# Real Analysis Structures

William G. Faris

June 27, 2006

# Contents

# Preface

These lectures are an introduction to various structures in real analysis. These include the following:

- Sets and functions
- Ordered sets and order-preserving mappings
- Metric spaces and contraction mappings
- Metric spaces and Lipschitz mappings
- Metric spaces and uniformly continuous mappings
- Topological spaces and continuous mappings
- Measurable spaces and measurable mappings

Measurable spaces are the setting for the integral. The construction of an integral is done starting from an elementary integral defined on a vector lattice of functions. Measure is a special case of integral. The goal is to exhibit the simplicity of this remarkable theory.

The concept of topological space seems much more elegant than the concept of metric space. One thesis of these lectures is that metric spaces are important in their own right. In particular, the notion of complete metric space is crucial. A Polish space is a separable complete metric space, and the exposition tends to focus on Polish spaces rather than on locally compact Hausdorff spaces. Second countable locally compact Hausdorff spaces are Polish spaces. An infinite dimensional separable Banach space is always a Polish space and is never locally compact.

In general, topology and measure coexist somewhat uneasily. However, measure theory is much simpler in the case of Polish spaces. One highlight is a remarkable uniqueness result. While Polish topological spaces include a huge variety of spaces that occur in analysis, the measurable spaces associated with uncountable Polish spaces are all isomorphic.

The integral and metric space notions come together in functional analysis. The concept of Banach space is central. A Banach space is a vector space with a norm that makes it a complete metric space. The fact that the function space

$L^p$ is a Banach space is a landmark result that sheds much light on such subjects as Fourier analysis and probability.

For linear mapping of Banach spaces the concepts of Lipschitz, uniformly continuous, and continuous coincide. The space of Lipshitz linear mappings from one Banach space to another is again a Banach space. In particular, the dual space of a Banach space is a Banach space. There are two notions of convergence in a dual Banach space, the metric notion of norm convergence and the topological notion of weak$^*$ convergence. Some spaces of measures are dual Banach spaces, and this leads to the useful concept of weak$^*$ convergence of measures.

This book is an introduction to real analysis structures. The goal is to produce a coherent account in a manageable scope. Standard references on real analysis should be consulted for more advanced topics. Folland [5] is an excellent general work. It has the results on locally compact Hausdorff spaces in full generality, and it gives thorough coverage both of theoretical topics and applications material. Another useful reference is Dudley [4]. It gives precise statements of the main results of real analysis. It also defends the thesis that Polish spaces are a natural setting for measure and integration, especially for applications to probability.

The reader should be warned that in these lecture positive means $\geq 0$ and strictly positive means $> 0$. Similar warnings apply to the terms increasing and strictly increasing and to contraction (Lipschitz constant $\leq 1$) and strict contraction (Lipschitz constant $< 1$). This terminology is suggested by the practice of the eminent mathematician Nicolas Bourbaki; it avoids awkward negations. A few unusual definitions are introduced in the text; these are indicated in the index with a dagger †. In discussions of measurable spaces the term $\sigma$-algebra can refer either to the collection of measurable subsets or the corresponding space of measurable functions. In the same way, the term measure can refer to the measure defined on the measurable subsets or to the corresponding integral defined on the positive measurable functions.

The plan of this book is straightforward. Parts I through IV are foundation material. Part I is on Sets and Functions. Part II presents Order and Structure. Part III is on Measure and Integral. Part IV covers Metric Spaces. Parts III and IV may be read in either order.

Part V on Polish Spaces may be read at any point after Parts III and IV. Part VI on Function Spaces includes applications to Fourier analysis and to probability. Part VII on Topology and Measure covers ideas of general topology and their application to Banach spaces, in particular to weak$^*$ convergence of measures.

# Part I

# Sets and Functions

# Chapter 1

# Logical language and mathematical proof

## 1.1 Terms, predicates and atomic formulas

There are many useful ways to present mathematics; sometimes a picture or a physical analogy produces more understanding than a complicated equation. However, the language of mathematical logic has a unique advantage: it gives a standard form for presenting mathematical truth. If there is doubt about whether a mathematical formulation is clear or precise, this doubt can be resolved by converting to this format. The value of a mathematical discovery is enhanced if it is clear that the result and its proof could be stated in such a rigorous framework.

Here is a somewhat simplified model of the language of mathematical logic. There may be *function symbols*. These may be 0-place function symbols, or constants. These stand for objects in some set. Example: 8. Or they may be 1-place functions symbols. These express functions from some set to itself, that is, with one input and one output. Example: square. Or they may be 2-place function symbols. These express functions with two inputs and one output. Example: +.

Once the function symbols have been specified, then one can form *terms*. The language also has a collection of *variables* $x, y, z, x', y', z', \ldots$. Each variable is a term. Each constant $c$ is a term. If $t$ is a term, and $f$ is a 1-place function symbol, then $f(t)$ is a term. If $s$ and $t$ are terms, and $g$ is a 2-place function symbol, then $g(s, t)$ or $(sgt)$ is a term. Example: In an language with constant terms $1, 2, 3$ and 2-place function symbol $+$ the expression $(x + 2)$ is a term, and the expression $(3+(x+2))$ is a term. Note: Sometimes it is a convenient abbreviation to omit outer parentheses. Thus $3 + (x + 2)$ would be an abbreviation for $(3 + (x + 2))$.

The second ingredient is *predicate symbols*. These may be 0-place predicate symbols, or propositional symbols. They may stand for complete sentences. One useful symbol of this nature is $\perp$, which is interpreted as always false. Or they may be 1-place predicate symbols. These express properties. Example: even. Or they may be 2-place predicate symbols. These express relations. Example: $<$.

Once the terms have been specified, then the *atomic formulas* are specified. A propositional symbol is an atomic formula. If $p$ is a property symbol, and $t$ is a term, then $tp$ is an atomic formula. If $s$ and $t$ are terms, and $r$ is a relation symbol, then $srt$ is an atomic formula. Thus $(x + 2) < 3$ is an atomic formula. Note: This could be abbreviated $x + 2 < 3$.

## 1.2 Formulas

Finally there are logical symbols. Each of $\wedge$, $\vee$, $\Rightarrow$ is called a logical *connective*. The $\forall$, $\exists$ are each a *quantifier*. Also, there are parentheses. Once the atomic formulas are specified, then the other *formulas* are obtained by logical operations. For instance $\exists x\, x + 2 < 3$ is an existential formula.

**And** If $A$ and $B$ are formulas, then so is $(A \wedge B)$.

**Or** If $A$ and $B$ are formulas, then so is $(A \vee B)$.

**Implies** If $A$ and $B$ are formulas, the so is $(A \Rightarrow B)$.

The implication $A \Rightarrow B$ is also written

if $A$, then $B$

$A$ only if $B$

$B$ if $A$.

**Not** We shall often abbreviate $(A \Rightarrow \perp)$ by $\neg A$. Thus facts about negation will be special cases of facts about implication.

**Equivalence** Another useful abbreviation is $(A \Leftrightarrow B)$ for $((A \Rightarrow B) \wedge (B \Rightarrow A))$.

The equivalence $A \Leftrightarrow B$ is also written

$A$ if and only if $B$.

**All** If $x$ is a variable and $A(x)$ is a formula, then so is $\forall x\, A(x)$.

The universal quantified formula $\forall x\, A(x)$ is also written

for all $x$ $A(x)$

for each $x$ $A(x)$

for every $x$ $A(x)$.

**Exists** If $x$ is a variable and $A(x)$ is a formula, then so is $\exists x\, A(x)$.

The existential quantified formula $\exists x\, A(x)$ is also written

there exists $x$ with $A(x)$

for some $x$ $A(x)$.

In writing a formula, we often omit the outermost parentheses. However this is just an abbreviation.

The *converse* of $A \Rightarrow B$ is $B \Rightarrow A$. The *contrapositive* of $A \Rightarrow B$ is $\neg B \Rightarrow \neg A$. Note the logical equivalence

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A). \tag{1.1}$$

When $A$ is defined by $B$, the definition is usually written in the form $A$ if $B$. It has the logical force of $A \Leftrightarrow B$.

Note: Avoid expressions of the form "for any $x$ $A(x)$." The word "any" does not function as a quantifier in the usual way. For example, if one says "$z$ is special if and only if for any singular $x$ it is the case that $x$ is tied to $z$", it is not clear which quantifier on $x$ might be intended.

## 1.3 Restricted variables

Often a quantifier has a restriction. The restricted universal quantifier is $\forall x\,(C(x) \Rightarrow A(x))$. The restricted existential quantifier is $\exists x\,(C(x) \wedge A(x))$. Here $C(x)$ is a formula that places a restriction on the $x$ for which the assertion is made.

It is common to have implicit restrictions. For example, say that the context of a discussion is real numbers $x$. There may be an implicit restriction $x \in \mathbb{R}$. Since the entire discussion is about real numbers, it may not be necessary to make this explicit in each formula. This, instead of $\forall x\,(x \in \mathbb{R} \Rightarrow x^2 \geq 0)$ one would write just $\forall x\, x^2 \geq 0$.

