forever.

14.2 Definition

We assert two additonal self-evident and indispensable properties of these natural numbers. First, one is the successor of no other element. Second, if we have a subset of the naturals containing one with the property that it contains successors of its elements, then that set is equal to the natural numbers. We call this second property the **principle of mathematical induction.**

These two properties, along with the existence and uniqueness of successors are together called **Peano's axioms** for the natural numbers. When in familiar company, we freely assume Peano's axioms.

14.3 Notation

As an exercise in the notation assumed so far, we can write Peano's axioms: N is a set along with a function $s : N \to N$

such that

1. $s(n)$ is the successor of n for all $n \in N$.
2. s is one-to-one; $s(n) = s(m) \implies m = n$ for all $m, n \in N$.
3. There does not exist $n \in N$ such that $s(n) = 1$.
4. If $T \subset N$, $1 \in T$, and $s(n) \in T$ for all $n \in T$, then $T = N$.

15 Function Composition

15.1 Why

We want a notion for applying two functions one after the other. We apply a first function then a second function.

15.2 Definition

Consider two functions for which the codomain of the first function is the domain of the second function.

The **composition** of the second function with the first function is the function which associates each element in the first's domain with the element in the second's codomain that the second function associates with the result of the first function.

The idea is that we take an element in the first domain. We apply the first function to it. We obtain an element in the first's codomain. This result is an element of the second's domain. We apply the second function to this result. We obtain an element in the second's codomain. The composition of the second function with the first is the function so constructed.

15.2.1 Notation

Let A, B, C be non-empty sets. Let $f : A \to B$ and $g : B \to C$. We denote the composition of g with f by $g \circ f$ read aloud as "g composed with f." To make clear the domain and comdomain, we denote the composition $g \circ f : A \to C$.

In previously introduced notation, $g \circ f$ satisfies

$$(g \circ f)(a) = g(f(a))$$

for all $a \in A$.

16 Function Inverses

16.1 Why

We want a notion of reversing functions.

16.2 Definition

An **identity function** is a relation on a set which is functional and reflexive. It associates each element in the set with itself. There is only one identity function associated to each set.

Consider two functions for which the codomain of the first function is the domain of the second function and the codomain of the second function is the domain of the first function. These functions are **inverse functions** if the composition of the second with the first is the identity function on the first's domain and the composition of the first with the second is the identity function on the second's domain.

In this case we say that the second function is an **inverse** of the second, and vice versa. When an inverse exists, it is unique, so we refer to the **inverse** of a function.

16.2.1 Notation

Let A a non-empty set. We denote the identity function on A by id_A, read aloud as "identity on A." id_A maps A onto A.

Let A, B be non-empty sets. Let $f : A \to B$ and $g : B \to A$ be functions. f and g are inverse functions if $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.

16.3 The Inverse

We discuss existence and uniqueness of an inverse.

Proposition 1. *Let $f : A \to B$, $g : B \to A$, and $h : B \to A$.*

If g and h are both inverse functions of f, then $g = h$.

Proof. □

Proposition 2. *If a function is one-to-one and onto, it has an inverse.*

Proof. □

17 Order Relations

17.1 Why

We want to handle elements of a set in a particular order.

17.2 Definition

Let R be a relation on a non-empty set A. R is a **partial order** if it is reflexive, transitive, and anti-symmetric.

A **partially ordered set** is a set and a partial order. The language partial is meant to suggest that two elements need not be comparable. For example, suppose R is $\{(a, a) \mid a \in A\}$; we may justifiably call this no order at all and call A totally unordered, but it is a partial order by our definition.

Often we want all elements of the set A to be comparable. We call R **connexive** if for all $a, b \in A$, $(a, b) \in R$ or $(b, a) \in R$. If R is a partial order and connexive, we call it a **total order**.

A **totally ordered set** is a set together with a total order. The language is a faithful guide: we can compare any two elements. Still, we prefer one word to three, and so we will use the shorter term **chain** for a totally ordered set; other terms include **simply ordered set** and **linearly ordered set**.

17.2.1 Notation

We denote total and partial orders on a set A by $\preceq$. We read $\preceq$ aloud as "precedes or equal to" and so read $a \preceq b$ aloud as "a precedes or is equal to b." If $a \preceq b$ but $a \neq b$, we write $a \prec b$, read aloud as "a precedes b."

18 Operations

18.1 Why

We want to “combine” elements of a set.

18.2 Definition

Let A be a non-empty set. An **operation** on A is a function from ordered pairs of elements in the set to the same set. We use operations to combine the elements. We operate on pairs.

18.2.1 Notation

Let A be a set and $g : A \times A \to A$. We tend to forego the notation $g(a, b)$ and write $a\,g\,b$ instead. We call this **infix notation**.

Using lower case latin letters for elements and for operators confuses, so we tend to use special symbols for operations. For example, $+$, $-$, $\cdot$, $\circ$, and $\star$.

Let A be a non-empty set and $+ : A \times A \to A$ be an operation on A. According to the above paragraph, we tend to write $a + b$ for the result of applying $+$ to (a, b).

19 Commutative Operations

19.1 Why

We introduce language for the case in which an operation does not depend on the order in which it operates.

19.2 Definition

An operation **commutes** if the result of two elements is the same regardless of their order.

19.2.1 Notation

Let A be a non-empty set and let $+ : A \times A \to A$ be an operation. If $+$ commutes, then

$$a + b = b + a$$

for all $a, b \in A$.

20 Associative Operations

20.1 Why

We name operations for which we the first two or the second two elements in an ordered list of three elements.

20.2 Definition

An operation **associates** if given any three elements in order it doesn't matter whether we first operate on the first two and then with the result of the first two the third, or the second two and with the result of the second two the first.

20.2.1 Notation

Let A be a non-empty set and let $+ : A \times A \to A$ be an operation. If $+$ associates, then

$$(a + b) + c = (a + b) + c$$

for all $a, b, c \in A$.

21 Algebras

21.1 Why

We name a set together with an operation.

21.2 Definition

An **algebra** is an ordered pair whose first element is a non-empty set and whose second element is an operation on that set. The **ground set** of the algebra is the set on which the operation is defined.

21.2.1 Notation

Let A be a non-empty set and let $+ : A \times A \to A$ be an operation on A. As usual, we denote the ordered pair by $(A, +)$.

22 Element Functions

22.1 Why

Take an element of an algebra, and consider the function defined on the ground set which maps elements to the result of the operation applied to the fixed element and the given element.

22.2 Definition

Let $(A, +)$ be an algebra. For each $a \in A$, denote by $+_a : A \to A$ the function defined by

$$+_a(b) = a + b.$$

We call $+_a$ the **left element function** of a.

Similarly, denote by $+^a : A \to A$ the function defined by

$$+^a(b) = b + a.$$

We call $+^a$ the **right element function** of a

23 Identity Elements

23.1 Why

We can construct functions on the ground set of an algebra by fixing an element in the ground set and defining a function which maps elements to the result of the operation applied to the fixed element and the given element.

23.2 Definition

Let $(A, +)$ be an algebra. For each $a \in A$, denote by $+_a : A \to A$ the function defined by

$$+_a(b) = a + b.$$

If $+_a$ is the identity function on A then we call a a **left identity element** of the algebra.

Similarly, denote by $+^a : A \to A$ the function defined by

$$+^a(b) = b + a.$$

If $+^a$ is the identity function on A then we call a a **right identity element**. of the algebra.

An **identity element** of the algbera is an element which is both a left and right identity. If the operation commutes, then a left identity and right identities are the same.

24 Inverse Elements

24.1 Why

Is the inverse of an element function the element function of a different element?

24.2 Definition

TODO

25 Solving Equations

25.1 Why

If I am holding 3 stones in both hands, and I have two of them in my left hand, how many might I have in my right hand? If I am holding fewer than 5 pebbles, how many might I have in my right hand?

25.2 Overview

An **equation** is an assertion that two objects are the same. An indeterminate equation is one in which some value is unknown. For example, I am asserting above that I have three pebbles and that I have two in my left hand. We know that the total nu A solution of the equation is any possible number of pebbles I could be holding in my right hand such that the sum is three.

Discuss solutions in a set. Discuss existence of solution. Discuss uniqueness of solution.

25.2.1 Notation

We can denote the pebble equation as find n so that

$$2 + n = 3$$

26 Graphs

26.1 Why

We want to visualize relations.

26.2 Definition

A **graph** is a set and a relation on the set. The graph is **undirected** if the relation is symmetric; otherwise the graph is **directed**.

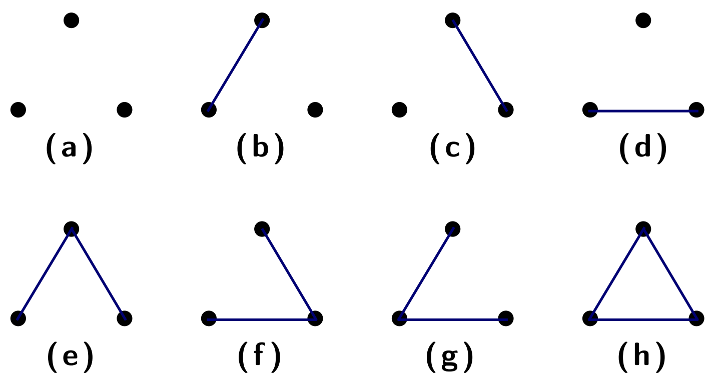
A **vertex** of the graph is an element of the set. The set is called the **vertex set**. An **edge** of the graph is an element of the relation. The relation is called the **edge set**.

26.2.1 Notation

We denote the vertex set by V, a mnemonic for vertex. We denote the edge set by E, a mnemonic for edge. We denote a graph by (V, E). If the vertex set is assumed we can unambiguously refer to the graph by E.

26.2.2 Visualization

We visualize a graph by drawing a point for each vertex. If two vertices u and v are in relation, we draw a line from the point corresponding to u to the point corresponding to v with an arrow at the point corresponding to v. If the graph is undirected, we omit arrows. Here are all undirected graphs on three vertices.



26.3 Paths

A path in a graph is a sequence of vertices with the property that consecutive vertices are related. A path **cycles** if a vertex appears more than once. A path is **finite** if the sequence is finite. A **loop** is a finite path that cycles once. A finite path from vertex u to vertex v is a path starting with u and ending with v. The **length** of a finite path is the length of the sequence.

26.4 Properties

A graph is **connected** if there is a path between every pair of vertices. A graph is **acyclic** if none of its paths cycle.

27 Trees

27.1 Why

Tree branches split and do not recombine. We formalize this property in the language of graphs.

27.2 Definition

A **tree** is a connected acyclic graph.

27.2.1 Notation

Let (V, E) be a tree. When the vertex set is clear from context, we use use T, a mnemonic for "tree," to denote the edge set. We denote the set of trees on the vertex set V by $T(V)$.

27.3 Properties

Proposition 3. *In any tree, there is only one path between any two vertices.*

28 Graph Cliques

28.1 Why

We speak of the complete subgraphs of a graph.

28.2 Definition

A **complete** graph is one for which an edge exists between any two nodes.

A **subgraph** of a given graph is a graph whose vertex set is a subset of the given vertex set and whose edge set is the subset of given edges connecting vertices in the vertex subset. With reference to the underling graph, then, a subgraph can be specified completely by its vertex set.

A **clique** of a given graph is a complete subgraph of that graph. When speaking of the cliques of a given graph, we identify the cliques with their vertex set. The relation contained in gives a partial order on cliques. A clique is **maximal** if it maximal with respect to this relation; i.e., it is contained in no other clique. As a convention, we include $\varnothing$ as a clique.

28.2.1 Notation

Let (V, E) a graph. We denote a clique by $C \subset V$, a mnemonic for clique.

29 Set Operations

29.1 Why

We want to consider the elements of two sets together at once, and other sets created from two sets.

29.2 Definitions

Let A and B be two sets.

The **union** of A with B is the set whose elements are in either A *or* B *or* both. The key word in the definition is *or*.

The **intersection** of A with B is the set whose elements are in both A *and* B. The keyword in the definition is *and*.

Viewed as operations, both union and intersection commute; this property justifies the language "with." The intersection is a subset of A, of B, and of the union of A with B.

The **symmetric difference** of A and B is the set whose elements are in the union but not in the intersection. The symmetric difference commutes because both union and intersection commute; this property justifies the language "and." The symmetric difference is a subset of the union.

Let C be a set containing A. The **complement** of A in C is the symmetric difference of A and C. Since $A \subset C$, the union is C and the intersection is A. So the complement is the "left-over" elements of B after removing the elements of A.

We call these four operations **set-algebraic operations**.

29.2.1 Notation

Let A, B be sets. We denote the union of A with B by $A \cup B$, read aloud as "A union B." $\cup$ is a stylized U. We denote the intersection of A with B by $A \cap B$, read aloud as "A intersect B." We denote the symmetric difference of A and B by $A \Delta B$, read aloud as "A symdiff B." "Delta" is a mnemonic for difference.

Let C be a set containing A. We denote the complement of A in C by $C - A$, read aloud as "C minus A."

29.2.2 Results

Proposition 4. *For all sets A and B the operations $\cup$, $\cap$, and Δ commute.*

Proposition 5. *Let S a set. For all sets $A, B \subset S$,*

(1) $S - (A \cup B) = (S - A) \cap (S - B)$

(2) $S - (A \cap B) = (S - A) \cup (S - B)$.

Proposition 6. *Let S a set. For all sets $A, B \subset S$,*

$$A \Delta B = (A \cup B) \cap C_S(A \cap B)$$

$\boxed{TODO : notation}$

30 Arithmetic

30.1 Why

Counting one by one is slow so we define an algebra on the naturals.

30.2 Sums and Addition

Let m and n be two natural numbers. If we apply the successor function to m n times we obtain a number. If we apply the successor function to n m times we obtain a number. Indeed, we obtain the same number in both cases. We call this number the **sum** of m and n. We say we **add** m to n, or vice versa. We call this correspondence, between (m, n) and the sum, **addition**.

30.2.1 Notation

We denote the function addition by $+$ and so denote the sum of the naturals m and n by $m + n$.

30.3 Products and Multiplication

Let m and n naturals. If we add n copies of m we obtain a number. If we add m copies of n we obtain a number. Indeed, we obtain the same number in both cases. We call this number the **product** of m and n. We say we **multiply** m to n, or vice versa. We call this symmetric operation mapping (m, n) to their product **multiplication**.

30.3.1 Notation

We denote the operation of multiplication by $\cdot$ and so denote the product of the naturals m and n by $m \cdot n$.

31 Equivalence Relations

31.1 Why

We want to handle at once all elements which are indistinguishable or equivalent in some aspect.

31.2 Definition

A relation R on a set A is an **equivalence relation** if it is reflexive, symmetric, and transitive.

For an element $a \in A$, we call the set of elements in relation R to a the **equivalence class** of a. The key observation, recorded and proven below, is that the equivalence classes partition the set A. A frequent technique is to define an appropriate equivalence relation on a large set A and then to work with the set of equivalence classes of A.

We call the set of equivalence classes the **quotient set** of A under R. An equally good name is the divided set of A under R, but this terminology is not standard. The language in both cases reminds us that $\sim$ partitions the set A into equivalence classes.