Sometimes restrictions are indicated by use of special letters for the variables. Thus often $i, j, k, l, m, n$ are used for integers. Instead of saying that $m$ is odd if and only if $\exists y\,(y \in \mathbf{N} \wedge m = 2y + 1)$ one would just write that $m$ is odd if and only if $\exists k\, m = 2k + 1$.

The letters $\epsilon, \delta$ are used for strictly positive real numbers. The corresponding restrictions are $\epsilon > 0$ and $\delta > 0$. Thus instead of writing $\forall x\,(x > 0 \Rightarrow \exists y\,(y > 0 \wedge y < x))$ one would write $\forall \epsilon \exists \delta\, \delta < \epsilon$;.

Other common restrictions are to use $f, g, h$ for functions or to indicate sets by capital letters. Reasoning with restricted variables should work smoothly, provided that one keeps the restriction in mind at the appropriate stages of the argument.

## 1.4 Free and bound variables

In a formula each occurrence of a variable is either free or bound. The occurrence of a variable $x$ is *bound* if it is in a subformula of the form $\forall x\, B(x)$ or $\exists x\, B(x)$. (There may also be other operations, such as the set builder operation, that produce bound variables.) If the occurrence is not bound, then it is said to be *free*.

In general, a bound variable may be replaced by a new bound variable without changing the meaning of the formula. Thus, for instance, if $y'$ is a variable that does not occur in the formula, one could replace the occurrences of $y$ in the subformula $\forall y\, B(y)$ by $y'$, so the new subformula would now be $\forall y'\, B(y')$. Of course if the variables are restricted, then the change of variable should respect the restriction.

Example: Let the formula be $\exists y\, x < y$. This says that there is a number greater than $x$. In this formula $x$ is free and $y$ is bound. The formula $\exists y'\, x < y'$ has the same meaning. In this formula $x$ is free and $y'$ is bound. On the other hand, the formula $\exists y\, x' < y$ has a different meaning. This formula says that there is a number greater than $x'$.

We wish to define *careful substitution* of a term $t$ for the free occurrences of a variable $x$ in $A(x)$. The resulting formula will be denoted $A(t)$ There is no particular problem in defining substitution in the case when the term $t$ has no variables that already occur in $A(x)$. The care is needed when there is a subformula in which $y$ is a bound variable and when the term $t$ contains the variable $y$. Then mere substitution might produce an unwanted situation in which the $y$ in the term $t$ becomes a bound variable. So one first makes a change of bound variable in the subformula. Now the subformula contains a bound variable $y'$ that cannot be confused with $y$. Then one substitutes $t$ for the free occurrences of $x$ in the modified formula. Then $y$ will be a free variable after the substitution, as desired.

Example: Let the formula be $\exists y\, x < y$. Say that one wished to substitute $y+1$ for the free occurrences of $x$. This should say that there is a number greater than $y+1$. It would be wrong to make the careless substitution $\exists y\, y+1 < y$. This statement is not only false, but worse, it does not have the intended meaning. The careful substitution proceeds by first changing the original formula to $\exists y'\, x < y'$. The careful substitution then produces $\exists y'\, y+1 < y'$. This says that there is a number greater than $y+1$, as desired.

The general rule is that if $y$ is a variable with bound occurrences in the formula, and one wants to substitute a term $t$ containing $y$ for the free occurrences of $x$ in the formula, then one should change the bound occurrences of $y$ to bound occurrences of a new variable $y'$ before the substitution. This gives the kind of careful substitution that preserves the intended meaning.

## 1.5 Quantifier logic

Here are some useful logical equivalences. The law of double negation states that

$$\neg\neg A \Leftrightarrow A. \tag{1.2}$$

**??***De Morgan's laws for connectives* state that

$$\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B) \tag{1.3}$$

and that

$$\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B). \tag{1.4}$$

**??***De Morgan's laws for quantifiers* state that

$$\neg\forall x\, A(x) \Leftrightarrow \exists x\, \neg A(x) \tag{1.5}$$

and

$$\neg\exists x\, A(x) \Leftrightarrow \forall x\, \neg A(x). \tag{1.6}$$

Since $\neg(A \Rightarrow B) \Leftrightarrow (A \wedge \neg B)$ and $\neg(A \wedge B) \Leftrightarrow (A \Rightarrow \neg B)$, De Morgan's laws continue to work with restricted quantifiers.

Examples:

1. The function $f$ is continuous if $\forall a\forall\epsilon\exists\delta\forall x(|x-a| < \delta \Rightarrow |f(x)-f(a)| < \epsilon)$. It is assumed that $a, x, \epsilon, \delta$ are real numbers with $\epsilon > 0, \delta > 0$.

2. The function $f$ is not continuous if $\exists a\exists\epsilon\forall\delta\exists x(|x-a| < \delta \wedge \neg|f(x)-f(a)| < \epsilon)$. This is a mechanical application of De Morgan's laws.

Similarly, the function $f$ is uniformly continuous if $\forall\epsilon\exists\delta\forall a\forall x(|x-a| < \delta \Rightarrow |f(x)-f(a)| < \epsilon)$. Notice that the only difference is the order of the quantifiers.

Examples:

1. Consider the proof that $f(x) = x^2$ is continuous. The heart of the proof is to prove the existence of $\delta$. The key computation is $|x^2 - a^2| = |x + a||x - a| = |x - a + 2a||x - a|$. If $|x - a| < 1$ then this is bounded by $(2|a| + 1)|x - a|$.

   Here is the proof. Let $\epsilon > 0$. Suppose $|x - a| < \min(1, \epsilon/(2|a| + 1))$. From the above computation it is easy to see that $|x^2 - a^2| < \epsilon$. Hence $|x-a| < \min(1, \epsilon/(2|a|+1)) \Rightarrow |x^2 - a^2| < \epsilon$. Since in this last statement $x$ is arbitrary, $\forall x\, (|x - a| < \min(1, \epsilon/(2|a| + 1)) \Rightarrow |x^2 - a^2| < \epsilon)$. Hence $\exists\delta\forall x\, (|x - a| < \delta \Rightarrow |x^2 - a^2| < \epsilon)$. Since $\epsilon > 0$ and $a$ are arbitrary, the final result is that $\forall a\forall\epsilon\exists\delta\forall x\, (|x - a| < \delta \Rightarrow |x^2 - a^2| < \epsilon)$.

2. Consider the proof that $f(x) = x^2$ is not uniformly continuous. Now the idea is to take $x-a = \delta/2$ and use $x^2-a^2 = (x+a)(x-a) = (2a+\delta/2)(\delta/2)$.

   Here is the proof. With the choice of $x-a = \delta/2$ and with $a = 1/\delta$ we have that $|x-a| < \delta$ and $|x^2-a^2| \geq 1$. Hence $\exists a\exists x\, (|x-a| < \delta \wedge |x^2-a^2| \geq 1)$. Since $\delta > 0$ is arbitrary, it follows that $\forall\delta\exists a\exists x\, (|x-a| < \delta \wedge |x^2-a^2| \geq 1)$. Finally we conclude that $\exists\epsilon\forall\delta\exists a\exists x\, (|x - a| < \delta \wedge |x^2 - a^2| \geq \epsilon)$.

It is a general fact that $f$ uniformly continuous implies $f$ continuous. This is pure logic; the only problem is to interchange the $\exists\delta$ quantifier with the $\forall a$ quantifier. This can be done in one direction. Suppose that $\exists\delta\forall a\, A(\delta, a)$. Let $\delta'$ be a temporary name for the number that exists, so that $\forall a A(\delta', a)$. In particular, $A(\delta', a')$. It follows that $\exists\delta A(\delta, a')$. This conclusion does not depend on the name, so it follows from the original supposition. Since $a'$ is arbitrary, it follows that $\forall a\exists\delta\, A(\delta, a)$.

What goes wrong with the converse argument? Suppose that $\forall a \exists \delta\, A(\delta, a)$. Then $\exists \delta\, A(\delta, a')$. Let $a'$ satisfy $A(\delta', a')$. The trouble is that $a'$ is not arbitrary, because something special has been supposed about it. So the generalization is not permitted.

## 1.6 Natural deduction

The formalization of logic that corresponds most closely to the practice of mathematical proof is *natural deduction*. Natural deduction proofs are constructed so that they may be read from the top down. (On the other hand, to construct a natural deduction proof, it is often helpful to work from the top down and the bottom up and try to meet in the middle.)

In natural deduction each **Suppose** introduces a new hypothesis to the set of hypotheses. Each matching **Thus** removes the hypothesis. Each line is a claim that the formula on this line follows logically from the hypotheses above that have been introduced by a **Suppose** and not yet eliminated by a matching **Thus** .

Example: Say that one wants to show that if one knows the algebraic fact $\forall x\,(x > 0 \Rightarrow (x+1) > 0)$, then one is forced by pure logic to accept that $\forall y\,(y > 0 \Rightarrow ((y+1)+1) > 0)$. Here is the argument, showing every logical step.