31.2.1 Notation

If R is an equivalence relation on a set A, we use the symbol $\sim$. When alone, $\sim$ is read aloud as "sim," but we still read $a \sim b$ aloud as "a equivalent to b." We denote the quotient set of A under $\sim$ by $A/\sim$, read aloud as "A quotient sim".

31.2.2 Results

$\boxed{TODO}$

32 Families

32.1 Why

We want to generalize operations beyond two objects.

32.2 Definition

Let A, B be non-empty sets. A **family** of elements of a first set **indexed** by elements of a second set is the range of a function from the second set to the first set. We call second set the **index set**.

If the index set is a finite set, we call the family a **finite family**. If the index set a countable set, we call the family a **countable family**. If the index set is an uncountable set, we call the family a **uncountable family**.

If the codomain is a set of sets, we call the family a **family of sets**. We often use a subset of the whole natural numbers as the index set. In this case, and for other indexed sets with orders, we call the family an **ordered family**

32.2.1 Notation

Let A be a non-empty set. We denote the index set by I, a mnemonic for index. For $i \in I$, let we denote the result of applying the function to i by a_i; the notation evokes evokes function notation but avoids naming the function.

We denote the family of a_α indexed with I by $\{a_\alpha\}_{\alpha \in I}$, which is short-hand for set-builder notation. We read this notation "a sub-alpha, alpha in I."

32.3 Operations

The **pairwise extension** of a commutative operation is the function from finite families of the ground set to the ground set obtained by applying the operation pairwise to elements.

The **ordered pairwise extension** of an operation is the function from finite families ground set to the ground set obtained by applying the operation pairwise to elements in order.

32.3.1 Notation

Let $(A, +)$ be an algebra and $\{A_i\}_{i=1}^n$ a finite family of elements of A. We denote the pairwise extension by

$$\mathop{+}_{i=1}^{n} A_i$$

32.4 Family Set Algebra

We define the set whose elements are the objects which are contained in at least one family member the **family union**. We define the set whose elements are the objects which are contained in all of the family members the **family intersection**.

32.4.1 Notation

We denote the family union by $\cup_{\alpha \in I} A_\alpha$. We read this notation as "union over alpha in I of A sub-alpha." We denote family intersection by $\cap_{\alpha \in I} A_\alpha$. We read this notation as "intersection over alpha in I of A sub-alpha."

32.4.2 Results

Proposition 7. *For an indexed family $\{A_\alpha\}_{\alpha \in I}$ in S, if $I = \{i, j\}$ then*

$$\cup_{\alpha \in I} A_\alpha = A_i \cup A_j$$

and

$$\cap_{\alpha \in I} A_\alpha = A_i \cap A_j.$$

Proposition 8. *For an indexed family $\{A_\alpha\}_{\alpha \in I}$ in S, if $I = \varnothing$, then*

$$\cup_{\alpha \in I} A_\alpha = \varnothing$$

and

$$\cap_{\alpha \in I} A_\alpha = S.$$

Proposition 9. *For an indexed family* $\{A_\alpha\}_{\alpha \in I}$ *in* S.

$$C_S(\cup_{\alpha \in I} A_\alpha) = \cap_{\alpha \in I} C_S(A_\alpha)$$

and

$$C_S(\cap_{\alpha \in I} A_\alpha) = \cup_{\alpha \in I} C_S(A_\alpha).$$

33 Partitions

33.1 Why

We divide a set into disjoint subsets whose union is the whole set. In this way we can handle each subset of the main set individually, and so handle the entire set piece by piece.

33.2 Definition

A **disjoint family** of sets is a family for which the intersection of any two member sets is empty. A **partition** of a set is a disjoint family of subsets of the set whose union is the set. A **piece** of a partition is an element of the family.

33.2.1 Notation

No new notation for partitions. Instead, we record the properties of partitions in previously introduced notation.

Let A be a set and $\{A_\alpha\}_{\alpha \in I}$ a family of subsets of A. We denote the condition that the family is disjoint by $A_\alpha \cap A_\beta = \varnothing$, for all $\alpha, \beta \in I$, We denote the condition that the family union is A by $\cup_{\alpha \in I} A_\alpha = A$.

34 Direct Products

34.1 Why

We can profitably generalize the notion of cartesian product to families of sets indexed by the natural numbers.

34.2 Direct Products

The **direct product** of family indexed by a subset of the naturals is the set whose elements are ordered sequences of elements from each set in the family. The ordering on the sequences comes from the natural ordering on N. If the index set is finite, we call the elements of the direct product **n-tuples**. If the index set is the natural numbers, and every set in the family is the same set A, we call the elements of the direct product the **sequences** in A.

34.2.1 Notation

For a family $\{A_\alpha\}_{\alpha \in I}$ of S with $I = \{1, \dots, n\}$, we denote the direct product by

$$\prod_{i=1}^{n} A_i.$$

We read this notation as "product over alpha in I of A sub-alpha." We denote an element of $\prod_{i=1}^{n} A_i$ by $(a_1, a_2, \ldots, a_n)$ with the understanding that $a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n$.

If I is the set of natural numbers we denote the direct product by

$$\prod_{i=1}^{\infty} A_i.$$

We denote an element of $\prod_{i=1}^{\infty} A_i$ by (a_i) with the understanding that $a_i \in A_i$ for all $i = 1, 2, 3, \ldots$. If $A_i = A$ for all $i = 1, 2, 3, \ldots$, then (a_i) is a sequence in A.

35 Sequences

35.1 Why

We introduce language for the steps of an infinite process.

35.2 Definition

A **sequence** is a function from the natural numbers to a set.

Equivalently, a sequence is an element of a direct product of a family of sets for which each set in the family is identical and the index set is the natural numbers.

35.2.1 Notation

36 Monotone Sequences

36.1 Why

If the base set of a sequence has a partial order, then we can discuss its relation to the order of sequence.

36.2 Definition

A sequence on a partially ordered set is **non-decreasing** if whenever a first index precedes a second index the element associated with the first index precedes the element associated with the second element. A sequence on a partially ordered set is **increasing** if it is non-decreasing and no two elements are the same.

A sequence on a partially ordered set is **non-increasing** if whenever a first index precedes a second index the element associated with the first index succedes the element associated with the second element. A sequence on a partially ordered set is **decreasing** if it is non-increasing and no two elements are the same.

A sequence on a partially ordered set is **monotone** if it is non-decreasing, or non-increasing. An increasing sequence is non-decreasing. A decreasing sequences is non-increasing. A sequence on a partially ordered set is **strictly monotone** if it is decreasing, or increasing.

36.2.1 Notation

Let A a non-empty set with partial order $\preceq$. Let $\{a_n\}_n$ a sequence in A.

The sequence is non-decreasing if $n \leq m \implies a_n \preceq a_m$, and increasing if $n < m \implies a_n \prec a_m$. The sequence is non-increasing if $n \leq m \implies a_n \succeq a_m$, and decreasing if $n < m \implies a_n > a_m$.

36.3 Examples

Example 10. *Let A a non-empty set and $\{A_n\}_n$ a sequence of sets in 2^A. Partially order elements of 2^A by the relation contained in.*

37 Nets

37.1 Why

We generalize the notion of sequence to index sets beyond the naturals.

37.2 Definition

A sequence is a function on the natural numbers; this set has two important properties: (a) we can order the natural numbers and (b) we can always go "further out."

To elaborate on property (b): if handed two natural numbers m and n, we can always find another, for example $\max\{m, n\}+1$, larger than m and n. We might think of larger as "further out" from the first natural number: 1.

Combining these to observations, we define a directed set:

Definition 11. *A **directed set** is a set D with a partial order $\preceq$ satisfying one additional property: for all $a, b \in D$, there exists $c \in D$ such that $a \preceq c$ and $b \preceq c$.*

Definition 12. *A **net** is a function on a directed set.*

A sequence, then, is a net. The directed set is the set of natural numbers and the partial order is $m \preceq n$ if $m \leq n$.

37.2.1 Notation

Directed sets involve a set and a partial order. We commonly assume the partial order, and just denote the set. We use the letter D as a mnemonic for directed.

For nets, we use function notation and generalize sequence notation. We denote the net $x : D \to A$ by $\{a_\alpha\}$, emulating notation for sequences. The use of α rather than n reminds us that D need not be the set of natural numbers.

38 Categories

38.1 Why

We generalize the notion of sets and functions.

38.2 Definition

A **category** is a collection of objects together with a set of **category maps** for each ordered pair of objects. The set of maps has a binary operation called **category composition**, whose induced algebra is associative and contains identities.

As the fundamental example, consider the category whose objects are sets and whose maps are functions. The sets are the objects of the category. The functions are the maps. The rule of composition is ordinary function composition. The map identities are the identity functions. We call this category the **category of sets**.

38.2.1 Notation

Our notation for categories is guided by our generalizing the notions of set and functions.

We denote categories with upper-case latin letters in script; for example, $\mathcal{C}$. We read $\mathcal{C}$ aloud as "script C." Upper case latin letters remind that the category is a set of objects. The script form reminds that these objects may themselves be sets.

We denote the objects of a category by upper-case latin letters, for example A, B, C; an allusion to the idea that these generalize sets. We denote the set of maps for an ordered pair of objects (A, B) by $A \to B$; an allusion to the function notation. We denote members of $A \to B$ using lower case latin letters, for example f, g, h; an allusion to our function notation.

39 Groups

39.1 Why

We generalize the algebraic structure of addition over the integers.

39.2 Definition

A **group** is an algebra with: (1) an associative operation, (2) an identity element, and (3) an inverse for each element. We call the operation of the algebra **group addition**. A **commutative group** is a group whose operation commutes.

39.2.1 Notation

TODO

40 Rings

40.1 Why

We generalize the algebraic structure of addition and multiplication over the integers.

40.2 Definition

A **ring** is two algebras over the same ground set with: (1) the first algebra a commutative group (2) an identity element in the second algebra, and (3) the operation of the second algebra distributes over the operation of the first algebra.

We call the operation of the first algebra **ring addition**. We call the operation of the second algebra **ring multiplication**.

40.2.1 Notation

TODO

41 Fields

41.1 Why

We generalize the algebraic structure of addition and multiplication over the rationals.

41.2 Definition

A **field** is two algebras over the same ground set with: (1) both algebras are commutative groups (2) the operation of the second algebra distributes over the operation of the first algebra.

We call the operation of the first algebra **field addition**. We call the operation of the second algebra **field multiplication**.

41.2.1 Notation

$\boxed{TODO}$

42 Vectors

42.1 Why

We speak of objects which we can add and scale.

42.2 Definition

42.2.1 Notation

43 Homomorphism

43.1 Why

We name a function which preserves group structure.

43.2 Definition

A **homomorphism** from group $(A, +)$ to group $(B, \tilde{+})$ is a function $f : A \to B$ such that $f(e_A) = f(e_B)$ for identities $e_A \in A$ and $e_B \in B$ and $f(a + a') = f(a) \tilde{+} f(a')$ for all $a, a' \in A$.

43.2.1 Notation

$\boxed{TODO}$

44 Cardinality

44.1 Why

We want to speak of the number of elements of a set. Subtetly arises when we can not finish counting the set's elements.

44.2 Finite Definition

If a set A is contained in a set B and not equal to B, we say that B is a **larger set** than A. Conversely, we say that A is a **smaller set** than B. We reason that we could pair the elements of B with themselves in A and still have some elements of B left over.

A **finite set** is one whose elements we can count and the process terminates. For example, $\{1, 2, 3\}$ or $\{a, b, c, d\}$. The **cardinality** of a finite set is the number of elements it contains. The cardinality of $\{1, 2, 3\}$ is 3 and the cardinality of $\{a, b, c, d\}$ is 4.

44.2.1 Notation

Let A be a non-empty set. We denote the cardinality of A by $|A|$.

44.3 Infinite Definition

Suppose we know that the counting process could never terminate. This situation superficially seems bizarre, but is in fact built in to some of our fundamental notions: namely, the natural numbers. We defined the natural numbers in a manner which made them not finite.

If we had a bag of natural numbers, we could use the total order to find the largest, and then use the existence of a successor to add a new largest number. Therefore, bizarrely, the process of counting the natural numbers can not terminate.

An **infinite set** is a non-empty set which is not finite. So the natural numbers are an infinite set. Alternatively we say that there are **infinitely many** natural numbers. The negating prefix "in" emphasizes that we have defined the nature of the size of the naturals indirectly: their size is not something we understand from the simple intuition of counting, but in contrast to the simple intuition of counting.

Still, we imagine that if we could go on forever, we could count the natural numbers; so in an infinite sense, they are countable. A **countable** set is one which is either (a) finite or (b) one for which there exists a one-to-one function mapping the natural numbers onto the set.

The natural numbers are countable: we exhibit the identity function. Less obviously the integer numbers and rational numbers are countable. Even more bizarre, the real numbers are not countable. An **uncountable** set is one which is not countable.

44.3.1 Notation

We denote the cardinality of the natural numbers by $\aleph_0$.

45 Subset System

45.1 Why

We speak of a set and a set of its subsets satisfying properties. The utility of this abstract concept is proved by its examples, in future sheets.

45.2 Definition

A **subset system** is a a pair of sets: the second set contains subsets of the first.

We call the first set the **base set**. If the base set is finite, we call the subset system a **finite subset system**. A **distinguished subset** is an element of the second set. An **undistingished subset** is a subset of the first set which is not distinguished.

45.2.1 Notation

Let A be a set and $\mathcal{A} \subset 2^A$. We denote the subset system of A and $\mathcal{A}$ by $(A, \mathcal{A})$, read aloud as "A, script A."

45.3 Example

Example 13. *Let A be a nonempty set. Let $\mathcal{A}$ be 2^A. Then $(A, \mathcal{A})$ is a subset system.*

46 Topological Space

46.1 Why

We want to generalize the notion of continuity.

46.2 Definition

A **topological space** is a subset system for which: (1) the empty set and the base set are distinguished, (2) the intersection of a finite family of distinguished subsets is distinguished, and (3) the union of a family of distinguished subsets is distinguished. We call the set of distinguished subsets the **topology**. We call the distinguished subsets the **open sets**.

46.2.1 Notation

Let A be a non-empty set. For the set of distinguished sets, we use $\mathcal{T}$, a mnemonic for topology, read aloud as "script T". We denote elements of $\mathcal{T}$ by O, a mnemonic for open. We denote the topological space with base set A and topology $\mathcal{T}$ by $(A, \mathcal{T})$. We denote the properties satisfied by elements of $\mathcal{T}$:

1. $X, \varnothing \in \mathcal{T}$

2. $\{O_i\}_{i=1}^n \subset \mathcal{T} \implies \cap_{i=1}^n O_i \in \mathcal{T}$

3. $\{O_\alpha\}_{\alpha \in I} \subset \mathcal{T} \implies \cup_{\alpha \in I} \in \mathcal{T}$

47 Monotone Classes

47.1 Why

47.2 Definition

The **limit** of an increasing sequence of sets is the family union of the sequence. The **limit** of a decreasing sequence of sets is the family intersection of the sequence.

A **monotone limit** of an sequence of sets is the limit of a monotone sequence.

A **monotone space** is a subset space in which monotone limits of monotone sequences of distinguished sets are distinguished. We call the distinguished sets a **monotone class**.

47.2.1 Notation

Let A a non-empty set with partial order $\preceq$. Let $(A, \mathcal{A})$ be a subset space on A.

Let $\{A_n\}_n$ be an increasing or decreasing sequence in $\mathcal{A}$. We denote the limit of $\{A_n\}_n$ by $\lim_n A_n$.

If $\{A_n\}_n$ is increasing, $\lim_n A_n = \cup_n A_n$. If $\{A_n\}_n$ is decreasing, $\lim_n A_n = \cap_n A_n$.

If $(A, \mathcal{A})$ is a monotone space, then for all monotone $\{A_n\}_n$ in $\mathcal{A}$, $\lim_n A_n \in \mathcal{A}$. In this case, $\mathcal{A}$ is a montone class.

48 Subset Algebra

48.1 Why

We speak of a subset space with set-algebraic properties.

48.2 Definition

A **subset algebra** is a subset space for which (1) the base set is distinguished (2) the complement of a distinguished set is distinguished (3) the union of two distinguished sets is distinguished.

We call the set of distinguished sets an **algebra** on the the base set. We justify this language by showing that the standard set operations applied to distinguished sets result in distinguished sets.

If a set of subsets is closed under complements it contains the base set if and only if it contains the empty set. So we can replace condition (1) by insisting that the algebra contain the empty set. Similarly, if a non-empty set of subsets is closed under complements and unions then it contains the base set: the union of a distinguished set and its complement. Thus we can replace condition (1) by insisting that the algebra be non-empty.

48.2.1 Notation

The notation follows that of a subset space. Let $(A, \mathcal{A})$ be a subset algebra. We also say "let $\mathcal{A}$ be an algebra on A." Moreover, since the largest element of the algebra is the base set, we can say without ambiguity: "let $\mathcal{A}$ be an algebra."

48.3 Properties

Proposition 14. *For any subset algebra, $\varnothing$ is distinguished.*

Proposition 15. *For any subset algebra, for any distinguished sets, (a) the intersection is distinguished and (b) their symmetric difference is distinguished. So, if one contains the other, the complement of the smaller in the larger is distinguished.*

Proposition 16. *For any subset algebra, for any finite family of distinguished sets, (a) the finite family union and (b) the finite family intersection are both distinguished.*

So we could have defined an algebra by insisting it be closed under finite intersections.

48.4 Examples

Example 17. *For any set A, $(A, 2^A)$ is a subset algebra.*

Example 18. *For any set A, $(A, \{A, \varnothing\})$ is a subset algebra.*

Example 19. *For any infinite set A, let $\mathcal{A}$ be the set*

$$\{B \subset A \mid |B| < \aleph_0 \vee |C_A(B)| < \aleph_0\}.$$

$\mathcal{A}$ is an algebra; the ***finite/co-finite algebra****.*

Example 20. *For any infinite set A, let $\mathcal{A}$ be the set*

$$\{B \subset A \mid |B| \leq \aleph_0 \vee |C_A(B)| \leq \aleph_0\}.$$

$\mathcal{A}$ is an algebra; the ***countable/co-countable algebra****.*

Example 21. *For any infinite set A, let $\mathcal{A}$ be the set*

$$\{B \subset A \mid |B| \leq \aleph_0\}.$$

$\mathcal{A}$ is not an algebra.

Example 22. *Let A be an uncountable set. Let $\mathcal{A}$ be the collection of all countable subsets of A. $\mathcal{A}$ is not a sigma algebra.*

49 Monotone Algebra

49.1 Why

Closure under monotone limits is a weaker condition than that included in the definition of sigma algebras, but is sufficient if the set is also an algebra. TODO: why

49.2 Result

If a subset algebra is a monotone space, then it is a countably summable subset algebra.

Proposition 23. *A subset algebra is a countably summable if either:*

1. *the limit of a nondecreasing sequence of distinguished sets is distinguished*

2. *the limit of a nonincreasing sequence of distinguished sets is distinguished.*

Proof. TODO □

50 Monotone Class Theorem

50.1 Why

TODO

50.2 Result

Proposition 24. *The sigma algebra generated by an algebra of sets is the same as the monotone class generated by the algebra.*

Proof. TODO ∎

51 Rational Numbers

51.1 Why

51.2 Definition

52 Real Numbers

52.1 Why

52.2 Definition

53 Length Common Notions

53.1 Why

We want to define the length of a subset of real numbers.

53.2 Notions

We take two common notions:

1. The length of a whole is the sum of the lengths of its parts; the **additivity principle**.
2. The length of a whole is the at least the length of any whole it contains the **containment principle**.

The task is to make precise the use of “whole,”, “parts,” and “contains.” We start with intervals.

53.3 Definition

By whole we mean set. By part we mean an element of a partition. By contains we mean set containment.

The **length** of an interval is the difference of its endpoints: the larger minus the smaller.

Two intervals are **non-overlapping** if their intersection is a single point or empty. The **length** of the union of two non-overlapping intervals is the sum of their lengths.

A **simple** subset of the real numbers is a finite union of non-overlapping intervals. The length of a simple subset is the sum of the lengths of its family.

A **countably simple** subset of the real numbers is a countable union of non-overlapping intervals. The length of a countably simple subset is the limit of the sum of the lengths of its family; as we have defined it, length is positive, so this series is either bounded and increasing and so converges, or is infinite, and so converges to $+\infty$.

At this point, we must confront the obvious question: are all subsets of the real numbers countably simple? Answer: no. So, what can we say?

A **cover** of a set A of real numbers is a family whose union is a contains A. Since a cover always contains the set A, it's length, which we understand, must be larger (containment principles) than A. So what if we declare that the length of an arbitrary set A be the greatest lower bound of the lengths of all sequences of intervals covering A. Will this work?

53.3.1 Cuts

If a, b are real numbers and $a < b$, then we **cut** an interval with a and b as its endpoints by selecting c such that $a < c$ and $c < b$. We obtain two intervals, one with endpoints a, c and one with endpoints c, b; we call these two the **cut pieces**.

Given an interval, the length of the interval is the sum of any two cut pieces, because the pieces are non-overlapping.

53.4 All sets

Proposition 25. *Not all subsets of real numbers are simple.*

Exhibit: R is not finite.

Proposition 26. *Not all subsets of real numbers are countably simple.*

Exhibit: the rationals.

Here's the great insight: approximate a set by a countable family of intervals.

53.4.1 Notation

54 Intervals

54.1 Why

We name and denote subsets of the set of real numbers which correspond to segments of a line.

54.2 Definition

Take two real numbers, with the first less than the second.

An **interval** is one of four sets:

1. the set of real numbers larger than the first number and smaller than the second; we call the interval **open**.

2. the set of real numbers larger than or equal to the first number and smaller than or equal to the second number; we call the interval **closed**.

3. the set of real numbers larger than the first number and smaller than or equal to the second; we call the interval **open on the left** and **closed on the right**.

4. the set of real numbers larger than or equal to the first number and smaller than the second; we call the interval **closed on the left** and **open on the right**.

If an interval is neither open nor closed we call it **half-open** or **half-closed**

We call the two numbers the **endpoints** of the interval. An open interval does not contain its endpoints. A closed interval contains its endpoints. A half-open/half-closed interval contains only one of its endpoints. We say that the endpoints **delimit** the interval.

54.2.1 Notation

Let a, b be two real numbers which satisfy the relation $a < b$.

We denote the open interval from a to b by (a, b). This notation, although standard, is the same as that for ordered pairs; no confusion arises with adequate context.

We denote the closed interval from a to b by $[a, b]$. We record the fact $(a, b) \subset [a, b]$ in our new notation.

We denote the half-open interval from a to b, closed on the right, by $(a, b]$ and the half-open interval from a to b, closed on the left, by $[a, b)$.

55 Interval Length

55.1 Why

Toward defining the length of a subset of real numbers, we start by defining the length of an interval.

55.2 Definition

The **length** of an interval is the difference of its endpoints: the larger less the smaller.

55.3 Notation

Let a, b be real numbers which satisfy the relation $a < b$. The length of (a, b), $[a, b]$ $[a, b)$ and $(a, b]$ is, in each case, $b - a$.

For example, the length of the interval $(0, 1)$ is one.

56 Interval Partitions

56.1 Why

We partition a real interval into interval pieces.

56.2 Definition

An **interval partition** is a finite partition of a closed real interval.

An interval partition is **regular** if all pieces except the largest are closed on the left and open on the right and the largest is closed.

Any regular interval partition with $n-1$ elements can be represented by $n+1$ real numbers: the endpoints of each interval. We call these the **cut points** of the interval partition.

56.2.1 Notation

Let R denote the set of real numbers. Let $[a, b]$ a closed interval in R with endpoints $a, b \in R$.

Consider a regular partition. of $[a, b]$ with $n-1$ pieces. We can identify its cut points:

$$a = a_1 < a_2 < \dots a_{n-1} < a_n = b.$$

The pieces of the partition are:

$$[a_1, a_2), [a_2, a_3), \dots, [a_{n-2}, a_{n-1})[a_{n-1}, a_n].$$

57 Real Functions

57.1 Why

We define functions mapping real numbers to real numbers.

57.2 Definition

A **real function** is a function from subset of the real numbers into a subset of the real numbers. When clear from context, we call a real function a function.

Often, the domain is an interval. In this case, we say that the function is defined on a closed interval of the real line. We usually leave the codomain as the set of real numbers, unless we wish to speak of the function being onto.

57.2.1 Notation

Let R denote the set of real numbers. Let $f : R \to R$. Then f is a real function.

Let $a, b \in R$. Let $[a, b] \subset R$ a closed interval of real numbers. Let $f : [a, b] \to R$. Then f is a real function defined on a closed interval.

We regularly declare the interval and the function in one pass: Let $f : [a, b] \to R$, read aloud as "f from closed a b to R."

58 Absolute Value

58.1 Why

We want a notion of distance between elements of the real line.

58.2 Definition

We define a function mapping a real number to its length from zero.

59 Characteristic Functions

59.1 Why

We represent rectangles by functions.

59.2 Definition

The **characteristic function** of a subset of some base set is the function from the base set to the real numbers which maps elements contained in the subset to value one and maps all other elements to zero. The range of the funtion is the set consisting of the real numbers one and zero.

If the base set is the real numbers and the subset is an interval, then the characteristic function is a rectangle with height one and the width of the interval.

59.2.1 Notation

Let A be a non-empty set and $B \subset A$. We denote the characteristic function of B in A by $\chi_B : A \to R$. The Greek letter χ is a mnemonic for "characteristic".

The subscript indicates the set on which the function is one. In other words, for all $B \subset A$, $\chi_B^{-1}(\{1\}) = B$.

If B is an interval and α is a real number then $\alpha\chi_B$ is a rectangle with height α.

60 Simple Functions

60.1 Why

We want to define area under a real function. We define functions for which this notion is clear.

60.2 Definition

A **simple function** is a function whose range is a finite set.

Partition the range into the finite family of one-element sets The family whose members consist of the inverse images of these sets is a partition of the domain. We call this the **simple partition** of the function.

A **real simple function** is a simple function whose codomain is real. In this case, we can write the simple function as a sum of the characteristic functions of the inverse images elements.

60.2.1 Notation

Let A and B be non-empty sets. We denote the set of simple functions from A to B by $\mathcal{SF}(A, B)$.

We denote the set of simple real functions with domain A by by $\mathcal{SF}(A)$. We denote subset of non-negative simple real functions with domain A by by $\mathcal{SF}_+(A)$.

Let $f \in \mathcal{SF}(A,)$. Order the members of the range of f from 1 to n as $r_1, \ldots, r_n$. Define $A_i = f^{-1}(\{r_i\})$. Then $f = \sum_{i=1}^{n} r_i \chi_{A_i}$.

61 Distance

61.1 Why

We want to talk about the "distance" between objects in a set.

61.2 Common Notions

Our inspiration is the notion of distance in the plane of geometry. The objects are points and the distance between them is the length of the line segment joining them. We note a few properties of this notion of distance:

1. The distance between any two distinct objects is not zero.
2. The distance between any two objects does not depend on the order in which we consider them.
3. The distance between two objects is no larger than the sum of the distances of each with any third object

The first observation is natural: if two points are not the same, then they are some distance apart. In other words, the line segment between them has length.

The second observation is natural: the line segment connecting two points does not depend on the order specifying the points. This observation justifies the word "between." If it were not the case, then we should use different words, and be careful to speak of the distance "from" a first point "to" a second point.

The third property is a non-obvious property of distance in the plane. It says, in other words, that the length of any side of a triangle is no larger than the sum of the lengths of the two other sides. With experience in geometry, the observation may become natural. But it does not seem to be superficially so.

A more muddled but superficially natural justification for our concern with third observation is that it says something about the transitivity of closeness. Two objects are close if their distance is small. Small is a relative concept, and needs some standard of comparison. Let us fix two points, take the distance between them, and call it a unit. We call two objects close with respect to our unit if their distance is less than a unit.

In this language, the third observation says that if we know two objects are each half of a unit distance from a third object, then the two objects are close (their distance is less than a unit). We might call this third object the reference object. Here, then, is the usefulness of the third property: we can infer closeness of two objects if we know their distance to a reference object.

62 Distance Asymmetry

62.1 Why

Sometimes "distance" as used in the English language refers to an asymmetric concept. This apparent paradox further illuminates the symmetry property.

62.2 Apparent Paradox

Distance in the plane is symmetric: the distance from one point to another does not depend on the order of the points so considered. We took this observation as a definiting property of our abstract notion of distance. The meaning, strength, and limitation of this property is clarified by considering an asymmetric case.

Contrast walking up a hill with walking down it. The "distance" between these two points, the top of the hill and a point on its base, may not be symmetric with respect to the time taken or the effort involved. Experience might suggest that it will take longer to walk up the hill than to walk down it. A superficial justification may include reference to the some notion of uphill walking requiring more effort.

If we were going to model the top and base of the hill as points in space, however, the distance between them is the same: it is symmetric. It is even the same if we take into account that some specific path, a trail say, must be followed.

If planning a backpacking trip, such symmetry appears foolish. The distance between two locations must not be considered symmetric. Going up the mountain takes longer than going down. It may justify, in the English phrase, "going around, rather than going over."

63 Metrics

63.1 Why

We want to talk about a set with a prescribed quantitative degree of closeness (or distance) between its elements.

63.2 Definition

The correspondences which serve as a degree of closeness, or measure of distance, must satisify our notions of distances previously developed.

A function on ordered pairs which does not depend on the order of the elements so considered is **symmetric**. A function to the real numbers which takes only non-negative values is **non-negative**. A repeated pair is an ordered pair of the same element twice. A function which satisfies a triangle inequality for any three elements is **triangularly transitive**.

A **metric** (or **distance function**) is a function on ordered pairs of elements of a set which is symmetric, non-negative, zero only on repeated pairs, and triangularly transitive. A **metric space** is an ordered pair: a set with a metric on the set.

In a metric space, we say that one pair of objects is **closer** together if the metric of the first pair is smaller than the metric value of the second pair.