**Suppose** $\forall x(x > 0 \Rightarrow (x+1) > 0)$
  **Suppose** $z > 0$
  $z > 0 \Rightarrow (z+1) > 0$
  $(z+1) > 0$
  $(z+1) > 0 \Rightarrow ((z+1)+1) > 0$
  $((z+1)+1) > 0$
**Thus** $z > 0 \Rightarrow ((z+1)+1) > 0$
$\forall y\,(y > 0 \Rightarrow ((y+1)+1) > 0$

Notice that the indentation makes the hypotheses in force at each stage quite clear. On the other hand, the proof could also be written in narrative form. It could go like this.

Example: **Suppose** that for all $x$, if $x > 0$ then $(x+1) > 0$. **Suppose** $z > 0$. By specializing the hypothesis, obtain that if $z > 0$, then $(z+1) > 0$. It follows that $(z+1) > 0$. By specializing the hypothesis again, obtain that if $(z+1) > 0$, then $((z+1)+1) > 0$. It follows that $((z+1)+1) > 0$. **Thus** if $z > 0$, then $((z+1)+1) > 0$. Since $z$ is arbitrary, conclude that for all $y$, if $(y > 0$, then $((y+1)+1) > 0)$.

Mathematicians usually write in narrative form, but it is useful to practice proofs in outline form, with proper indentation to show the subarguments.

Natural deduction takes time to learn, and so a full exposition is not attempted here. However it worth being aware that there are systematic rules for logical deduction. The following pages present the rules for natural deduction, at least for certain of the logical operations. In each rule there is a connective or quantifier that is the center of attention. It may be in the hypothesis or in the conclusion. The rule shows how to reduce an argument involving this logical operation to one without the logical operation. (To accomplish this, the rule needs to be used just once, except for the all in hypothesis and exists in conclusion rules. If it were not for this exception, mathematics would be simple indeed.)

## 1.7 Rules for logical operations

Here is a complete set of natural deduction rules for the logical operations $\wedge$, $\forall$, $\Rightarrow$, $\neg$, and the falsity symbol $\perp$. Most of these these rules gives a practical method for using a hypothesis or for proving a conclusion that works in all circumstances. The exceptions are noted, but the supplement provides recipes for these cases too.

**And in hypothesis**

$A \wedge B$

$A$

$B$

**And in conclusion**

$A$

$B$

$A \wedge B$

**All in hypothesis**

$\forall x\, A(x)$

$A(t)$

Note: This rule may be used repeatedly with various terms.

**All in conclusion** If $z$ is a variable that does not occur free in a hypothesis in force or in $\forall x\, A$, then

$A(z)$

$\forall x\, A(x)$

Note: The restriction on the variable is usually signalled by an expression such as “since $z$ is arbitrary, conclude $\forall x\, A(x)$.”

**Implication in hypothesis**

$A \Rightarrow B$

$A$

$B$

Note: This rule by itself is an incomplete guide to practice, since it may not be clear how to prove $A$. See the supplement for a variant that always works.

**Implication in conclusion**

**Suppose** $A$

$B$

**Thus** $A \Rightarrow B$

The operation of negation $\neg A$ is regarded as an abbreviation for $A \Rightarrow \bot$. Thus we have the following specializations of the implication rules.

**Not in hypothesis**

$\neg A$

$A$

$\bot$

Note: This rule by itself is an incomplete guide to practice, since it may not be clear how to prove $A$. See the supplement for a variant that always works.

**Not in conclusion**

**Suppose** $A$

$\bot$

**Thus** $\neg A$

Finally, there is the famous law of contradiction.

**Contradiction**

**Suppose** $\neg C$

$\bot$

**Thus** $C$

So far there are no rules for $A \vee B$ and for $\exists x\, A(x)$. In principle one could not bother with such rules, because $A \vee B$ could always be replaced by $\neg(\neg A \wedge \neg B)$ and $\exists x\, A(x)$ could be replaced by $\neg\forall x\, \neg A(x)$. Such a replacement is rather clumsy and is not done in practice. Thus the following section gives additional rules that explicitly deal with $A \vee B$ and with $\exists x\, A(x)$.

## 1.8 Additional rules for or and exists

**Or in hypothesis**

$A \vee B$

**Suppose** A

C

**Instead suppose** $B$

$C$

**Thus** $C$

**Or in conclusion**

$A$

$A \vee B$

together with

$B$

$A \vee B$

Note: This rule is an incomplete guide to practice. See the supplement for a variant that always works.

**Exists in hypothesis** If $z$ is a variable that does not occur free in a hypothesis in force, in $\exists x\, A$, or in $C$, then

$\exists x\, A(x)$

**Let** $A(z)$

$C$

From this point on treat $C$ as a consequence of the existential hypothesis without the temporary supposition $A(z)$ or its temporary consequences.

The safe course is to take $z$ to be a variable that is used as a temporary name in this context, but which occurs nowhere else in the argument.

**Exists in conclusion**

$A(t)$

$\exists x\, A(x)$

Note: This rule is an incomplete guide to practice. See the supplement for a variant that always works.

## 1.9 Strategies for natural deduction

A natural deduction proof is read from top down. However it is often discovered by working simultaneously from the top and the bottom, until a meeting in the middle. The discoverer then obscures the origin of the proof by presenting it from the top down. This is convincing but not illuminating.

Example: Here is a natural deduction proof that $\forall x\,(x \text{ rich} \Rightarrow x \text{ happy})$ leads to $\forall x\,(\neg x \text{ happy} \Rightarrow \neg x \text{ rich})$.

**Suppose** $\forall x\,(x \text{ rich} \Rightarrow x \text{ happy})$
  **Suppose** $\neg w$ rich
    **Suppose** $w$ happy
    $w$ happy $\Rightarrow w$ rich
    $w$ rich
    $\perp$
  **Thus** $\neg w$ happy
**Thus** $\neg w$ rich $\Rightarrow \neg w$ happy
$\forall x\,(\neg x \text{ happy} \Rightarrow \neg x \text{ rich})$

There are 3 “Suppose” lines and 2 “Thus” lines. Each “Thus” removes a “Suppose.” Since 3-2 = 1, the bottom line follows from the top line alone.

Here is how to construct the proof. Start from the bottom up. To prove the general conclusion, prove the implication for an arbitrary variable. To prove the implication, make a supposition. This reduces the problem to proving a negation. Then work from outside to inside. Make a supposition without the

negation and try to get a contradiction. To accomplish this, specialize the hypothesis.
Example: Here is the same proof in narrative form.

**Suppose** $\forall x\,(x$ rich $\Rightarrow x$ happy$)$. **Suppose** $\neg w$ rich. **Suppose** $w$ happy. Specializing the hypothesis gives $w$ happy $\Rightarrow w$ rich. So $w$ rich. This gives a false conclusion $\perp$. **Thus** $\neg w$ happy. **Thus** $\neg w$ rich $\Rightarrow \neg w$ happy. Since $w$ is arbitrary $\forall x\,(\not x$ happy $\Rightarrow \neg x$ rich$)$

Example: Here is a natural deduction proof of the fact that $\exists x\,(x$ happy$\wedge x$ rich$)$ logically implies that $\exists x\, x$ happy $\wedge \exists x\, x$ rich.

**Suppose** $\exists x\,(x$ happy $\wedge\, x$ rich$)$
**Let** $z$ happy $\wedge\, z$ rich
$z$ happy
$z$ rich
$\exists x\, x$ happy
$\exists x\, x$ rich
$\exists x\, x$ happy $\wedge \exists x\, x$ rich

Example: Here is the same proof in narrative form.

**Suppose** $\exists x\,(x$ happy $\wedge\, x$ rich$)$. **Let** $z$ happy $\wedge\, z$ rich. Then $z$ happy and hence $\exists x\, x$ happy. Similarly, $z$ rich and hence $\exists x\, x$ rich. It follows that $\exists x\, x$ happy $\wedge \exists x\, x$ rich. Since $z$ is an arbitrary name, this conclusion holds on the basis of the original supposition of existence.

Example: We could try to reason in the other direction, from the existence of a happy individual and the existence of a rich individual to the existence of a happy, rich individual? What goes wrong?

**Suppose** $\exists x\, x$ happy$\wedge\exists x\, x$ rich. Then $\exists x\, x$ happy, $\exists x\, x$ rich. **Let** $z$ happy. **Let** $w$ rich. Then $z$ happy $\wedge\, w$ rich. This approach does not work.

Example: Here is another attempt at the other direction.

**Suppose** $\exists x\, x$ happy$\wedge\exists x\, x$ rich. Then $\exists x\, x$ happy, $\exists x\, x$ rich. **Let** $z$ happy. **Let** $z$ rich. Then $z$ happy $\wedge\, z$ rich. So $\exists x\,(x$ happy $\wedge\, x$ rich$)$. So this proves the conclusion, but we needed two temporary hypotheses on $z$. However we cannot conclude that we no longer need the last temporary hypothesis $z$ rich, but only need $\exists x\, x$ rich. The problem is that we have temporarily supposed also that $z$ happy, and so $z$ is not an arbitrary name for the rich individual. All this proves is that one can deduce logically from $z$ happy, $z$ rich that $\exists x\,(x$ happy $\wedge\, x$ rich$)$. So this approach also does not work.