63.2.1 Notation

Let A be a set and let R be the set of real numbers. We commonly denote a metric by the letter d, as a mnemonic for "distance." Let $d : A \times A \to R$. Then d is a metric if:

1. it is non-negative, which we tend to denote by
$$d(a,b) \geq 0, \quad \forall a, b \in A.$$
2. it is 0 only on repeated pairs, which we tend to denote by
$$d(a,b) = 0 \Leftrightarrow a = b, \quad \forall a, b \in A.$$
3. it is symmetric, which we tend to denote by:
$$d(a,b) = d(b,a), \quad \forall a, b \in A.$$
4. it is triangularly transitive, which we tend to denote by
$$d(a,b) \leq d(a,c) + d(c,b), \quad \forall a, b, c \in A.$$

As usual, we denote the metric space of A with d by (A, d).

64 Product Metrics

64.1 Why

Given n sets each with metrics, there is a standard way of turning the direct product of the sets into a metric space. In other words, defining a distance on the tuples of elements from the sets.

64.2 Motivating Result

Proposition 27. *Let $(A_1, d_1), (A_2, d_2), \dots, (A_n, d_n)$ be metric spaces. Let A be $\prod_{i=1}^{n} A_n$ and let R be the set of real numbers. Define $d : A \times A \to R$ by*

$$d(a, b) = \max\{d_1(a_1, b_1), \dots, d_n(a_n, b_n)\}.$$

Then (A, d) is a metric space.

65 Real Limits

65.1 Why

We want to speak of an infinite process which, although never arrives, does terminate.

65.2 Definition

A **limit** of a sequence of real numbers is a real number for which we can always find a final part of the sequence wholly contained in an interval around the limit, no matter how small the interval.

You propose a limit for a sequence. To test this proposal, I specify some small positive real number. Then we look for a final part wholly contained in the interval of that width. If we can always find the final part, no matter how small the positive number I specified, then the proposed limit is true.

65.2.1 Existence

Some sequences have no limits. Consider the sequence which alternates between the $+1$ and -1. To show that the limit does not exist, we argue indirectly. We take any real number and test it with the interval length one. No matter which real number we have selected, $+1$ and -1 are a distance two apart, and so can not both be contained in an interval of width one.

65.2.2 Uniqueness

If a sequence has a limit, it has only one limit. So, from here on, we will speak of **the limit** of the sequence.

To see this uniqueness, suppose that two real numbers satisfy the limiting property. We now argue indirectly: suppose also that they are not equal. Denote the distance between them by x. Then ask for final parts in intervals of width $x/2$ for both limits.

65.2.3 Approximation

We use limits to speak about the terminating behavior of infinite processes. We think about the sequence as approximating the limit. The sequence may never actually take the value of its limit, so the limit need be in the set of terms of the sequence, but it does get close.

The definition, moreover, ensures that the sequence will get arbitrarily close. We can operationalize this property, by taking the first element of that final part after which all elements are close to the limit. This element is an element of the sequence approximates the limit value well.

65.2.4 Notation

Let $\{a_n\}_n$ be a sequence of real numbers. Let a be a real number. We denote that a is the limit of $\{a_n\}_n$ by

$$a = \lim_{n \to \infty} a_n.$$

We read this statement aloud as “a is the limit of a sub n.” The above statement asserts two facts: (1) the sequence $\{a_n\}_n$ has a limit and (2) the limit is the real number a. We sometimes abbreviate the by writing $a = \lim_n a_n$.

66 Real Limiting Bounds

66.1 Why

We can think of a limit as existing when the limit of upper bounds and lower bounds on final parts of the sequence coincide.

66.2 Definition

The **limit superior** of a sequence of real numbers is the limit of the sequence of suprema of final parts of the sequence. Similarly, the **limit inferior** of a sequence of real numbers is the limit of the sequence of infima of final parts of the sequence.

The limit of the sequence exists if and only if the If the limit superior and the limit inferior coincide, then the sequence has a limit which is defined to be the limiting value of each of those two sequences.

$\liminf_n f_n$

$$\liminf_n f_n$$

67 Extended Real Numbers

67.1 Why

That some limits grow without bound leads us to add two elements to the set of real numbers.

67.2 Definition

The set of **extended real numbers** is the union of the set of real numbers with a set containing two elements: one we call **positive infinity** and we call **negative infinitty**.

67.2.1 Extended Arithmetic

We extend addition to all but one ordered pair of elements of the new set. The sum of any real number with a real number is defined as before. The sum any real number with positive infinity is positive infinity. The sum any real number with negative infinity is negative infinity. The sum of positive infinity with positive infinity is positive infinity. The sum of negative infinity with negative infinity is negative infinity. We do not define the sum of positive infinity and negative infinity. TODO

67.2.2 Notation

Let R denote the set of real numbers. We denote the element positive infinity by $+\infty$ and we denote the element negative infinity by $-\infty$. The set of extended real numbers is the set $R \cup \{+\infty, -\infty\}$.

67.3 Intervals

We

68 Real Length Impossible

68.1 Why

Given a subset of the real line, what is its length?

68.2 Background

Let $a, b \in R$ with $a \leq b$. The **length** of the closed interval of the real numbers $[a, b]$ is $b - a$. The length is non-negative.

A family $\{A_\alpha\}_{\alpha \in I}$ is **disjoint** if for $\alpha, \beta \in I$, $\alpha \neq \beta$, then $A_\alpha \cap A_\beta = \varnothing$. A set A can be **partioned** into a family if there exists a disjoint family whose union is A. A set $A \subset R$ is **simple** if it can be partitioned into a countable family whose members are closed intervals. The above discussion suggests that we should define the length of a simple set as the sum of the lengths of sets which parition it.

The above discussion suggests that if we wish to define a function length : $2^R \to R \cup \{-\infty, \infty\}$, we should ask that (1) length$(A) \geq 0$, (2) length$([a, b]) = b - a$, (3) for disjoint closed intervals $\{A_n\}_{n \in N}$, length$(A_i) = \sum_i$ length(A_i), and (4) for all $A \subset R$ and $a \in R$, length$(A + x) =$ length(A).

68.3 Converse

Define the equivalence relation $\sim$ on R by by $x \sim y$ if $x \sim y \in Q$

68.3.1 Notation

Let A be a set and $\mathcal{A} \subset 2^A$. We denote the subset algebra of A and $\mathcal{A}$ by $(A, \mathcal{A})$, read aloud as "A, script A."

68.4 Properties

Proposition 28. *For any set A, 2^A is a sigma algebra.*

Proposition 29. *The intersection of a family of sigma algebras is a sigma algebra.*

68.5 Generation

Proposition 30. *Let A a set and $\mathcal{B}$ a set of subsets. There is a unique smallest sigma algebra $(A, \mathcal{A})$ with $\mathcal{B} \subset \mathcal{A}$.*

We call the unique smallest sigma algebra containing B the **generated sigma algebra** of B.

69 Sigma Algebra

69.1 Why

For general measure theory, we need an algebra of sets closed under countable unions; we define such an object (TODO).

69.2 Definition

A **countably summable subset algebra** is a subset space for which (1) the base set is distinguished (2) the complement of a distinguished set is distinguished (3) the union of a sequence of distinguished sets is distinguished.

The name is justified, as each countably summable subset algebra is a subset algebra, because the union of $A_1, \dots, A_n$ coincides with the union of $A_1, \dots, A_n, A_n, A_n \dots$.

We say that the set of distinguished sets a **sigma algebra** on the base set; we justify this langauge, as for an algebra, by the closure properties under standard set operations.

69.2.1 Notation

The notation follows that of a subset space. Let $(A, \mathcal{A})$ be a countably summable subset algebra. We also say "let $\mathcal{A}$ be an sigma algebra on A." Moreover, since the largest element of the sigma algebra is the base set, we can say without ambiguity: "let $\mathcal{A}$ be a sigma algebra."

69.3 Examples

Example 31. *For any set A, 2^A is a sigma algebra.*

Example 32. *For any set A, $\{A, \varnothing\}$ is a sigma algebra.*

Example 33. *Let A be an infinite set. Let $\mathcal{A}$ the collection of finite subsets of A. $\mathcal{A}$ is not a sigma algebra.*

Example 34. *Let A be an infinite set. Let $\mathcal{A}$ be the collection subsets of A such that the set or its complement is finite. $\mathcal{A}$ is not a sigma algebra.*

Proposition 35. *The intersection of a family of sigma algebras is a sigma algebra.*

Example 36. *For any infinite set A, let $\mathcal{A}$ be the set*

$$\{B \subset A \mid |B| \leq \aleph_0 \vee |C_A(B)| \leq \aleph_0\}.$$

*$\mathcal{A}$ is an algebra; the **countable/co-countable algebra**.*

$\boxed{TOOD : cleanupexamples}$

70 Generated Sigma Algebra

70.1 Why

A simple way to obtain a sigma algebra, is to ask it to obtain some sets, and then to ask it to contain all the sets it needs to fulfill the properties.

70.2 Definition

The **generated sigma algebra** for a set of subsets is the smallest sigma algbera containing the set of subsets. We must prove the existence and uniqueness of this sigma algebra.

Proposition 37. *The intersection of a non-empty set of sigma algebras on the same base set is a sigma algebra.*

Proof. Let $\{(A, \mathcal{A}_\alpha\}_{\alpha \in I}$ a familiy of sigma algebras on the same base set. Define $\mathcal{A}$ as $\cap_{\alpha \in I} \mathcal{A}_\alpha$.

1. For all $\alpha \in I$, $A \in \mathcal{A}_\alpha$, thus $A \in \mathcal{A}$; condition (a).
2. For all $B \in \mathcal{A}$, for all $\alpha \in I$, $B \in \mathcal{A}_\alpha$. Thus, for all $\alpha \in I$, $C_A(B) \in \mathcal{A}_\alpha$. And so $C_A(B) \in \mathcal{A}$; condition (b).

3. For all sequences $\{B_n\} \subset \mathcal{A}$, $\{B_n\} \subset \mathcal{A}_\alpha$ for all α. Thus $\cup_n B_n \in \mathcal{A}_\alpha$ for all α and so $\cup_n B_n \in \mathcal{A}$; condition (c).

□

On the other hand, the union of a set of sigma algebras can fail to be a sigma algebra.

Proposition 38. *If A is a set and $\mathcal{A} \subset 2^A$, then there is a unique a smallest sigma algebra containing $\mathcal{A}$.*

Proof. We know of one sigma algebra containing $\mathcal{A}$: the power set of A. Thus, the set of sigma algebras containing $\mathcal{A}$ is not empty. Proposition 37 implies the intersection of all such sigma algebras (containing $\mathcal{A}$) is a sigma algebra. The intersection contains $\mathcal{A}$, and is contained in all other sigma algebras with this property, so is a smallest sigma algebra containing $\mathcal{A}$. If $\mathcal{B}, \mathcal{C}$ were two smallest sigma algebras, then $\mathcal{B} \subset \mathcal{C}$ and $\mathcal{C} \subset \mathcal{B}$, but then $\mathcal{B} = \mathcal{C}$; thus the smallest sigma algebra is unique. □

70.3 Notation

Let A be a set and $\mathcal{A} \subset 2^A$. We denote the sigma algebra generated by $\mathcal{A}$ by $\sigma(\mathcal{A})$.

71 Topological Sigma Algebra

71.1 Why

We often take the a the topology of a topological space as the generating set for the sigma algebra.

71.2 Definition

Given a topological space, the **topological sigma algebra** is the sigma algebra generated by the topology.

71.3 Notation

Let $(A, \mathcal{T})$ be a topological space. We denote the topological sigma algebra by $\sigma(\mathcal{T})$.

72 Borel Sigma Algebra

72.1 Why

We name and discuss the topological sigma algebra on the real numbers; the language and results generalize to finite direct products of the real numbers.

72.2 Definition

The **Borel sigma algebra** is the topological sigma algebra for the real numbers with the usual topology; we call its members the **Borel sets**.

72.3 Notation

Throughout this sheet we denote the real numbers by R. As usual, then, we denote the d-dimensional direct product of R by R^d. We denote the Borel sigma algebra on R^d by $\mathcal{B}(R^d)$. We denote $\mathcal{B}(R^1)$ by $\mathcal{B}(R)$.

72.4 Alternate Generations

The Borel sigma algebra is useful because it contains interesting sets besides the open sets. To make precise this statement, we show that the Borel sigma algebra is generated by, and therefore contains, other common subsets of R.

Proposition 39. *If a sigma algebra $\mathcal{A}$ includes a particular set of subsets $\mathcal{B}$, then $\mathcal{A}$ includes $\sigma(\mathcal{B})$.*

Proposition 40. *Each of*

(a) the collection of all closed subsets of R,

(b) the collection of all subintervals of R of the form $(-\infty, b]$,

(c) the collection of all subintervals of R of the form $(a, b]$,

generate $\mathcal{B}(R)$.

Proof. Denote the sigma algebra which corresponds to (a) by $\mathcal{B}_1$, that which corresponds to (b) by $\mathcal{B}_2$, and that which corresponds to (c) by $\mathcal{B}_3$. It suffices to establish $\mathcal{B}(R) \subset \mathcal{B}_3 \subset \mathcal{B}_2 \subset \mathcal{B}_1 \subset \mathcal{B}(R)$.

Start with $\mathcal{B}_1$. Closed sets are the complement of open sets. Thus $\mathcal{B}(R)$ contains all closed sets and so contains the sigma algebra generated by all closed sets, namely $\mathcal{B}_1$.

Next, $\mathcal{B}_2$. The intervals $(-\infty, b])$ are closed. Thus $\mathcal{B}_1$ contains all such intervals, and so contains the sigma algebra generated by such intervals, namely $\mathcal{B}_2$.

Next, $\mathcal{B}_3$. An interval $(a, b]$ is $(-\infty, b) \cap C_R((-\infty, a])$. Thus, all such intervals are contained in $\mathcal{B}_2$, and so $\mathcal{B}_2$ contains the sigma algebra generated by all such intervals, namely, $\mathcal{B}_3$.

Each open interval of R is the union of a sequence of sets $(a, b]$; namely $(a, b - 1/n]$. So $\mathcal{B}_3$ contains all open intervals (a, b). Each open set of R can be written as a countable union of open intervals (proof: $\boxed{TODO}$). Thus, $\mathcal{B}_3$ contains all open sets, and therefore contains the sigma algebra generated by the open subsets, namely $\mathcal{B}(R)$. □

Proposition 41. *Each of:*

(a) the collection of all closed subsets of R^d,

(b) the collection of all closed half-spaces of R^d of the form

$$\{(x_1, \ldots, x_d) \mid x_i \leq b_i\}$$

for some index i and some b in R, and

(c) the collection of all rectangles of R^d of the form

$$\{(x_1, \ldots, x_d) \mid a_i < x_i \leq b_i\}$$

generate $B(R^d)$.

Proof. Follow the proof of Proposition 40.

The complement of open sets are closed. Closed half spaces are closed. A strip of the form $\{(x_1, \ldots, x_d) \mid a_i < x_i \leq b_i\}$ is the intersection of two half-spaces in (b). Each rectangle in (c) is the union of d such strips.

$\boxed{TODO}$: two step last piece, open rectangles are unions of rectangles in (c) and open sets are union of open rectangles.

$\square$

73 Measures

73.1 Why

We want to generalize the notion of length, area, volume beyond the Lebesgue measure on the product spaces of real numbers.

73.2 Definition

An extended-real-valued non-negative function on an algebra is **finitely additive** if the result of the function applied to the union of a disjoint finite family of distinguished sets is the sum of the results of the function applied to each of the sets individually.

An extended-real-valued non-negative function on a sigma algebra is **countably additive** if the result of the function applied to the union of a disjoint countable family of distinguished sets is the limit of the partial sums of the results of the function applied to each of the sets individually.

A **finitely additive measure** is an extended-real-valued non-negative finitely additive function which associates the empty set with the real number 0. A **countably additive measure** is an extended-real-valued non-negative countably additive function which associates the empty set with the real number 0. We call countably additive measures **measures**, for short.

Every countably additive measure is finitely additive. On the other hand, there exist finitely additive measures which are not countable additive.

In the context of measure, we call a countably unitable subset algebra a **measurable space**. We call the distinguished sets **measurable** sets. A **measure space** is triple. As a pair, the first two objects are a measurable space. The third object is a measure defined on the sigma algebra of teh measurable space.

73.3 Notation

Let A a set. Let $\mathcal{A}$ a sigma algebra on A. The pair $(A, \mathcal{A})$ is a measurable space.

Let $\mu : \mathcal{A} \to [0, \infty]$ a measure; thus: (a) $\mu(\varnothing) = 0$ and (b) for disjoint $\{A_n\} \subset \mathcal{A}$, $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ The triple $(A, \mathcal{A}, \mu)$ is a measure space.

We use μ since it is a mnemonic for "measure". We often also us ν to denote measures, since it is after μ in the Greek alphabet, and λ, since it is before μ in the Greek alphabet.

73.4 Examples

Example 42. *Let $(A, \mathcal{A})$ a measurable space. Let $\mu : \mathcal{A} \to [0, +\infty]$ such that $\mu(A)$ is $|A|$ if A is finite and $\mu(A)$ is $+\infty$ otherwise. Then μ is a measure. We call μ the* ***counting mea-***

sure.

Example 43. *Let $(A, \mathcal{A})$ measurable. Fix $a \in A$. Let $\mu : \mathcal{A} \to [0, +\infty]$ such that $\mu(A)$ is 1 if $a \in A$ and $\mu(A)$ is 0 otherwise. Then μ is a measure. We call μ the* ***point mass*** *concentrated at a.*

Example 44. *Let R denote the real numbers. The Lebesgue measure on the measurable space $(R, \mathcal{B}(R))$ is a measure.*

Example 45. *Let N be the natural numbers. Let $\mathcal{A}$ the finite co-finite algebra on N. Let $\mu : \mathcal{A} \to [0, +\infty]$ be such that $\mu(A)$ is 1 if A is infinite or 0 otherwise. Then μ is a finitely additive measure. However it is impossible to extend μ to be a countably additive measure. Observe that if $A_n = \{n\}$ the $\mu(\cup_n A_n) = 1$ but $\sum_n \mu(A_n) = 0$.*

Example 46. *Let $(A, \mathcal{A})$ a measurable space. Let $\mu : \mathcal{A} \to [0, +\infty]$ be 0 if $A = \varnothing$ and $\mu(A)$ is $+\infty$ otherwise. Then μ is a measure.*

Example 47. *Let A be set with at least two elements ($|A| \geq 2$). Let $\mathcal{A} = 2^A$. Let $\mu : \mathcal{A} \to [0, +\infty]$ such that $\mu(A)$ is 0 if $A = \varnothing$ and $\mu(A) = 1$ otherwise. Then μ is not a measure, nor is μ finitely additive.*

Proof. Let $B, C \in \mathcal{A}$, $B \cap C = \varnothing$ then using finite additivity we obtain a contradiction $1 = \mu(B \cup C) = \mu(B) + \mu(C) = 2$. $\square$

74 Finite Measures

74.1 Why

Sometimes we want finite measures.

74.2 Definition

A measurable set is **finite** if its measure is a real number. The measure space itself is **finite** if the base set is finite.

A measurable set is **sigma-finite** if there exists a sequence of finite measurable sets whose union is the set. The measure space itself is **sigma-finite** if the base set is sigma finite.

74.2.1 Notation

We denote that a measure space is finite by saying "Let $(A, \mathcal{A}, \mu)$ and $\mu(A) < +\infty$."

Example 48. *Let $(A, \mathcal{A})$ be a measurable space.*

The counting measure on $(A, \mathcal{A})$ is finite if and only if the base set is finite. It is sigma finite if and only if the base set is a union of a sequence of finite sets.

If $\mathcal{A} = 2^A$, then the counting measure is sigma finite if and only if A is countable.

Example 49. *A point mass measure is finite.*

Example 50. *Let R be the set of real numbers. The Lebesgue measure on $(R, \mathcal{B}(R))$ is sigma finite.*

75 Measure Properties

75.1 Why

We expect measure to have the common sense properties we stated when trying to define a notion of length for the real line.

75.2 Monotonicity

An extended-real-valued function on an alebra is **monotone** if, given a first distinguished set contained in a distinguished second set, the result of the first is no greater than the result of the second.

Proposition 51. *All measures are monotone.*

Proof. Let $(A, \mathcal{A}, \mu)$ be a measure space. Let $A, B \in \mathcal{A}$ and $A \subset B$. Then $B = A \cup (B - A)$, a disjoint union. So

$$\mu(B) = \mu(A \cup (B - A)) = \mu(A) + \mu(B - A),$$

by the additivity of μ. Since $\mu(B - A) \geq 0$, we conclude $\mu(A) \leq \mu(B)$. $\square$

Proposition 52. *$A \subset B$ and B finite means $\mu(B - A) = \mu(B) - \mu(A)$.* $\boxed{TODO}$

75.3 Subadditivity

Monotonicity along with additivity of measures give us one other convenient property: subadditivity.

An extended-real-valued function on an algebra is **subadditive** if, given a sequence of distinguished sets, the result of union of the sequence is no greater than the limit of the partial sums of the results on each element of the sequence.

Proposition 53. *All measures are subadditive.*

Proof. Let $(A, \mathcal{A}, \mu)$ be a measure space.

Let $\{A_n\} \subset \mathcal{A}$. Define $\{B_n\} \subset \mathcal{A}$ with $B_n := A_n - \cup_{i=1}^{n-1} A_i$. Then $\cup_n A_n = \cup_n B_n$, $\{B_n\}$ is a disjoint sequence, and $B_n \subset A_n$ for each n. So

$$\mu(\cup_n A_n) = \mu(\cup_n B_n) = \sum_{i=1}^{\infty} \mu(B_n) \leq \sum_{i=1}^{\infty} \mu(A_n),$$

by additivity and then montonicity of measure. □

75.4 Limits

Measures also behave well under limits.

An extended-real-valued function on an algebra **resolves under increasing limits** if the result of the union of an increasing sequence of distinguished sets coincides with the limit of the sequence of results on the individual sets. An extended-real-valued function on an algebra **resolves under decreasing**

limits if the result of the intersection of a decreasing sequence of distinguished sets coincides with the limit of the sequence of results on the individual sets.

Proposition 54. *Measures resolve under increasing limits.*

Proof. Let $(A, \mathcal{A}, \mu)$ be a measure space. Let $\{A_n\}$ be an increasing sequence in $\mathcal{A}$. Then we want to show: $\mu(\cup_n A_n) = \lim_{n\to\infty} \mu(A_n)$.

Define $\{B_n\}$ such that $B_n := A_n - \cup_{i=1}^{n-1} A_i$. Then $\{B_n\}$ is disjoint, $A_n = \cup_{i=1}^{n} B_i$ for each n, $\cup_n A_n = \cup_n B_n$, and $\mu(\cup_{i=1}^{n} B_i) = \sum_{i=1}^{n} \mu(B_i)$, by additivity. So

$$\begin{aligned} \mu(\cup_n A_n) &= \mu(\cup_n B_n) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \mu(B_i) \\ &= \lim_{n\to\infty} \mu(\cup_{i=1}^{n} B_i) \\ &= \lim_{n\to\infty} \mu(A_n). \end{aligned}$$

□

Proposition 55. *Measures resolve under decreasing limits if there is a finite set in the decreasing sequence.*

Proof. Let $(A, \mathcal{A}, \mu)$ be a measure space. Let $\{A_n\}$ be a decreasing sequence in $\mathcal{A}$ with one element finite. Then we want to show: $\mu(\cap_n A_n) = \lim_{n\to\infty} \mu(A_n)$.

On one hand, let n_0 be the index of the first finite element of the sequence. Then for all $n \geq n_0$, the sequence is finite because of the monotonicity of measure. Denote this decreasing finite subsequence of sets by $\{B_n\}$. Then $\cap_n A_n = \cap_n B_n$ and $\lim_n A_n = \lim_n B_n$.

On the other hand, the sequence $\{B_1 - B_n\}$ is an increasing sequence in $\mathcal{A}$. Also $\cap_n B_n = B_1 - \cup_n (B_1 - B_n)$. So

$$\begin{aligned}
\mu(\cap_n B_n) &= \mu(B_1 - \cup_n (B_1 - B_n)) \\
&= \mu(B_1) - \mu(\cup_n (B_1 - B_n)) \\
&= \mu(B_1) - \lim_n \mu(B_1 - B_n) \\
&= \mu(B_1) - \left(\lim_n \mu(B_1) - \mu(B_n)\right) \\
&= \lim_n B_n.
\end{aligned}$$

$\square$

76 Measure Space

76.1 Why

We want to generalize the notions of length, area, and volume.

76.2 Definition

A **measurable space** is a sigma algebra. We call the distinguished subsets the **measurable sets**.

A **measure** on a measurable space is a function from the sigma algebra to the positive extended reals. A **measure space** is a measurable space and a measure.

76.2.1 Notation

76.2.2 Properties

Proposition 56. *Let $(A, \mathcal{A})$ be a measurable space and $m : \mathcal{A} \to [0, \infty]$ be a measure.*

If $B \subset C \subset A$, then $m(B) \leq m(C)$. We call this property the of measures ***monotonicity of measure****.*

Proposition 57. *For a measure space $(A, \mathcal{A}, m)$.*

If $B \subset C \subset A$, then $m(B) \leq m(C)$.

*We call this property the **monotonicity of measure**.*

Proposition 58. *For a measure space $(A, \mathcal{A}, m)$.*

If $\{A_n\} \subset \mathcal{A}$ a countable family, then $m(\cup A_n) \leq \sum_i m(A_i)$.

*We this property the **sub-additivty of measure**.*

Proposition 59. *For a measure space $(A, \mathcal{A}, m)$.*

If $\{A_n\} \subset \mathcal{A}$ a countable family, then $m(\cup A_n) \leq \sum_i m(A_i)$.

*We this property the **sub-additivty of measure**.*

Proposition 60. *For a measure space $(A, \mathcal{A}, m)$.*

$$m(\cup_{n=1}^{\infty} A_i) = \lim_{n\to\infty} m(A_i)$$

Proposition 61. *For a measure space $(A, \mathcal{A}, m)$.*

$$m(\cap_{n=1}^{\infty} A_i) = \lim_{n\to\infty} m(A_i)$$

76.2.3 Examples

Example 62. ***counting measure***

77 Measurable Functions

77.1 Why

We define integrals using an infinite process; in order for each step of the process to make sense, need functions to be measurable. Maybe: point to simple functions so that the why is clear.

77.2 Definition

A function between the base sets of two measurable spaces is **measurable** with respect to the distinguished sets of the two spaces if the inverse image of every distinguished subset of the codomain is a distinguished subset of the domain.

77.2.1 Notation

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. Then a function $f : X \to Y$ is measurable if $B \in \mathcal{B}$ implies $f^{-1}(B) \in \mathcal{A}$. We say that f is measurable with respect to $\mathcal{A}$ and $\mathcal{B}$.

In this case, we sometimes say f is a measurable function from $(X, \mathcal{A})$ to $(Y, \mathcal{B})$. We say, $f : (X, \mathcal{A}) \to (Y, \mathcal{B})$ is measurable, read aloud as “f from X, A to Y, B is measurable.”

78 Measurable Function Operations

78.1 Why

Under which operations is the set of measurable functions closed?

78.2 Overview

Measurable functions are closed under composition and concatenation. Taking these facts together with the observation that continuous functions are measurable, we conclude that measurable functions are closed under addition, multiplication, and (with suitable nonzero assumptions) division.

78.3 Results

Proposition 63. *The composition of two measurable functions is measurable.*

Proof. Let $(X, \mathcal{A})$, $(Y, \mathcal{B})$, $(Z, \mathcal{C})$ be measurable spaces. Let $f : X \to Y$ and $g : Y \to Z$ be measurable functions. Define $h = g \circ f$.

Let $C \in \mathcal{C}$. Measurability of g implies $g^{-1}(C) \in \mathcal{B}$. This fact, together with measurability of f, implies $f^{-1}(g^{-1}(C)) \in \mathcal{A}$. Since $h^{-1} = f^{-1} \circ g^{-1}$, we conclude that h is measurable. $\square$

Proposition 64. *A continuous function between topological spaces is measurable with respect to the topological sigma algebras.*

Proposition 65. *The concatentation of two measurable functions is measurable.*

Proof. Let $(X, \mathcal{A})$, $(Y, \mathcal{B})$ and $(Z, \mathcal{C})$ be measurable spaces. Let $f : X \to Y$ and $g : X \to Z$ be measurable. Define $h : X \to Y \times Z$ by $h(x) = (f(x), g(x))$. $\square$

Proposition 66. *Let $(X, \mathcal{A})$ be a measurable space and let R denote the real numbers. Let $f, g : X \to R$ be measurable.*

Then, $f + g$ and fg are measurable.

79 Almost Everywhere

79.1 Why

We treat properties failing on a set of measure zero as though they occur everyhwere; especially in discussions of convergence.

79.2 Definition

A subset of the base set of a measure space is **negligible** if there exists a measurable set with measure zero containing the subset. Negligible sets need not be measurable.

A property holds **almost everywhere** with respect to a measure on a measure space if the set of elements of the base set on which the property does not hold is negligible.

If the property holds everywhere, it holds almost everywhere. In this sense we call the almost everywhere sense "weaker" than the everywhere sense.

79.2.1 Notation

Let $(X, \mathcal{A}, \mu)$ be a measure space. A set $N \subset X$ is negligible if there exists $A \in \mathcal{A}$ with $N \subset A$ and $\mu(N) = 0$.

We abbreviate almost everywhere as "a.e.," read "almost everywhere". We say that a property "holds a.e." If the measure μ is not clear from context, we say that the property holds almost everywhere $[\mu]$ or μ-a.e., read "mu almost everywhere."

79.3 Examples

Let $(X, \mathcal{A}, \mu)$ be a measure space and let R be the real numbers.

79.3.1 Function Comparisons

Let $f, g : X \to R$ be two functions on X, not necessarily measurable. Then $f = g$ almost everywhere if the set of points at which the functions disagree is μ-negligible. Similarly, $f \geq g$ almost everywhere if the set of points where f is less than g is μ-negligible. If f and g are $\mathcal{A}$-measurable, then the sets

$$\{x \in X \mid f(x) \neq g(x)\} \text{ and } \{x \in X \mid f(x) < g(x)\}$$

are measurable; but they need not be measurable otherwise.