Example: Here is a natural deduction proof that $\exists y \forall x\, x \leq y$ gives $\forall x \exists y\, x \leq y$.

**Suppose** $\exists y \forall x\, x \leq y$
**Let** $\forall x\, x \leq y'$
$x' \leq y'$
$\exists y\, x' \leq y$
$\forall x \exists y\, x \leq y$

Example: Here is the same proof in narrative form.

**Suppose** $\exists y \forall x\, x \leq y$. **Let** $y'$ satisfy $\forall x\, x \leq y'$. In particular, $x' \leq y'$. Therefore $\exists y\, x' \leq y$. In fact, since $y'$ is just an arbitrary name, this follows on the basis of the original existential supposition. Finally, since $x'$ is arbitrary, conclude that $\forall x \exists y\, x \leq y$.

A useful strategy for natural deduction is to begin with writing the hypotheses at the top and the conclusion at the bottom. Then work toward the middle. The most important point is to try to use the all in conclusion rule and the exists in hypothesis rule early in this process of proof construction. This introduces new "arbitrary" variables. Then one uses the all in hypothesis rule and the exists in conclusion rule with terms formed from these variables. So it is reasonable to use these latter rules later in the proof construction process. They may need to be used repeatedly.

## 1.10 Lemmas and theorems

In statements of mathematical theorems it is common to have implicit universal quantifiers. For example say that we are dealing with real numbers. Instead of stating the theorem that

$$\forall x \forall y\, 2xy \leq x^2 + y^2 \tag{1.7}$$

one simply claims that

$$2uv \leq u^2 + v^2. \tag{1.8}$$

Clearly the second statement is a specialization of the first statement. But it seems to talk about $u$ and $v$, and it is not clear why this might apply for someone who wants to conclude something about $p$ and $q$, such as $2pq \leq p^2 + q^2$. Why is this permissible?

The answer is that the two displayed statements are logically equivalent, provided that there is no hypothesis in force that mentions the variables $u$ or $v$. Then given the second statement and the fact that the variables in it are arbitrary, the first statement is a valid generalization.

Notice that there is no similar principle for existential quantifiers. The statement

$$\exists x\, x^2 = x \tag{1.9}$$

is a theorem about real numbers, while the statement

$$u^2 = u \tag{1.10}$$

is a condition that is true for $u = 0$ or $u = 1$ and false for all other real numbers. It is certainly not a theorem about real numbers. It might occur in a context where there is a hypothesis that $u = 0$ or $u = 1$ in force, but then it would be incorrect to generalize.

One cannot be careless about inner quantifiers, even if they are universal. Thus there is a theorem

$$\exists x\, x < y. \tag{1.11}$$

This could be interpreted as saying that for each arbitrary $y$ there is a number that is smaller than $y$. Contrast this with the statement

$$\exists x \forall y\, x < y \tag{1.12}$$

with an inner universal quantifier. This is clearly false for the real number system.

## 1.11 Relaxed natural deduction

Mathematicians ordinarily do not care to put in all logical steps explicitly, as would be required by the natural deduction rules. However there is a more relaxed version of natural deduction which might be realistic in some contexts. This version omits certain trivial logical steps. Here is an outline of how it goes.

**And** The rules for eliminating "and" from a hypothesis and for introducing "and" in the conclusion are regarded as obvious.

**All** The rule for eliminating $\forall x\, A(x)$ from a hypothesis by replacing it with $A(t)$ is regarded as obvious. The rule for introducing $\forall x\, A(x)$ in a conclusion is indicated more explicitly, by some such phrase as "since $x$ is arbitrary", which means that at this stage $x$ does not occur as a free variable in any hypothesis in force.

**Implies** The rule for eliminating $\Rightarrow$ from a hypothesis is regarded as obvious. The rule for introducing $\Rightarrow$ in a conclusion requires special comment. At an earlier stage there was a **Suppose** $A$. After some logical reasoning there is a conclusion $B$. Then the removal of the supposition and the introduction of the implication is indicated by **Thus** $A \Rightarrow B$.

**Not** The rule for eliminating $\neg$ from a hypothesis is regarded as obvious. The rule for introducing $\neg$ in a conclusion requires special comment. At an earlier stage there was a **Suppose** $A$. After some logical reasoning there is a false conclusion $\perp$. Then the removal of the supposition and the introduction of the negation is indicated by **Thus** $\neg A$.

**Contradiction** The rule for proof by contradiction requires special comment. At an earlier stage there was a **Suppose** $\neg A$. After some logical reasoning there is a false conclusion $\perp$. Then the removal of the supposition and the introduction of the conclusion is indicated by **Thus** $A$.

**Or** The rule for eliminating $\vee$ for the hypothesis is by proof by cases. Start with $A \vee B$. **Suppose** $A$ and reason to conclusion $C$. **Instead suppose** $B$ and reason to the same conclusion $C$. **Thus** $C$. The rule for starting with $A$ (or with $B$) and introducing $A \vee B$ in the conclusion is regarded as obvious.

**Exists** Mathematicians tend to be somewhat casual about $\exists x\, A(x)$ in a hypothesis. The technique is to **Let** $A(z)$. Thus $z$ is a variable that may be used as a temporary name for the object that has been supposed to exist. (The safe course is to take a variable that will be used only in this context.) Then the reasoning leads to a conclusion $C$ that does not mention $z$. The conclusion actually holds as a consequence of the existential hypothesis, since it did not depend on the assumption about $z$. The rule for starting with $A(t)$ and introducing $\exists x\, A(x)$ is regarded as obvious.

Rules for equality (everything is equal to itself, equals may be substituted) are also used without comment.

One of the most important concepts of analysis is the concept of open set. This makes sense in the context of the real line, or in the more general case of Euclidean space, or in the even more general setting of a metric space. Here we use notation appropriate to the real line, but little change is required to deal with the other cases.

For all subsets $V$, we say that $V$ is open if $\forall a\,(a \in V \Rightarrow \exists\epsilon\forall x\,(|x-a| < \epsilon \Rightarrow x \in V))$.

Recall the definition of union of a collection $\Gamma$ of subsets. This says that for all $y$ we have $y \in \bigcup\Gamma$ if and only if $\exists W\,(W \in \Gamma \wedge y \in W)$.

Here is a proof of the theorem that for all collections of subsets $\Gamma$ the hypothesis $\forall U\,(U \in \Gamma \Rightarrow U$ open) implies the conclusion $\bigcup\Gamma$ open. The style of the proof is a relaxed form of natural deduction in which some trivial steps are skipped.

**Suppose** $\forall U\,(U \in \Gamma \Rightarrow U$ open). **Suppose** $a \in \bigcup\Gamma$. By definition $\exists W(W \in \Gamma \wedge a \in W)$. **Let** $W' \in \Gamma \wedge a \in W'$. Since $W' \in \Gamma$ and $W' \in \Gamma \Rightarrow W'$ open, it follows that $W'$ open. Since $a \in W'$ it follows from the definition that $\exists\epsilon\forall x\,(|x-a| < \epsilon \Rightarrow x \in W')$. **Let** $\forall x\,(|x-a| < \epsilon' \Rightarrow x \in W')$. **Suppose** $|x-a| < \epsilon'$. Then $x \in W'$. Since $W' \in \Gamma \wedge x \in W'$, it follows that $\exists W(W \in \Gamma \wedge x \in W')$. Then from the definition $x \in \bigcup\Gamma$. **Thus** $|x-a| < \epsilon' \Rightarrow x \in \bigcup\Gamma$. Since $x$ is arbitrary, $\forall x\,(|x-a| < \epsilon' \Rightarrow x \in \bigcup\Gamma)$. So $\exists\epsilon\forall x\,(|x-a| < \epsilon \Rightarrow x \in \bigcup\Gamma)$. **Thus** $a \in \bigcup\Gamma \Rightarrow \exists\epsilon\forall x\,(|x-a| < \epsilon \Rightarrow x \in \bigcup\Gamma)$. Since $a$ is arbitrary, $\forall a\,(a \in \bigcup\Gamma \Rightarrow \exists\epsilon\forall x\,(|x-a| < \epsilon \Rightarrow x \in \bigcup\Gamma))$. So by definition $\bigcup\Gamma$ open. **Thus** $\forall U\,(U \in \Gamma \Rightarrow U$ open) $\Rightarrow \bigcup\Gamma$ open.

In practice, the natural deduction rules are useful only for the construction of small proofs and for verification of a proof after the fact. The way to make progress in mathematics is find concepts that have meaningful interpretations. In order to prove a major theorem, one prepares by proving smaller theorems or lemmas. Each of these may have a rather elementary proof. But the choice of the statements of the lemmas is crucial in making progress. So while the micro structure of mathematical argument is based on the rules of proof, the global structure is a network of lemmas, theorems, and theories based on astute selection of mathematical concepts.