79.3.2 Function Limits

Let $f_n : X \to R$ for each natural number n and let $f : X \to R$ be a function. The sequence $\{f_n\}_n$ **converges to f almost everywhere** if

$$\left\{x \in X \,\middle|\, \lim_n f_n \text{ does not exist, or } f(x) \neq \lim_n f_n\right\}$$

is μ-negligible. In this case, we write " $f = \lim_n f_n$ almost everywhere."

80 Almost Everywhere Measurability

80.1 Why

Does convergence almost everywhere of a sequence of measurable functions guarantee measurability of the limit function? It does on complete measure spaces, and we can use this result to "weaken" the hypotheses of many theorems.

80.2 Results

A measure is **complete** if every subset of a measurable set of measure zero is measurable. If the measure is complete, then every negligible set must be measurable.

We begin with a transitivity property: almost everywhere equality of two functions allows us to infer measurability of one from the other.

Proposition 67. *Let $(X, \mathcal{A}, \mu)$ be a measure space Let $f, g : X \to [-\infty, \infty]$ with $f = g$ almost everywhere. If μ is complete and f is $\mathcal{A}$-measurable, then g is $\mathcal{A}$-measurable.*

Proof. □

Proposition 68. *Let* $(X, \mathcal{A}, \mu)$ *be a measure space. Let* $f_n : X \to [-\infty, \infty]$ *for all natural numbers* n *and* $f : X \to [-\infty, \infty]$ *with* $\{f_n\}_n$ *converging to* f *almost everywhere. If* μ *is complete and and* f_n *is measurable for each* n*, then* f *is* $\mathcal{A}$*-measurable.*

Proof. $\square$

81 Matroids

81.1 Why

We generalize the notion of linear dependence.

81.2 Definition

A **matroid** is a finite subset algebra satisfying:

1. The subset of a distinguished set is distinguished.
2. For two distinguished subsets of nonequal cardinality, there is an element of the base set in the complement of the smaller set in the bigger set whose singleton union with the smaller set is a distinguished set.

An **independent subset** of a matroid is a distinguished subset. A **depenedent subset** of a matroid is an undistinguished subset.

81.2.1 Notation

We follow the notation of subset algebras, but use M for the base set, a mnemonic for matroid, and $\mathcal{I}$ for the distinguished sets, a menomic for independent.

Let $(M, \mathcal{I})$ a matroid. We denote the properties by

1. $A \in \mathcal{I} \wedge B \subset A \implies B \in \mathcal{I}$.
2. $A, B \in \mathcal{I} \wedge |A| < |B| \implies \exists x \in M : (A \cup \{x\}) \in \mathcal{I}$

82 Simple Integrals

82.1 Why

We want to define area under a real function. We begin with functions whose area under the curve is self-evident.

82.2 Definition

Consider a measure space. The characteristic function of any measurable set is measurable. A simple function is measurable if and only if each element of its simple partition is measurable.

The **integral** of a measurable non-negative simple function is the sum of the products of the measure of each piece with the value of the function on that piece. For example, the integral of a measurable characteristic function of a subset is the measure of that subset.

The **integral operator** is the real-valued function which associates each measurable non-negative simple function with its integral. The simple integral is non-negative, so the integral operator is a non-negative function.

82.2.1 Notation

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let R be the set of real numbers.

Let $f : X \to R$ be a measurable simple function. So there exist $A_1, \ldots, A_n \in \mathcal{A}$ and $a_1, \ldots, a_n \in R$ with:

$$f = \sum_{i=1}^{n} a_i \chi_{A_i}.$$

We denote the integral of f with respect to measure μ by $\int f d\mu$. We defined:

$$\int f d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

83 Non-negative Integrals

83.1 Why

We want to define area under an extended real function. We use the infinite process to approximate the area under a non-negative extended real function using simple functions.

83.2 Definition

Consider a measure space.

The **integral** of a measurable nonnegative function is the supremum of integrals over non-negative simple functions pointwise less than or equal to the function.

83.2.1 Notation

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to [0, \infty]$ be measurable. We denote the integral of f with respect to the measure μ by $\int f d\mu$. We defined:

$$\int f d\mu = \sup\left\{ \int g d\mu \;\middle|\; g \in \mathcal{SF}_+(X) \text{ and } g \leq f \right\}.$$

84 Real Integrals

84.1 Why

We define the area under an extended real function.

84.2 Definition

The **positive part** of an extended-real-valued function is the function mapping each element to the maximum of the function's result and zero. The **negative part** of an extended-real-valued function is the function mapping each element to the maximum of the negative of function's result and zero.

We decompose an extended-real-valued function as the difference of its positive part and its negative part. Both the positive and negative parts are non-negative extended-real-valued functions.

Consider a measure space. An **integrable** function is a measurable extended real function for which the non-negative integral of the posititve part and the non-negative integral of the negative part of the function are finite.

The **integral** of an integrable function is the difference of the non-negative integral of the posititive part and and the non-negative integral of the negative part.

If one but not both of the parts of the function are finite, we say that the integral **exists** and again define it as before. in this way we avoid arithmetic between two infinities.

84.2.1 Notation

Let A a non-empty set. Let $g : A \to [-\infty, \infty]$. We denote the positive part of g by g^+ and the negative part of g by g^-:

$$g^+(x) = \max\{g(x), 0\} \quad \text{and} \quad g^-(x) = \max\{-g(x), 0\}.$$

Moreover, we decompose g as $g = g^+ - g^-$. We observed that $g^+(x) \geq 0$ and $g^-(x) \geq 0$ for all $x \in X$.

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to [-\infty, +\infty]$ measurable and one of $\int f^+ d\mu$ or $\int f^- d\mu$ is finite (if both are finite, f is integrable).

We denote the integral of f with respect to the measure μ by $\int f d\mu$. We defined:

$$\int f d\mu = \left(\int f^+ d\mu\right) - \left(\int f^- d\mu\right).$$

85 Simple Integral Homogeneity

85.1 Why

If we stack a rectangle on top of itself we have a rectangle twice the height. The additivity principle says that the area of the so-formed rectangle is the sum of the areas of the stacked rectangles. Our definition of integral for simple functions has this property.

85.2 Result

Proposition 69. *The simple non-negative integral operator is homogenous over non-negative real values.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\mathcal{SF}_+(X)$ denote the non-negative real-valued simple functions on X. Define $s : \mathcal{SF}_+(X) \to [0, \infty]$ by $s(f) = \int f d\mu$ for $f \in \mathcal{SF}_+(X)$.

In this notation, we want to show that $s(\alpha f) = \alpha s(f)$ for all $\alpha \in [0, \infty)$ and $f \in \mathcal{SF}_+(X)$. Toward this end, let $f \in \mathcal{SF}_+(X)$ with the simple partition $\{A_n\} \subset \mathcal{A}$ and $\{a_n\} \subset [0, \infty]$.

First, let $\alpha \in (0, \infty)$. Then $\alpha f \in \mathcal{SF}_+(X)$, with the simple partition $\{A_n\} \subset \mathcal{A}$ and $\{\alpha a_n\} \subset [0, \infty]$.

$$s(\alpha f) = \sum_{i=1}^{n} \alpha a_n \mu(A_i) \quad = \alpha \sum_{i=1}^{n} a_n \mu(A_i) = \alpha s(f).$$

If $\alpha = 0$, then αf is uniformly zero; it is the non-negative simple with partition $\{X\}$ and $\{0\}$. Regardless of the measure of X, this non-negative simple function is zero Recall that we define $0 \cdot \infty = \infty \cdot 0 = 0$. $\square$

86 Simple Integral Additivity

86.1 Why

If we stack a two rectangles, with equal base lengths but different heights, on top of each other, the additivity principle says that the area of the so-formed rectangle is the sum of the areas of the stacked rectangles. Our definition of integral for simple functions has this property.

86.2 Result

Proposition 70. *The simple non-negative integral operator is additive.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\mathcal{SF}_+(X)$ denote the non-negative real-valued simple functions on X. Define $s : \mathcal{SF}_+(X) \to [0, \infty]$ by $s(f) = \int f d\mu$ for $f \in \mathcal{SF}_+(X)$.

In this notation, we want to show that $s(f+g) = s(f)+s(g)$ for all $f, g \in \mathcal{SF}_+(X)$. Toward this end, let $f, g \in \mathcal{SF}_+(X)$ with the simple partitions:

$$\{A_i\}_{i=1}^m, \{B_j\}_{j=1}^n \subset \mathcal{A} \quad \text{and} \quad \{a_i\}_{i=1}^m, \{b_j\}_{j=1}^n \subset [0, \infty].$$

We consider the refinement of the two partitions. TODO: this is why you don't do the unique maximal partition business. $\{A_i \cap B_j\}_{i,j=1}^{i=m,j=n}$.

First, let $\alpha \in (0, \infty)$. Then $\alpha f \in \mathcal{SF}_+(X)$, with the simple partition $\{A_n\} \subset \mathcal{A}$ and $\{\alpha a_n\} \subset [0, \infty]$.

$$s(\alpha f) = \sum_{i=1}^{n} \alpha a_n \mu(A_i) \quad = \alpha \sum_{i=1}^{n} a_n \mu(A_i) = \alpha s(f).$$

If $\alpha = 0$, then αf is uniformly zero; it is the non-negative simple with partition $\{X\}$ and $\{0\}$. Regardless of the measure of X, this non-negative simple function is zero Recall that we define $0 \cdot \infty = \infty \cdot 0 = 0$. □

87 Simple Integral Monotonicity

87.1 Why

If one rectangle contains another rectangle, the area of the first should be larger than the area of the second. Our definition of integral for simple functions carries this property. TODO: area sheet.

87.2 Result

Proposition 71. *The simple non-negative integral operator is monotone.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g \in \mathcal{SF}_+(X)$ with $f \leq g$. Then $f - g \in \mathcal{SF}_+(X)$, so

$$\begin{aligned}\int g d\mu &= \int (f + (g - f)) d\mu \\ &\overset{(a)}{=} \int f d\mu + \int (g - f) d\mu \\ &\overset{(b)}{\geq} \int f d\mu\end{aligned}$$

where (a) follows from linearity and (b) follows from non-negativity; properties of the non-negative simple integral operator.

$\square$

88 Real Integral Monotone Convergence

88.1 Why

An integral is a limit. When can we exchange this limit with another? We give a first result in the search for sufficient conditions to do so.

88.2 Result

When context is clear, we refer to the following proposition as the **monotone convergence theorem**.

Proposition 72. *The integral of the almost everywhere limit of an almost-everywhere nondecreasing sequence of measurable, nonnegative, extended-real-valued functions is the limit of the sequence of integrals of the functions.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $f_n :\to [-\infty, \infty]$ a $\mathcal{A}$-measurable function for every natural number n and let $f : X \to [-\infty, \infty]$ a $\mathcal{A}$-measurable function. We want to show that if

$$f_n(x) \leq f_{n+1}(x) \quad \text{ and } \quad f(x) = \lim_n f_n(x)$$

hold for all natural n and almost every x in X, then

$$\int f d\mu = \lim_n \int f_n d\mu.$$

$\square$

89 Real Integral Series Convergence

89.1 Why

Sums of non-negative functions are increasing, and workable with the monotone convergence theorem.

89.2 Result

Proposition 73. *The integral of the limit of the partial sums of a sequence of measurable, nonnegative, extended-real-valued functions is the limit of the partial sums of the integrals.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $f_n :\to [0, \infty]$ a $\mathcal{A}$-measurable function for every natural number n. We want to show that:

$$\int \sum_{k=1}^{\infty} f_k d\mu = \sum_{k=1}^{\infty} \int f_k d\mu.$$

We apply the monotone convergence theorem to the sequence $\{\sum_{i=1}^{n} f_i\}_n$. This sequence is nondecreasing because $f_n \geq 0$ for all n. $\square$

90 Real Integral Limit Inferior Bound

90.1 Why

TODO

90.2 Result

Proposition 74. *The integral of the limit inferior of a sequence of measurable, nonnegative, extended-real-valued functions is no larger than the limit inferior of the sequence of integrals.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $f_n :\to [0, \infty]$ a $\mathcal{A}$-measurable function for every natural number n. We want to show that if

$$\int \liminf_n f_n d\mu \leq \liminf_n \int f_n d\mu.$$

□

91 Real Integral Dominated Convergence

91.1 Why

An integral is a limit. When can we exchange this limit with another? We give a first result in the search for sufficient conditions to do so.

91.2 Result

When context is clear, we refer to the following proposition as the **dominated convergence theorem**.

Proposition 75. *The integral of the almost everywhere limit of a sequence of measurable, extended-real-valued, almost-everywhere bounded functions is the limit of the sequence of integrals of the functions.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to [-\infty, \infty]$ be a $\mathcal{A}$-measurable function. Let $f_n :\to [-\infty, \infty]$ a $\mathcal{A}$-measurable function for every natural number n so that $\{f_n\}_n$ converges almost everywhere to f. Let $g : X \to [0, \infty]$ be an integrable function which dominates f_n almost everywhere for each n. We want to show that:

$$\int f d\mu = \lim_n \int f_n d\mu.$$

$\square$

92 Real Integral Limit Theorems

92.1 Why

A vista sheet on exchanging integrals and limits.

92.2 Discussion

The monotone convergence theorem and the dominated convergence theorem give conditions under which we can exchange a limit with an integral. Since an integral is a limit, these theorems give conditions under which we can exchange limits.

We remember in exchanging these limits that we are really exchanging infinite processes. It is no suprise that this exchange does not always make sense. Nonetheless, the dominated convergence theorem gives a simple criteria which is easy and guarantees the desired exchange is valid.

TODO: Structure of the theory. Building from the monotone convergence for simple functions through to the dominated convergence.

93 Image Measures

93.1 Why

A measurable function from a first measure space to a second measurable space induces a measure on the latter.

93.2 Definition

Consider two measurable spaces and a measurable function between them. The **image measure** of a measure on the first space **under** the measurable function is the measure on the second space which assigns to each measurable set the measure of the inverse image of that measurable set.

We say that the function **induces** the image measure on the codomain. Alternatively, we say that we **push forward** the measure to the codomain, and so call the image measure a **push forward measure**.

93.2.1 Notation

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be two measurable spaces. Let $f : X \to Y$ be a measurable function. Let $\mu : \mathcal{A} \to [0, \infty]$ be a measure. We denote the image measure of μ under f by $\mu \circ f^{-1}$, for the reason that it

$$\mu \circ f^{-1}(B) = \mu(f^{-1}(B))$$

for every $B \in \mathcal{B}$.

93.3 Change of Variables

The main property we would like to hold is that integration on the new measure space is the same as integration on the old.

Proposition 76. *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces and let R denote the real numbers. Let $f : X \to Y$ be a measurable function and let $\mu : \mathcal{A} \to [0, \infty]$ be a measure.*

Then $g : Y \to R$ is integrable with respect to $\mu \circ f^{-1}$ if and only if $g \circ f$ is integrable with respect to μ. In this case,

$$\int g d(\mu \circ f^{-1}) = \int g \circ f d\mu.$$

Proof. □

94 Functionals

94.1 Why

We speak of maps from a vector space to a field of scalars.