To illustrate this, here is a final version of the proof that the union of a collection $\Gamma$ of open subsets is open. The simplification is due to the notion of open ball $B(a,\epsilon)$, $\epsilon > 0$, and the use of set notation and facts about sets. It allows us to say that $V$ is open if $\forall a\,(a \in V \Rightarrow \exists\epsilon B(a,\epsilon) \subset V)$.

**Suppose** $\forall U\,(U \in \Gamma \Rightarrow U$ open). **Suppose** $a \in \bigcup\Gamma$. By definition $\exists W(W \in \Gamma \wedge a \in W)$. **Let** $W' \in \Gamma \wedge a \in W'$. Since $W' \in \Gamma$ and $W' \in \Gamma \Rightarrow W'$ open, it follows that $W'$ open. Since $a \in W'$ it follows from the definition that $\exists\epsilon B(a,\epsilon) \subset W'$. **Let** $B(a,\epsilon') \subset W'$. Use the fact that $W' \in \Gamma$ implies $W' \subset \bigcup\Gamma$. It follows that $W' \subset \bigcup\Gamma$. Conclude that $B(a,\epsilon') \subset \bigcup\Gamma$. So $\exists\epsilon B(a,\epsilon) \subset \bigcup\Gamma$. **Thus** $a \in \bigcup\Gamma \Rightarrow \exists\epsilon B(a,\epsilon) \subset \bigcup\Gamma$. Since $a$ is arbitrary, $\forall a\,(a \in \bigcup\Gamma \Rightarrow B(a,\epsilon) \subset \bigcup\Gamma)$. So by definition $\bigcup\Gamma$ open. **Thus** $\forall U\,(U \in \Gamma \Rightarrow U$ open) $\Rightarrow \bigcup\Gamma$ open.

## 1.12 Supplement: Templates

Here are templates that show how to use the contradiction rule to remove the imperfections of certain of the natural deduction rules presented above. These templates are sometimes less convenient, but they always work to produce a proof, if one exists.

**Implication in hypothesis template**

$A \Rightarrow B$

**Suppose** $\neg C$

$A$

$B$

$\perp$

**Thus** $C$

**Negation in hypothesis template**

$\neg A$

**Suppose** $\neg C$

$A$

$\perp$

**Thus** $C$

Note: The role of this rule to make use of a negated hypothesis $\neg A$. When the conclusion $C$ has no useful logical structure, but $A$ does, then the rule effectively switches $A$ for $C$.

**Or in conclusion template** Replace $A \vee B$ in a conclusion by $\neg(\neg A \wedge \neg B)$. Thus the template is

$\neg(\neg A \wedge \neg B)$

$A \vee B$.

**Exists in conclusion template** Replace $\exists x\, A(x)$ in a conclusion by $\neg(\forall x\, \neg A(x))$. Thus the template is

$\neg(\forall x\, \neg A(x))$

$\exists x\, A(x)$.

The Gödel completeness theorem says given a set of hypotheses and a conclusion, then either there is a proof using the natural deduction rules, or there is an interpretation in which the hypotheses are all true and the conclusion is false. Furthermore, in the case when there is a proof, it may be constructed via these templates.

Whey does this theorem not make mathematics trivial? The problem is that if there is no proof, then the unsuccessful search for one may not terminate. The problem is with the rule for "all" in the hypothesis. This may be specialized in more than one way, and there is not upper bound to the number of unsuccessful attempts.

## 1.13 Supplement: Existential hypotheses

Most expositions of natural deduction give a different version of the rule for an existential hypothesis. This rule displays the logical pattern much more clearly. Unfortunately, it is not part of everyday mathematical practice.

For the record, here is the rule:

If $z$ is a variable that does not occur free in a hypothesis in force, in $\exists x\, A$, or in $C$, then

$\exists x\, A(x)$
        **Suppose** $A(z)$

        $C$
    **Thus** $C$         exists in hypothesis

Note: The restriction on the variable could be signalled by an expression such as "since $z$ is arbitrary, conclude $C$ on the basis of the existential hypothesis $\exists x\, A(x)$."

As we have seen, mathematicians tend not to use this version of the rule. They simply suppose that some convenient variable may be used as a name for the thing that exists. They reason with this name up to a point at which they get a conclusion that no longer mentions it. At this point they conveniently forget the temporary supposition.

Example: Here is a natural deduction proof of the fact that $\exists x\, (x \text{ happy} \wedge x \text{ rich})$ logically implies that $\exists x\, x \text{ happy} \wedge \exists x\, x \text{ rich}$.

    **Suppose** $\exists x\, (x \text{ happy} \wedge x \text{ rich})$
        **Suppose** $z \text{ happy} \wedge z \text{ rich}$
        $z$ happy
        $z$ rich
        $\exists x\, x$ happy
        $\exists x\, x$ rich
        $\exists x\, x \text{ happy} \wedge \exists x\, x \text{ rich}$
    **Thus** $\exists x\, x \text{ happy} \wedge \exists x\, x \text{ rich}$

Example: Here is a natural deduction proof that $\exists y \forall x\, x \leq y$ gives $\forall x \exists y\, x \leq y$.

**Suppose** $\exists y \forall x\, x \leq y$

  **Suppose** $\forall x\, x \leq y'$

  $x' \leq y'$

  $\exists y\, x' \leq y$

**Thus** $\exists y\, x' \leq y$

$\forall x \exists y\, x \leq y$

# Problems

1. Quantifiers. A sequence of functions $f_n$ converges pointwise (on some set of real numbers) to $f$ as $n$ tends to infinity if $\forall x \forall \epsilon \exists N \forall n (n \geq N \Rightarrow |f_n(x) - f(x)| < \epsilon)$. Here the restrictions are that $x$ is in the set and $\epsilon > 0$. Show that for $f_n(x) = x^n$ and for suitable $f(x)$ there is pointwise convergence on the closed interval $[0, 1]$.

2. Quantifiers. A sequence of functions $f_n$ converges uniformly (on some set of real numbers) to $f$ as $n$ tends to infinity if $\forall \epsilon \exists N \forall x \forall n (n \geq N \Rightarrow |f_n(x) - f(x)| < \epsilon)$. Show that for $f_n(x) = x^n$ and the same $f(x)$ the convergence is not uniform on $[0, 1]$.

3. Quantifiers. Show that uniform convergence implies pointwise convergence.

4. Quantifiers. Show that if $f_n$ converges uniformly to $f$ and if each $f_n$ is continuous, then $f$ is continuous.

   Hint: The first hypothesis is $\forall \epsilon \exists N \forall x \forall n\, (n \geq N \Rightarrow |f_n(x) - f(x)| < \epsilon)$. Deduce that $\exists N \forall x \forall n\, (n \geq N \Rightarrow |f_n(x) - f(x)| < \epsilon'/3)$. Temporarily suppose $\forall x \forall n\, (n \geq N' \Rightarrow |f_n(x) - f(x)| < \epsilon'/3)$.

   The second hypothesis is $\forall n \forall a \forall \epsilon \exists \delta \forall x\, (|x - a| < \delta \Rightarrow |f_n(x) - f_n(a)| < \epsilon)$. Deduce that $\exists \delta \forall x\, (|x - a| < \delta \Rightarrow |f_{N'}(x) - f_{N'}(a)| < \epsilon'/3)$. Temporarily suppose that $\forall x\, (|x - a| < \delta' \Rightarrow |f_{N'}(x) - f_{N'}(a)| < \epsilon'/3)$.

   Suppose $|x - a| < \delta'$. Use the temporary suppositions above to deduce that $|f(x) - f(a)| < \epsilon'$. Thus $|x - a| < \delta' \Rightarrow |f(x) - f(a)| < \epsilon'$. This is well on the way to the desired conclusion. However be cautious: At this point $x$ is arbitrary, but $a$ is not arbitrary. (Why?) Explain in detail the additional arguments to reach the goal $\forall a \forall \epsilon \exists \delta \forall x (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon)$.

5. Quantifiers. A function $f : \mathbf{R} \to \mathbf{R}$ is uniformly continuous iff $\forall \epsilon \exists \delta \forall x \forall y\, (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$. Here $\epsilon > 0$ and $\delta > 0$ are variables restricted to be strictly positive. Describe the class of functions such that $\forall x \forall y \forall \epsilon \exists \delta\, (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$.

6. Logical deduction. Here is a mathematical argument that shows that there is no largest prime number. Assume that there were a largest prime number. Call it $a$. Then $a$ is prime, and for every number $j$ with $a < j$,

$j$ is not prime. However, for every number $m$, there is a number $k$ that divides $m$ and is prime. Hence there is a number $k$ that divides $a! + 1$ and is prime. Call it $b$. Now every number $k > 1$ that divides $n! + 1$ must satisfy $n < k$. (Otherwise it would have a remainder of 1.) Hence $a < b$. But then $b$ is not prime. This is a contradiction.

Write a complete proof in outline form to show that from pure logic it follows that the hypotheses

$$\forall m \exists k \, (k \text{ prime} \wedge k \text{ divides } m) \tag{1.13}$$

$$\forall n \forall k \, (k \text{ divides } n! + 1 \Rightarrow n < k) \tag{1.14}$$

logically imply the conclusion

$$\neg \exists n \, (n \text{ prime} \wedge \forall j \, (n < j \Rightarrow \neg \, j \text{ prime})). \tag{1.15}$$

7. Logical deduction. It is a well-known mathematical fact that $\sqrt{2}$ is irrational. In fact, if it were rational, so that $\sqrt{2} = m/n$, then we would have $2n^2 = m^2$. Thus $m^2$ would have an even number of factors of 2, while $2n^2$ would have an odd number of factors of two. This would be a contradiction.

   Show that from logic alone it follows that

$$\forall i \, i^2 \text{ even-twos} \tag{1.16}$$

   and

$$\forall j \, (j \text{ even-twos} \Rightarrow \neg (2 \cdot j) \text{ even-twos}) \tag{1.17}$$

   give

$$\neg \exists m \exists n \, (2 \cdot n^2) = m^2. \tag{1.18}$$

8. Logical deduction. If $X$ is a set, then $P(X)$ is the set of all subsets of $X$. If $X$ is finite with $n$ elements, then $P(X)$ is finite with $2^n$ elements. A famous theorem of Cantor states that there is no function $f$ from $X$ to $P(X)$ that is onto $P(X)$. Thus in some sense there are more elements in $P(X)$ than in $X$. This is obvious when $X$ is finite, but the interesting case is when $X$ is infinite.

   Here is an outline of a proof. Consider an arbitrary function $f$ from $X$ to $P(X)$. We want to show that there exists a set $V$ such that for each $x$ in $X$ we have $f(x) \neq V$. Consider the condition that $x \notin f(x)$. This condition defines a set. That is, there exists a set $U$ such that for all $x$, $x \in U$ is equivalent to $x \notin f(x)$. Call this set $S$. Let $p$ be arbitrary. Suppose $f(p) = S$. Suppose $p \in S$. Then $p \notin f(p)$, that is, $p \notin S$. This is a contradiction. Thus $p \notin S$. Then $p \in f(p)$, that is, $p \in S$. This is a contradiction. Thus $f(p) \neq S$. Since this is true for arbitrary $p$, it follows that for each $x$ in $X$ we have $f(x) \neq S$. Thus there is a set that is not in the range of $f$.

Show that from logic alone it follows that from

$$\exists U\,\forall x\,((x \in U \Rightarrow \neg x \in f(x)) \wedge (\neg x \in f(x) \Rightarrow x \in U)) \tag{1.19}$$

one can conclude that

$$\exists V\,\forall x\,\neg f(x) = V. \tag{1.20}$$

9. Logical deduction. Here is an argument that if $f$ and $g$ are continuous functions, then the composite function $g \circ f$ defined by $(g \circ f)(x) = g(f(x))$ is a continuous function.

   Assume that $f$ and $g$ are continuous. Consider an arbitrary point $a'$ and an arbitrary $\epsilon' > 0$. Since $g$ is continuous at $f(a')$, there exists a $\delta > 0$ such that for all $y$ the condition $|y - f(a')| < \delta$ implies that $|g(y) - g(f(a'))| < \epsilon'$. Call it $\delta_1$. Since $f$ is continuous at $a'$, there exists a $\delta > 0$ such that for all $x$ the condition $|x - a'| < \delta$ implies $|f(x) - f(a')| < \delta_1$. Call it $\delta_2$. Consider an arbitrary $x'$. Suppose $|x' - a'| < \delta_2$. Then $|f(x') - f(a')| < \delta_1$. Hence $|g(f(x')) - g(f(a'))| < \epsilon'$. Thus $|x' - a'| < \delta_2$ implies $|g(f(x')) - g(f(a'))| < \epsilon'$. Since $x'$ is arbitrary, this shows that for all $x$ we have the implication $|x - a'| < \delta_2$ implies $|g(f(x)) - g(f(a'))| < \epsilon'$. It follows that there exists $\delta > 0$ such that all $x$ we have the implication $|x - a'| < \delta$ implies $|g(f(x)) - g(f(a'))| < \epsilon'$. Since $\epsilon'$ is arbitrary, the composite function $g \circ f$ is continuous at $a'$. Since $a'$ is arbitrary, the composite function $g \circ f$ is continuous.

   In the following proof the restrictions that $\epsilon > 0$ and $\delta > 0$ are implicit. They are understood because this is a convention associated with the use of the variables $\epsilon$ and $\delta$.

   Prove that from

$$\forall a \forall \epsilon \exists \delta \forall x\,(|x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon) \tag{1.21}$$

   and

$$\forall b \forall \epsilon \exists \delta \forall y\,(|y - b| < \delta \Rightarrow |g(y) - g(b)| < \epsilon) \tag{1.22}$$

   one can conclude that

$$\forall a \forall \epsilon \exists \delta \forall x\,(|x - a| < \delta \Rightarrow |g(f(x)) - g(f(a))| < \epsilon). \tag{1.23}$$

10. Relaxed natural deduction. Take the proof that the union of open sets is open and put it in outline form, with one formula per line. Indent at every **Suppose** line. Remove the indentation at every **Thus** line. (However, do not indent at a **Let** line.)

11. Draw a picture to illustrate the proof in the preceding problem.

12. Relaxed natural deduction. Prove that for all subsets $U, V$ that ($U$ open $\wedge$ $V$ open) $\Rightarrow$ $U \cap V$ open. Recall that $U \cap V = \bigcap\{U, V\}$ is defined by requiring that for all $y$ that $y \in U \cap V \Leftrightarrow (y \in U \wedge y \in V)$. It may be helpful to use the general fact that for all $t, \epsilon_1 > 0, \epsilon_2 > 0$ there is an implication $t < \min(\epsilon_1, \epsilon_2) \Rightarrow (t < \epsilon_1 \wedge t < \epsilon_2)$. Use a relaxed natural deduction format. Put in outline form, with one formula per line.

13. Draw a picture to illustrate the proof in the preceding problem.

14. Relaxed natural deduction. Recall that for all functions $f$, sets $W$, and elements $t$ we have $t \in f^{-1}[W] \Leftrightarrow f(t) \in W$. Prove that $f$ continuous (with the usual $\epsilon$-$\delta$ definition) implies $\forall U\, (U \text{ open} \Rightarrow f^{-1}[U] \text{ open})$.

15. Relaxed natural deduction. It is not hard to prove the lemma $\{y \mid |y-b| < \epsilon\}$ open. Use this lemma and the appropriate definitions to prove that $\forall U\, (U \text{ open} \Rightarrow f^{-1}[U] \text{ open})$ implies $f$ continuous.

# Chapter 2

# Sets

## 2.1 Zermelo axioms

Mathematical objects include sets, functions, and numbers. It is natural to begin with sets. If $A$ is a set, the expression

$$t \in A \tag{2.1}$$

can be read simply "$t$ *in* $A$". Alternatives are "$t$ is a member of $A$, or "$t$ is an element of $A$", or "$t$ belongs to $A$", or "$t$ is in $A$". The expression $\neg t \in A$ is often abbreviated $t \notin A$ and read "$t$ not in $A$".

If $A$ and $B$ are sets, the expression

$$A \subset B \tag{2.2}$$

is defined in terms of membership by

$$\forall t\,(t \in A \Rightarrow t \in B). \tag{2.3}$$

This can be read simply "$A$ *subset* $B$." Alternatives are "$A$ is included in $B$" or "$A$ is a subset of $B$". (Some people write $A \subseteq B$ to emphasize that $A = B$ is allowed, but this is a less common convention.) It may be safer to avoid such phrases as "$t$ is contained in $A$" or "$A$ is contained in $B$", since here practice is ambiguous. Perhaps the latter is more common.

The following axioms are the starting point for *Zermelo set theory*. They will be supplemented later with the axiom of infinity and the axiom of choice. These axioms are taken by some to be the foundations of mathematics; however they also serve as a review of important constructions.

**Extensionality** A set is defined by its members. For all sets $A, B$

$$(A \subset B \land B \subset A) \Rightarrow A = B. \tag{2.4}$$

**Empty set** Nothing belongs to the empty set.

$$\forall y\, y \notin \emptyset. \tag{2.5}$$

**Unordered pair** For all objects $a, b$ the unordered pair set $\{a, b\}$ satisfies

$$\forall y\, (y \in \{a, b\} \Leftrightarrow (y = a \vee y = b)). \tag{2.6}$$

**Union** If $\Gamma$ is a set of sets, then its union $\bigcup \Gamma$ satisfies

$$\forall x\, (x \in \bigcup \Gamma \Leftrightarrow \exists A\, (A \in \Gamma \wedge x \in A)) \tag{2.7}$$

**Power set** If $X$ is a set, the power set $P(X)$ is the set of all subsets of $X$, so

$$\forall A\, (A \in P(X) \Leftrightarrow A \subset X). \tag{2.8}$$

**Selection** Consider an arbitrary condition $p(x)$ expressed in the language of set theory. If $B$ is a set, then the subset of $B$ consisting of elements that satisfy that condition is a set $\{x \in B \mid p(x)\}$ satisfying

$$\forall y\, (y \in \{x \in B \mid p(x)\} \Leftrightarrow (y \in B \wedge p(y))). \tag{2.9}$$

## 2.2 Comments on the axioms

Usually in a logical language there is the logical relation symbol = and a number of additional relation symbols and function symbols. The Zermelo axioms could be stated in an austere language in which the only non-logical relation symbol is $\in$, and there are no function symbols. The only terms are variables. While this is not at all convenient, it helps to give a more precise formulation of the selection axiom. The following list repeats the axioms in this limited language. However, in practice the other more convenient expressions for forming terms are used.