94.2 Definition

A **functional** is a function from a set of vectors to a field. It is natural, and common, for the field of scalars to be the base field.

A real-valued functional is **non-negative** if its range is a subset of the non-negative real numbers. A real-valued functional if **definite** if the only it maps to zero is the zero element of the vector space.

A real-valued functional on a real or complex vector space is **absolutely homogeneous** if the result of a scaled vector is the same as the result of the vector scaled by the absolute value of the scalar.

95 Signed Measures

95.1 Why

Can we view the set of measures as a vector space?

Not quite: the difference of two measures may take negative values on some set. This functional will be countably additive, however, and so behaves similar to a measure.

95.2 Definition

An extended-real-valued function on a sigma algebra is **countably additive** if the result of the function applied to the union of a disjoint countable family of distinguished sets is the limit of the partial sums of the results of the function applied to each of the sets individually. The limit of the partial sums must exist irregardless of the summand order.

A **signed measure** is an extended-real-valued function on a sigma algebra that is (1) zero on the empty set and (2) countably additive. We call the result of the function applied to a set in the sigma algebra the **signed measure** (or when no ambiguity arises, the **measure**) of the set.

When speaking of a measure, which is non-negative, in contrast to a signed measure, we will call the former a **positive measure**.

95.2.1 Notation

Let $(X, \mathcal{A})$ be a measurable space and let $\mu : \mathcal{A} \to [-\infty, \infty]$. Then μ is a signed measure if

1. $\mu(\varnothing) = 0$ and
2. $\mu(\cup_i A_i) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(A_k)$ for all disjoint $\{A_n\}_n$.

Proposition 77. *A signed measure never takes both positive infinity and negative infinity.*

Proof. Let $(X, \mathcal{A})$ be a measurable space. Let $\mu : \mathcal{A} \to [-\infty, \infty]$ be a signed measure. First, suppose $\mu(X)$ is finite, Then by Proposition 78 μ is finite for each $A \in \mathcal{A}$.

Suppose $\mu(X) = \infty$. Let $A \in \mathcal{A}$. As before, $\mu(X) = \mu(A) + \mu(X - A)$. Since $\mu(X) = +\infty$, then both of $\mu(A)$ and $\mu(X - A)$ must be either finite or $+\infty$. Argue similarly for $\mu(X) = -\infty$. $\square$

96 Finite Signed Measures

96.1 Why

For the difference of two (signed) measures to be well-defined, we need one of the two to be finite. Otherwise, the measure of the difference on the base set involves subtracting ∞ from ∞.

96.2 Definition

A **finite** signed measure is one for which the measure of every set is finite. This condition is equivalent to the base set having finite measure (see below).

96.3 Result

Proposition 78. *A signed measure is finite if and only if it is finite on the base set.*

Proof. Let $(X, \mathcal{A})$ be a measurable space. Let $\mu : \mathcal{A} \to [-\infty, \infty]$ be a signed measure. ($\Rightarrow$) If μ is finite, then $\mu(X)$ is finite since $X \in \mathcal{A}$. ($\Leftarrow$) Next, suppose $\mu(X)$ is finite. Let $A \in \mathcal{A}$. Then $X = A \cup (X - A)$, with these sets disjoint, so by countable additivity of μ, $\mu(X) = \mu(A) + \mu(X - A)$. Since $\mu(X)$ finite, $\mu(A)$ and $\mu(X - A)$ are both finite. $\square$

97 Measure Vector Space

97.1 Why

If both signed measures are finite, then their difference is always well-defined. Is the difference a finite signed measure?

97.2 Preliminary Result

Proposition 79. *A linear combination of finite signed measures is a finite signed measure.*

Proof. Let $(X, \mathcal{A})$ be a measurable space. Let μ and ν be finite signed measures. Let R denote the real numbers. Then $(\alpha\mu)(\varnothing) = \alpha \cdot \mu(\varnothing) = \alpha \cdot 0 = 0$. Also for $\{A_n\}_n \subset \mathcal{A}$ disjoint,

$$\begin{aligned}(\alpha\mu)(\cup A_n) &= \alpha\mu(\cup A_n) = \alpha \sum_{n=1}^{\infty} \mu(A_n) \\ &= \sum_{n=1}^{\infty} \alpha\mu(A_n) = (\alpha\mu)(A_n)\end{aligned}$$

.

Similarly, $(\mu + \nu)(\varnothing) = \mu(\varnothing) + \nu(\varnothing) = 0$. And, for $\{A_n\}_n \subset$

$\mathcal{A}$ disjoint,

$$(\mu+\nu)(\cup A_n) = \mu(\cup A_n) + \nu(\cup A_n) = \sum_{n=1}^{\infty} \mu(A_n) + \sum_{n=1}^{\infty} \nu(A_n)$$
$$= \sum_{n=1}^{\infty} \mu(A_n) + \nu(A_n) = \sum_{n=1}^{\infty} (\mu+\nu)(A_n)$$

. □

97.3 Main Result

Proposition 80. *The set of finite signed measures is a vector space.*

Proof. Use the previous proposition. Observe that the function $\mu \equiv 0$ is a measure. and $\nu + \mu = \nu$ for all measures ν. □

97.3.1 Notation

Denote the set of real numbers by R. We denote the vector space of signed measures on measurable space $(X, \mathcal{A})$ by by $M(X, \mathcal{A}, R)$.

98 Signed Set Decomposition

98.1 Why

Are all signed measures the difference of two positive measures?

Suppose we could partition the base set into two sets, one containing all the sets with positive measure and one containing all sets with negative measure. We could "restrict" the measure to the former and it would be positive, and we could "restrict" it to the latter and it would be negative.

Any measurable set could be partitioned into a piece in the former and a piece in the latter, and so its signed measure could be written as a sum of measures of these pieces.

98.2 Definition

By "positive" and "negative" we mean "non-negative" and "non-positive." Let $(X, \mathcal{A})$ be a measurable space. Let $\mu : \mathcal{A} \to [-\infty, \infty]$ be a signed measure.

A **positive set** is a measurable set with the property that each of its subsets have non-negative measure under μ. A **negative set** is a measurable set with the property that each of its subsets have non-positive measure under μ.

A **signed-set decomposition** of X under μ is a partition of X into a positive and a negative set. Some authors call it a **Hahn decomposition**.

98.2.1 Notation

Denote by P a positive and by N a negative set. When we say "let (P, N) be a signed-set decomposition of X under μ", we mean that P is the positive set and N is the negative set.

98.3 Motivating Implication

Does such a decomposition always exist? Is it unique? We are motivated to find answers by the following observation.

Suppose there was a signed-set decomposition of $(X, \mathcal{A})$ under μ; denote it by (P, N). Then $\mu(A \cap P) \geq 0$ and $\mu(A \cap N) \leq 0$ for all $A \in \mathcal{A}$.

Define $\mu_1 = \mu(A \cap P)$ and $\mu_2 = -\mu(A \cap P)$, then μ_1 and μ_2 are finite measures. Moreover, $\mu = \mu_1 - \mu_2$. Thus, if we had a signed-set decomposition we could write μ as the difference of two measures.

99 Signed-Set Decomposition Existence

99.1 Why

Does a signed-set decomposition exist for any signed measure?

99.2 Result

The answer is yes.

Proposition 81. *Let $(X, \mathcal{A})$ be a measurable space. Let $\mu : \mathcal{A} \to [-\infty, \infty]$ be a signed measure. There exists a signed-set decomposition of X under μ.*

Proof. TODO $\square$

99.2.1 Uniqueness

100 Product Sigma Algebras

100.1 Why

We want to generalize the construction of cover area as generated as a product of two cover lengths, and more generally for arbitrary measure spaces. TODO

100.2 Definition

Consider two measurable spaces. The **product base set** is the cartesian product of the first base set with the second base set. A first distinguished set is a distinguished set of the first measurable space, and likewise for a second distinguished set.

A **rectangle with measurable sides** is a set in product base set which is a product of a first distinguished set with a second distinguished set. The **product sigma algebra** is the sigma algebra generated by the rectangles with measurable sides.

The **product measurable space** is the measurable space whose base set is the product base set and whose sigma algebra is the product sigma algebra.

100.2.1 Notation

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. The product base set is $X \times Y$. A set $R \in X \times Y$ is a rectangle with measurable sides if $R = A \times B$ for $A \in \mathcal{A}$ and $B \in \mathcal{B}$. We denote the product sigma algebra of $\mathcal{A}$ and $\mathcal{B}$ by $\mathcal{A} \times \mathcal{B}$. $(X \times Y, \mathcal{A} \times \mathcal{B})$ is the product measurable space.

101 Product Sections

101.1 Why

Toward a theory of iterated integrals, we need to generalize rectangular strips to arbitrary products.

101.2 Definition

Consider the product of two non-empty sets.

First, consider a subset of this product. For a specified element in the first set, the **set section** of the subset with respect to that element is the set of elements in the second set for which the ordered pair of the specified element and that element is in the subset; the section is a subset of the second set. For elements of the second set, we define sections similarly.

Second, consider a function on the product. For a specified element in the first set, the **function section** of the function for that element is the function from the second set to the codomain of the function which maps elements of the second set to the result of the function applied to the ordered pair of the specified element and the element of the second set. For elements of the second set, we define sections similarly.

101.2.1 Notation

Let X, Y be non-empty sets.

Let $E \subset X \times Y$. For $x \in X$, we denote the section of E with respect to x by E_x. For $y \in Y$, we denote the section of E with respect to x by E^y. For every $x \in X$ and $y \in Y$,

$$E_x = \{y \in Y \mid (x, y) \in E\} \quad \text{and} \quad E^y = \{x \in X \mid (x, y) \in E\}.$$

$E_x \subset Y$ and $E_y \subset X$.

Let $f : X \times Y \to Z$. For $x \in X$, we denote the section of f with respect to x by $f_x : Y \to Z$. For $y \in Y$, we denote the section of f with respect to x by $f^y : X \to Z$. For every $x \in X$ and $y \in Y$,

$$f_x(y) = f(x, y) \quad \text{and} \quad f^y(z) = f(x, y).$$

102 Measurable Sections

102.1 Why

Toward a theory of iterated integrals, we need to know that set and function sections are measurable.

102.2 Results

Proposition 82. *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. For any $E \in \mathcal{A} \times \mathcal{B}$, the sections E_x and E^y are measurable for any $x \in X$ and $y \in Y$.*

Proof. TODO □

Proposition 83. *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. Let $f : X \times Y \to F$, where F is the extended real numnbers or the complex numbers, and f is measurable (using the appropriate sigma algebra of the codomain). The sections $f_x : Y \to F$ and $f^y : X \to F$ are measurable for each $x \in X$ and $y \in Y$.*

Proof. TODO □

103 Sections Measures

103.1 Why

Toward a theory of iterated integrals, we need to know the function measuring a section is integrable.

103.2 Results

Proposition 84. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be sigma-finite measurable spaces. Let $E \in \mathcal{A} \times \mathcal{B}$. The function $x \mapsto \nu(E_x)$ is $\mathcal{A}$-measurable and the function $y \mapsto \mu(E^y)$ is $\mathcal{B}$-measurable.*

Proof. TODO $\square$

104 Product Measures

104.1 Why

We want to generalize the construction of cover area as generated as a product of two cover lengths, and more generally for arbitrary measure spaces. TODO

104.2 Definition

The **product measure** of the measures of two sigma finite measure spaces is the unique measure which assigns to every rectangle with measurable sides the product of the measures of the sides. We prove that such a measure exists, and is unique.

104.2.1 Defining Result

Proposition 85. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be sigma-finite measurable spaces. There is a unique measure π on $\mathcal{A} \times \mathcal{B}$ such that*

$$\pi(A \times B) = \mu(A) \times \nu(B)$$

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Furthermore, for any $E \in \mathcal{A} \times \mathcal{B}$.

$$\pi(E) = \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\nu(y).$$

Proof. TODO □

104.2.2 Notation

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be sigma-finite measurable spaces. We denote the product measure by $\mu \times \nu$. For all $A \in \mathcal{A}$ and $B \in \mathcal{B}$,

$$(\mu \times \nu)(A \times B) = \mu(A) \times \nu(B).$$

105 Norms

105.1 Why

We want to measure the size of an element in a vector space.

105.2 Definition

A **norm** is a real-valued functional that is (a) non-negative, (b) definite, (c) absolutely homogeneous, (d) and satisifies a triangle inequality. The triangle inequality property requires that the norm applied to the sum of any two vectors is less than the sum of the norms.

105.2.1 Examples

Example 86. *The absolute value function is a norm on the vector space of real numbers.*

Example 87. *The Euclidean distance is a norm on the various real spaces.*

105.2.2 Notation

Let (X, F) be a vector space where F is the field of real numbers or the field of complex numbers. Let R denote the set of real numbers. Let $f : X \to R$. The functional f is a norm if

1. $f(v) \geq 0$ for all $x \in V$
2. $f(v) = 0$ if and only if $x = 0 \in X$.
3. $f(\alpha x) = |\alpha| f(x)$ for all $\alpha \in F$, $x \in X$
4. $f(x + y) \leq f(x) + f(y)$ for all $x, y \in X$.

In this case, for $x \in X$, we denote $f(x)$ by $|x|$, read aloud "norm x". The notation follows the notation of absolute value as a norm. When we wish to distinguish the norm from the absolute value function, we may write $\|x\|$. In some cases, we go further, and for a norm indexed by some parameter α or set A we write $\|x\|_\alpha$ or $\|x\|_A$.

106 Integrable Function Space

106.1 Why

The integrable functions are a vector space.

106.2 Definition

The **integrable function space** corresponding to a measure space is the set of real-valued functions which are integrable with respect to the measure. The term space is appropriate because this set is a real vector space. If we scale an integrable function, it remains integrable. If we add two integrable functions, the sum is integrable. Thus, a linear combination of integrable functions is integrable. The zero function is the zero element of the vector space.

TODO The open question is: what elements of our geometric intuition can we bring to a space of functions. Do functions have a size? Are certain functions near each other?

106.2.1 Notation

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let R denote the set of real numbers and let C denote the set of complex numbers.

We denote set the real-valued integrable functions on X by $\mathcal{I}(X, \mathcal{A}, \mu, R)$, read aloud as "the real integrable functions on the measure space X script A mu." We denote set the complex-valued integrable functions on X by $\mathcal{I}(X, \mathcal{A}, \mu, C)$, read aloud as "the complex integrable functions on the measure space X script A mu." When the field is irrelevant, we denote them by $\mathcal{I}(X, \mathcal{A}, \mu)$, read aloud as "integrable functions on the measure space X script A mu." The $\mathcal{I}$ is a mnemonic for "integrable."

107 Supremum Norm

107.1 Why

We want a norm on the vector space of continuous functions.

107.2 Definition

Consider a function from a closed real interval to the real numbers. The **absolute supremum** of the function is the absolute value of its results on the interval. Since the function is continuous and defined on a closed interval, the supremum is finite.