The philosophy of Zermelo set theory is that everything is a set. However it is helpful at times to think of a hierarchy of objects of different types. An object whose internal structure is of no interest is a *point*. A *set* is defined by its members, which may be points. A *collection* is a set whose members are themselves sets. Sometimes a collection is called a *family*.

**Extensionality**

$$\forall A \forall B\, (\forall t\, (t \in A \Leftrightarrow t \in B) \Rightarrow A = B). \tag{2.10}$$

The axiom of extensionality says that a set is defined by its members. Thus, if $A$ is the set consisting of the digits that occur at least once in my car's license plate 5373, and if $B$ is the set consisting of the odd one digit prime numbers, then $A = B$ is the same three element set. All that matters are that its members are the numbers 7,3,5.

**Empty set**

$$\exists N \forall y \, \neg y \in N. \tag{2.11}$$

By the axiom of extensionality there is only one empty set, and in practice it is denoted by the conventional name $\emptyset$.

**Unordered pair**

$$\forall a \forall b \exists E \forall y \, (y \in E \Leftrightarrow (y = a \vee y = b)). \tag{2.12}$$

By the axiom of extensionality, for each $a, b$ there is only one unordered pair $\{a, b\}$. The unordered pair construction has this name because the order does not matter: $\{a, b\} = \{b, a\}$. Notice that this set can have either one or two elements, depending on whether $a = b$ or $a \neq b$. In the case when it has only one element, it is written $\{a\}$ and is called a singleton set.

If $a, b, c$ are objects, then there is a set $\{a, b, c\}$ defined by the condition that for all $y$

$$y \in \{a, b, c\} \Leftrightarrow (y = a \vee y = b \vee y = c). \tag{2.13}$$

This is the corresponding unordered triple construction. The existence of this object is easily seen by noting that both $\{a, b\}$ and $\{b, c\}$ exist by the unordered pair construction. Again by the unordered pair construction the set $\{\{a, b\}, \{b, c\}\}$ exists. But then by the union construction the set $\bigcup\{\{a, b\}, \{b, c\}\}$ exists. A similar construction works for any finite number of objects.

**Union**

$$\forall \Gamma \exists U \forall x \, (x \in U \Leftrightarrow \exists A \, (A \in \Gamma \wedge x \in A)) \tag{2.14}$$

The standard name for the *union* is $\bigcup \Gamma$. Notice that $\bigcup \emptyset = \emptyset$ and $\bigcup P(X) = X$. A special case of the union construction is $A \cup B = \bigcup\{A, B\}$. This satisfies the property that for all $x$

$$x \in A \cup B \Leftrightarrow (x \in A \vee x \in B). \tag{2.15}$$

Suppose that $C \subset X$ is a given subset of $X$ and that $\Gamma$ is a collection of subsets of $X$. Then $\Gamma$ is said to be a *cover* of $C$ provided $C \subset \bigcup \Gamma$.

If $\Gamma \neq \emptyset$ is a set of sets, then the *intersection* $\bigcap \Gamma$ is defined by requiring that for all $x$

$$x \in \bigcap \Gamma \Leftrightarrow \forall A \, (A \in \Gamma \Rightarrow x \in A) \tag{2.16}$$

The existence of this intersection follows from the union axiom and the selection axiom: $\bigcap \Gamma = \{x \in \bigcup \Gamma \mid \forall A \, (A \in \Gamma \Rightarrow x \in A)\}$.

There is a peculiarity in the definition of $\bigcap \Gamma$ when $\Gamma = \emptyset$. If there is a context where $X$ is a set and $\Gamma \subset P(X)$, then we can define

$$\bigcap \Gamma = \{x \in X \mid \forall A \, (A \in \Gamma \Rightarrow x \in A)\}. \tag{2.17}$$

If $\Gamma \neq \emptyset$, then this definition is independent of $X$ and is equivalent to the previous definition. On the other hand, by this definition $\bigcap \emptyset = X$. This might seem strange, since the left hand side does not depend on $X$. However in most contexts there is a natural choice of $X$, and this is the definition that is appropriate to such contexts. There is a nice symmetry with the case of union, since for the intersection $\bigcap \emptyset = X$ and $\bigcap P(X) = \emptyset$.

A special case of the intersection construction is $A \cap B = \bigcap\{A, B\}$. This satisfies the property that for all $x$

$$x \in A \cap B \Leftrightarrow (x \in A \wedge A \in B). \tag{2.18}$$

If $A \subset X$, the *relative complement* $X \setminus A$ is characterized by saying that for all $x$

$$x \in X \setminus A \Leftrightarrow (x \in X \wedge x \notin A). \tag{2.19}$$

The existence again follows from the selection axiom: $X \setminus A = \{x \in X \mid x \notin A\}$. Sometimes when the set $X$ is understood the *complement* $X \setminus A$ of $A$ is denoted $A^c$.

The constructions $A \cap B$, $A \cup B$, $\bigcap \Gamma$, $\bigcup \Gamma$, and $X \setminus A$ are means of producing objects that have a special relationship to the corresponding logical operations $\wedge, \vee, \forall, \exists, \neg$. A look at the definitions makes this apparent.

Two sets $A, B$ are *disjoint* if $A \cap B = \emptyset$. (In that case it is customary to write the union of $A$ and $B$ as $A \sqcup B$.) More generally, a set $\Gamma \subset P(X)$ of sets is disjoint if for each $A$ in $\Gamma$ and $B \in \Gamma$ with $A \neq B$ we have $A \cap B = \emptyset$. A *partition* of $X$ is a set $\Gamma \subset P(X)$ such that $\Gamma$ is disjoint and $\emptyset \notin \Gamma$ and $\bigcup \Gamma = X$.

**Power set**

$$\forall X \exists P \forall A\, (A \in P \Leftrightarrow \forall t\, (t \in A \Rightarrow t \in X)). \tag{2.20}$$

The *power set* is the set of all subsets of $X$, and it is denoted $P(X)$. Since a large set has a huge number of subsets, this axiom has strong consequences for the size of the mathematical universe.

**Selection** The selection axiom is really an infinite family of axioms, one for each formula $p(x)$ expressed in the language of set theory.

$$\forall B \exists S \forall y\, (y \in S \Leftrightarrow (y \in B \wedge p(y))). \tag{2.21}$$

The selection axiom says that if there is a set $B$, then one may select a subset $\{x \in B \mid p(x)\}$ defined by a condition expressed in the language of set theory. The language of set theory is the language where the only non-logical relation symbol is $\in$. This is why it is important to realize that in principle the other axioms may be expressed in this limited language. The nice feature is that one can characterize the language as the one with just one non-logical relation symbol. However the fact that the separation axiom is stated in this linguistic way is troubling for one who believes that we are talking about a Platonic universe of sets.

Of course in practice one uses other ways of producing terms in the language, and this causes no particular difficulty. Often when the set $B$ is understood the set is denoted more simply as $\{x \mid p(x)\}$. In the defining condition the quantified variable is implicitly restricted to range over $B$, so that the defining condition is that for all $y$ we have $y \in \{x \mid p(x)\} \Leftrightarrow p(y)$. The variables in the set builder construction are bound variables, so, for instance, $\{u \mid p(u)\}$ is the same set as $\{t \mid p(t)\}$.

The famous paradox of Bertrand Russell consisted of the discovery that there is no sensible way to define sets by conditions in a completely unrestricted way. The idea is to note that $y \in x$ is defined for every ordered pair of sets $x, y$. Consider the diagonal where $y = x$. Either $x \in x$ or $x \notin x$. If there were a set $a = \{x \mid x \notin x\}$, then $a \in a$ would be equivalent to $a \notin a$, which is a contradiction.

Say that it is known that for every $x$ in $A$ there is another corresponding object $\phi(x)$ in $B$. Then another useful notation is

$$\{\phi(x) \in B \mid x \in A \wedge p(x)\}. \tag{2.22}$$

This can be defined to be the set

$$\{y \in B \mid \exists x\, (x \in A \wedge p(x) \wedge y = \phi(x)\}. \tag{2.23}$$

So it is a special case. Again, this is often abbreviated as $\{\phi(x) \mid p(x)\}$ when the restrictions on $x$ and $\phi(x)$ are clear. In this abbreviated notion one could also write the definition as $\{y \mid \exists x\, (p(x) \wedge y = \phi(x))\}$.