Proposition 88. *The functional mapping $f \in C[a, b]$ to its absolute supremum is a norm.*

Proof. Let R denote the set of real numbers. Define $\phi : C[a, b] \to R$ by:

$$\phi(f) = \sup\{|f(x)| \mid x \in [a, b]\}.$$

1. $|f(x)| \geq 0$ for all $x \in [a, b]$, so $\phi(f) \geq 0$.

2. If $\phi(f) = 0$ then $|f(x)| \leq 0$ for all x and so $f(x) = 0$ for all $x \in [a, b]$. If $f = 0$, then $|f(x)| = 0$ for all $x \in [a, b]$

3. For all α real, $|\alpha f(x)| = |\alpha||f(x)|$. so $\phi(\alpha f) = |\alpha|\phi(f)$

4. For all $f, g \in C[a, b]$, and $x \in [a, b]$, $|f(x)+g(x)| \leq |f(x)| + |g(x)|$ by the triangle inequality for absolute value. Thus,

$$\begin{aligned}\phi(f+g) &\leq \sup\{|f(x)| + |g(x)| \mid x \in [a, b]\} \\ &\leq \sup\{|f(x)| \mid x \in [a, b]\} + \sup\{|g(x)| \mid x \in [a, b]\} \\ &= \phi(f) + \phi(g)\end{aligned}$$

$\square$

We call the functional ϕ defined above the **supremum norm**.

107.2.1 Notation

Let $f \in C[a, b]$. We denote the supremum norm of f by $|f|_{\text{sup}}$.

108 Supremum Norm Complete

108.1 Why

We want a complete norm on the vector space of continuous functions.

108.2 Result

Proposition 89. *The supremum norm is complete.*

Proof. Let R denote the real numbers. Let $\{f_n\}_n$ be an egoprox sequence in $C[a, b]$.

Candidate. $\{f_n\}_n$ is egoprox means $\forall \epsilon > 0, \exists N$ so that

$$m, n > N \implies |f_n - f_m|_{\text{sup}} < \epsilon.$$

Since $|f_n - f_m|_{\text{sup}} < \epsilon \implies |f_n(x) - f_m(x)| < \epsilon$ for all $x \in [a, b]$, the sequence of real numbers $\{f_n(x)\}_n$ is egoprox for each $x \in [a, b]$. Since the metric space $(R, |\cdot|)$ is complete, there is a limit $l_x \in R$ such that $f_n(x) \longrightarrow l_x$ as $n \longrightarrow \infty$, for each $x \in [a, b]$. Define $f : [a, b] \to R$ by $f(x) = l_x$ for each $x \in [a, b]$.

Candidate is Limit. First, we argue that $|f_n - f|_{\text{sup}} \longrightarrow 0$ as $n \longrightarrow \infty$. Since $\{f_n\}_n$ is an egoprox sequence, there exists n_0

so that

$$n, m \geq n_0 \implies |f_n - f_m|_{\sup} < \epsilon/2.$$

So for all $x \in [a, b]$,

$$n, m \geq n_0 \implies |f_n(x) - f_m(x)| < \epsilon/2.$$

For all $x \in [a, b]$, and $n \geq n_0$,

$$\lim_{m \to \infty} |f_n(x) - f_m(x)| \leq \epsilon/2 < \epsilon.$$

The sequence $\{f_k(x)\}_{k=m}^{\infty}$ is a final part of $\{f_k(x)\}_{k=1}^{\infty}$, and so has the same limit, $f(x)$. Therefore, using continuity of subtraction and the absolute value,

$$\lim_{m \to \infty} |f_n(x) - f_m(x)| = |f_n(x) - f(x)|.$$

We conclude that for $n \geq n_0$, $x \in [a, b]$, $|f_n(x) - f(x)| < \epsilon$, from which we deduce $|f_n - f|_{\sup} < \epsilon$. Thus $f_n \longrightarrow f$ as $n \longrightarrow \infty$.

Limit is Continuous. Next, we argue that f is continuous. Let $x_0 \in [a, b]$. Let $\epsilon > 0$. Since $f_n \longrightarrow f$ there exists n_0 so that

$$|f_{n_0} - f|_{\sup} < \epsilon/3.$$

By the triangle inequality,

$$|f(x_0) - f(x)| \leq |f(x_0) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x)|,$$

for all $x \in [a, b]$. Using $|f(x_0) - f_{n_0}(x_0)| < \epsilon/3$,

$$|f(x_0) - f(x)| < \epsilon/3 + |f_{n_0}(x_0) - f(x)|,$$

for all $x \in [a, b]$. Using the triangle inequality,

$$|f(x_0) - f(x)| < \epsilon/3 + |f_{n_0}(x_0) - f_{n_0}(x)| + |f_{n_0}(x) - f(x)|$$

for all $x \in [a,b]$. Using $|f_{n_0}(x_0) - f(x)| < \epsilon/3$

$$|f(x_0) - f(x)| < \epsilon/3 + |f_{n_0}(x_0) - f_{n_0}(x)| + \epsilon/3$$

for all $x \in [a,b]$. Since f_{n_0} is continuous, there exists $\delta > 0$ so that

$$|x_0 - x| < \delta \implies |f_{n_0}(x_0) - f_{n_0}(x)| < \epsilon/3,$$

for $x \in [a,b]$. In this case,

$$|f(x_0) - f(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

Since ϵ was arbitrary, f is continuous at x_0. Since x_0 was arbitrary, f is continuous everywhere. Some call the above the **three epsilon argument**. $\square$

109 Complex Measures

109.1 Why

We allow measures to take complex values. TODO

109.2 Definition

A complex-valued function on a sigma algebra is **countably additive** if the result of the function applied to the union of a disjoint countable family of distinguished sets is the limit of the partial sums of the results of the function applied to each of the sets individually. The limit of the partial sums must exist irregardless of the summand order.

A **complex measure** is an complex-valued function on a sigma algebra that is (1) zero on the empty set and (2) countably additive. We call the result of the function applied to a set in the sigma algebra the **complex measure** (or when no ambiguity arises, the **measure**) of the set.

Since the codomain of a complex measures is the complex numbers, the sum corresponding to every countable union must be absolutely convergent (?) (TODO: define).

109.2.1 Notation

We denote complex measures by μ a mnemonic for "measure." Let C denote the set of complex numbers. Let $(X, \mathcal{A})$ be a measurable space and let $\mu : \mathcal{A} \to C$. Then μ is a complex measure if

1. $\mu(\varnothing) = 0$ and
2. $\mu(\cup_i A_i) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(A_k)$ for all disjoint families $\{A_n\}_n$.

110 Sigma Algebra Independence

110.1 Why

TODO

110.2 Definition

A **sub sigma algebra** is a subset of a sigma algebra which is itself a sigma algebra.

A collection of sub sigma algebras is **independent** if the measure of every set which is an intersection of distinguished sets from the sub sigma algebras is the product of the individual measures of the sets. An arbitrary (not finite) collection of sigma algebras is **independent** if any finite sub-collection is independent.

110.2.1 Notation

Let $(X, \mathcal{A}, \mu)$ be a probability space. Let $Y_1, \dots, Y_n$ be sub-sigma-algebras of X. Then $\{Y_1, \dots, Y_n\}$ are independent if for all $A_1 \in Y_1, A_2 \in A_2, \dots, A_n \in Y_n$,

$$\mu\left(\cap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mu(A_i).$$

In this case we write $Y_1 \perp Y_2 \perp \cdots \perp Y_n$.

111 Event Independence

111.1 Why

TODO

111.2 Definition

The sigma algebra **generated by an event** is the sigma algebra consisting of the empty set, the event, the complement (in the base set) of the event, and the base set.

A family of events events are **independent** if the sigma algebras generated by the events are independent.

111.2.1 Notation

Let $(X, \mathcal{A}, \mu)$ be a probability space. Let $A \in \mathcal{A}$ be an event. The sigma algebra generated by A is $\{\varnothing, A, X - A, X\}$. We denote it by $\sigma(A)$.

Let $B \in \mathcal{A}$. If A is independent of B we write $A \perp B$.

111.3 Equivalent Condition

Proposition 90. *Two events are independent if and only if the measure of their intersection is the product of their measures.*

Proof. Let $(X, \mathcal{A}, \mu)$ be a probability space. Let $A, B \in \mathcal{A}$.

($\Rightarrow$) If $A \perp B$, then by definition $A \in \sigma(A)$ and $B \in \sigma(B)$ and so:

$$\mu(A \cap B) = \mu(A)\mu(B).$$

($\Leftarrow$) Conversely, let $a \in \sigma(A)$ and $b \in \sigma(B)$. If $a = \varnothing$ or $b = \varnothing$ then $a \cap b = \varnothing$. So

$$\mu(a \cap b) = \mu(\varnothing) = \mu(a)\mu(b),$$

since one of the two measures on the right hand side is zero. On the other hand, if $a = X$, then $a \cap b = b$ and so

$$\mu(a \cap b) = \mu(b) = \mu(a)\mu(b),$$

since $\mu(a) = \mu(X) = 1$. Likewise if $b = X$.

So it remains to verify $\mu(a \cap b) = \mu(a)\mu(b)$ for the cases $a \in \{A, X - A\}$ and $b \in \{B, X - B\}$. If $a = A$, and $b = B$, then the identity follows by hypothesis. Next, observe that $A \cap (X - B) = A - (A \cap B)$ and $(A \cap B) \subset A$ so $\mu(X) < \infty$ allows us to deduce:

$$\begin{aligned}
\mu(A \cap (X - B)) &= \mu(A - (A \cap B)) \\
&= \mu(A) - \mu(A \cap B) \\
&= \mu(A)(1 - \mu(B)) \\
&= \mu(A)\mu(X - A).
\end{aligned}$$

Similar for $X - A$ and B. Finally, recall that $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$. So then,

$$\begin{aligned}
\mu((X - A) \cap (X - B)) &= 1 - \mu(A \cup B) \\
&= 1 - \mu(A) - \mu(B) + \mu(A \cap B) \\
&= 1 - \mu(A) - \mu(B) + \mu(A)\mu(B) \\
&= (1 - \mu(A))(1 - \mu(B)) \\
&= \mu(X - A)\mu(X - B).
\end{aligned}$$

$\square$

112 Random Variable Sigma Algebra

112.1 Why

What does it mean for two random variables to be independent? What are the events associated with a random variable? TODO

112.2 Definition

112.2.1 Defining Result

Proposition 91. *The set of inverse images the distinguished sets of a measurable space under a function from a set to that space is a sigma algebra.*

If the first set and the function are measurable, the sigma algebra is a sub sigma algebra of the domain sigma algebra.

Proof. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. Let $f : X \to Y$ be a measurable function. Define

$$\mathcal{C} := \left\{ f^{-1}(B) \;\middle|\; B \in \mathcal{B} \right\}.$$

First, since $f^{-1}(Y) = X$, $X \in \mathcal{C}$.

Second, let $C \in \mathcal{C}$. Then there is B such that $C = f^{-1}(B)$. Then $X - C = X - f^{-1}(B) = f^{-1}(Y - B)$. Since $\mathcal{B}$ is a sigma algebra, $B \in \mathcal{B}, Y - B \in \mathcal{B}$ and so $X - C \in \mathcal{C}$.

Finally, let $\{C_n\}_n \subset \mathcal{C}$. Then for every n there exists a $B_n \in \mathcal{B}$ so that $C_n = f^{-1}(B_n)$. Then:

$$\cup_n C_n = \cup_n f^{-1}(B_n) = f^{-1}(\cup B_n).$$

Since $\mathcal{B}$ is a sigma algebra, $\cup_n B_n \in \mathcal{B}$ and so $\cup_n C_n \in \mathcal{C}$.

Since f is measurable, $f^{-1}(B) \in \mathcal{A}$ for every $B \in \mathcal{B}$, and so $\mathcal{C} \subset \mathcal{A}$. $\square$

The sigma algebra **generated by a random variable** is the sigma algebra consisting of the inverse images of every measurable set of the codomain.

The sigma algebra generated by a family of random variables is the sigma algebra generated by the union of the sigma algebras generated individually by each of the random variables.

112.2.2 Notation

Let $(X, \mathcal{A}, \mu)$ be a probability space and $(Y, \mathcal{B})$ be a measurable space. Let $f : X \to Y$ be a random variable. Denote by $\sigma(f)$ the sigma algebra generated by f.

112.3 Results

Proposition 92. *The sigma algebra generated by a family of random variables is the smallest sigma algebra for with respect* *to which each random variable is measurably.*

113 Random Variable Independence

113.1 Why

What does it mean for two random variables to be independent? What are the events associated with a random variable? TODO

113.2 Definition

A family of random variables are **independent** if the sigma algebras generated by the random variables are independent.

113.2.1 Notation

Let $(X, \mathcal{A}, \mu)$ be a probability space and $(Y, \mathcal{B})$ be a measurable space. Let $f_1, f_2 : X \to Y$ be a random variables. If the random variables are independent we write $f_1 \perp f_2$.

113.2.2 Results

Proposition 93. *Let $f_1, \dots, f_n$ be independent real-valued random variables defined on a probability space $(X, \mathcal{A}, \mu)$.*

Let $B_1, \ldots, B_n$ be Borel sets of real numbers and let $A_i = f_i^{-1}(B_i)$. Let $A = \cap_{i=1}^n f_i^{-1}(B_i)$. Then

$$\mu(A) = \prod_{i=1}^{n} \mu(A_i)$$

Proof. Since f_i are independent, so are the sigma algebras they generate. A_i are in each of these sigma algebras, so by definition of independence the measure of the intersection is the product of the measures. □

114 Tail Sigma Algebra

114.1 Why

114.2 Definition

The **tail sigma algebra** of a sequence of random variables is the sigma algebra which is the intersection of the sigma algebras of all final parts of the sequence. A **tail event** is an element of the tail sigma-algebra.

The tail sigma algebra coincides with the sigma algebra generated by the union of the sigma algebras of each of of the random variables.

114.2.1 Notation

Let $\{f_n\}_n$ be a sequence of random variables. Denote the tail sigma algebra by $T(\{f_n\}_n)$. We defined it as:

$$T(\{f_n\}_n) = \bigcap_{n=1}^{\infty} \sigma(\{X_{n+k}\}_k).$$

In other words, for all natural n, the event is in the sigma algebra of the final part of

114.3 Results

Proposition 94. *The tail sigma algebra of a sequence of random variables is the same equals the sigma algebra generated by the union of the sigma algebras of each of the random variables.*

115 Zero-One Law

115.1 Why

115.2 Result

Proposition 95. *Every event in the tail sigma algebra of a sequence of independent random variables defined on the same probability space has probability zero or one.*

Proof. □

116 Probability

116.1 Why

We want to talk about things which may happen.

116.2 Definition

117 Empirical Law

117.1 Why

Suppose we have collected data.

117.2 Definition

Let A be a non-empty set. Let n be a natural number. A **data set** of size n for A is a function from $\{1, \dots, n\}$ into A. It may be that $a_i = a_j$ for some $i \neq j$.

To each data set we associate an **empirical law** which is a probability measure P on the measurable space $(A, 2^A)$ that assigns to each set $B \subset A$ the number

$$P(B) = \frac{|\{i \in [n] \mid a_i \in B\}|}{n},$$