## 2.3 Ordered pairs and Cartesian product

There is also a very important *ordered pair* construction. If $a, b$ are objects, then there is an object $(a, b)$. This ordered pair has the following fundamental property: For all $a, b, p, q$ we have

$$(a, b) = (p, q) \Leftrightarrow (a = p \wedge b = q). \tag{2.24}$$

If $y = (a, b)$ is an ordered pair, then the first coordinates of $y$ is $a$ and the second coordinate of $y$ is $b$.

Some mathematicians like to think of the ordered pair $(a, b)$ as the set $(a, b) = \{\{a\}, \{a, b\}\}$. The purpose of this rather artificial construction is to make it a mathematical object that is a set, so that one only needs axioms for sets, and not for other kinds of mathematical objects. However this definition does not play much of a role in mathematical practice.

There are also ordered triples and so on. The ordered triple $(a, b, c)$ is equal to the ordered triple $(p, q, r)$ precisely when $a = p$ and $b = q$ and $c = r$. If $z = (a, b, c)$ is an ordered triple, then the coordinates of $z$ are $a$, $b$ and $c$. One can construct the ordered triple from ordered pairs by $(a, b, c) = ((a, b), c)$. The ordered $n$-tuple construction has similar properties.

There are degenerate cases. There is an ordered 1-tuple $(a)$. If $x = (a)$, then its only coordinate is $a$. Furthermore, there is an ordered 0-tuple $(\ ) = 0 = \emptyset$.

Corresponding to these constructions there is a set construction called *Cartesian product*. If $A, B$ are sets, then $A \times B$ is the set of all ordered pairs $(a, b)$ with $a \in A$ and $b \in B$. This is a set for the following reason. Let $U = A \cup B$. Then each of $\{a\}$ and $\{a, b\}$ belongs to $P(U)$. Therefore the ordered pair $(a, b)$ belongs to $P(P(U))$. This is a set, by the power set axiom. So by the selection axiom $A \times B = \{(a, b) \in P(P(U)) \mid a \in A \wedge b \in B\}$ is a set.

One can also construct Cartesian products with more factors. Thus $A \times B \times C$ consists of all ordered triples $(a, b, c)$ with $a \in A$ and $b \in B$ and $c \in C$.

The Cartesian product with only one factor is the set whose elements are the $(a)$ with $a \in A$. There is a natural correspondence between this somewhat trivial product and the set $A$ itself. The correspondence is that which associates to each $(a)$ the corresponding coordinate $a$. The Cartesian product with zero factors is a set $1 = \{0\}$ with precisely one element $0 = \emptyset$.

There is a notion of sum of sets that is dual to the notion of product of sets. This is the *disjoint union* of two sets. The idea is to attach labels to the elements of $A$ and $B$. Thus, for example, for each element $a$ of $A$ consider the ordered pair $(0, a)$, while for each element $b$ of $B$ consider the ordered pair $(1, b)$. Then even if there are elements common to $A$ and $B$, their tagged versions will be distinct. Thus the sets $\{0\} \times A$ and $\{1\} \times B$ are disjoint. The disjoint union of $A$ and $B$ is the set $A + B$ such that for all $y$

$$y \in A + B \Leftrightarrow (y \in \{0\} \times A \vee y \in \{1\} \times B). \tag{2.25}$$

One can also construct disjoint unions with more summands in the obvious way.

## 2.4 Relations and functions

A *relation* $R$ between sets $A$ and $B$ is a subset of $A \times B$. A *function* $F$ from $A$ to $B$ is a relation with the following two properties:

$$\forall x \exists y\, (x, y) \in F. \tag{2.26}$$

$$\forall y \forall y'\, (\exists x\, ((x, y) \in F \wedge (x, y') \in F) \Rightarrow y = y'). \tag{2.27}$$

In these statements the variable $x$ is restricted to $A$ and the variables $y, y'$ are restricted to $B$. A function $F$ from $A$ to $B$ is a *surjection* if

$$\forall y \exists x\, (x, y) \in F. \tag{2.28}$$

A function $F$ from $A$ to $B$ is an *injection* if

$$\forall x \forall x'\, (\exists y\, ((x, y) \in F \wedge (x', y) \in F) \Rightarrow x = x'). \tag{2.29}$$

Notice the same pattern in these definitions as in the two conditions that define a function. As usual, if $F$ is a function, and $(x, y) \in F$, then we write $F(x) = y$.

In this view a function is regarded as being identical with its graph as a subset of the Cartesian product. On the other hand, there is something to be said for a point of view that makes the notion of a function just as fundamental as the notion of set. In that perspective, each function from $A$ to $B$ would have a graph that would be a subset of $A \times B$. But the function would be regarded as an operation with an input and output, and the graph would be a set that is merely one means to describe the function.

There is a useful function builder notation that corresponds to the set builder notation. Say that it is known that for every $x$ in $A$ there is another corresponding object $\phi(x)$ in $B$. Then another useful notation is

$$\{x \mapsto \phi(x) : A \to B\} = \{(x, \phi(x)) \in A \times B \mid x \in A\}. \tag{2.30}$$

This is an explicit definition of a function from $A$ to $B$. This could be abbreviated as $\{x \mapsto \phi(x)\}$ when the restrictions on $x$ and $\phi(x)$ are clear. The variables in such an expression are of course bound variables, so, for instance, the squaring function $u \mapsto u^2$ is the same as the squaring function $t \mapsto t^2$.

## 2.5 Number systems

The axiom of infinity states that there is an infinite set. In fact, it is handy to have a specific infinite set, the set of all natural numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$. The mathematician von Neumann gave a construction of the natural numbers that is perhaps too clever to be taken entirely seriously. He defined $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, and so on. Each natural number is the set of all its predecessors. Furthermore, the operation $s$ of adding one has a simple definition:

$$s(n) = n \cup \{n\}. \tag{2.31}$$

Thus $4 = 3 \cup \{3\} = \{0, 1, 2\} \cup \{3\} = \{0, 1, 2, 3\}$. Notice that each of these sets representing a natural number is a finite set. There is as yet no requirement that the natural numbers may be combined into a single set.

This construction gives one way of formulating the *axiom of infinity*. Say that a set $I$ is inductive if $0 \in I$ and $\forall n\,(n \in I \Rightarrow s(n) \in I)$. The axiom of infinity says that there exists an inductive set. Then the set $\mathbb{N}$ of natural numbers may be defined as the intersection of the inductive subsets of this set.

According to this definition the natural number system $\mathbb{N}\{0, 1, 2, 3, \ldots\}$ has 0 as an element. It is reasonable to consider 0 as a natural number, since it is a possible result of a counting process. However it is sometimes useful to consider the set of natural numbers with zero removed. In this following we denote this set by by $\mathbb{N}_+ = \{1, 2, 3, \ldots\}$.

According to the von Neuman construction, the natural number $n$ is defined by $n = \{0, 1, 2, \ldots, n-1\}$. This is a convenient way produce an $n$ element index set, but in other contexts it can also be convenient to use $\{1, 2, 3, \ldots, n\}$.

This von Neumann construction is only one way of thinking of the set of natural numbers $\mathbb{N}$. However, once we have this infinite set, it is not difficult to construct a set $\mathbb{Z}$ consisting of all integers $\{\ldots,-3,-2,-1,0,1,2,3,\ldots\}$. Furthermore, there is a set $\mathbb{Q}$ of rational numbers, consisting of all quotients of integers, where the denominator is not allowed to be zero. The next step after this is to construct the set $\mathbb{R}$ of real numbers. This is done by a process of completion, to be described later. The transition from $\mathbb{Q}$ to $\mathbb{R}$ is the transition from algebra to analysis. The result is that it is possible to solve equations by approximation rather than by algebraic means.

After that, next important number system is $\mathbb{C}$, the set of complex numbers. Each complex number is of the form $a + bi$, where $a, b$ are real numbers, and $i^2 = -1$. Finally, there is $\mathbb{H}$, the set of quaternions. Each quaternion is of the form $t + ai + bj + ck$, where $t, a, b, c$ are real numbers. Here $i^2 = -1, j^2 = -1, k^2 = -1$, $ij = k$, $jk = i$, $ki = j$, $ji = -k$, $kj = -i$, $ik = -j$. A pure quaternion is one of the form $ai + bj + ck$. The product of two pure quaternions is $(ai+bj+ck)(a'i+b'j+c'k) = -(aa'+bb'+cc')+(bc'-cb')i+(ca'-ac')j+(ab'-ba')k$. Thus quaternion multiplication includes both the dot product and the cross product in a single operation.

In summary, the number systems of mathematics are $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{H}$. The systems $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ each have a natural linear order, and there are natural order preserving injective functions from $\mathbb{N}$ to $\mathbb{Z}$, from $\mathbb{Z}$ to $\mathbb{Q}$, and from $\mathbb{Q}$ to $\mathbb{R}$. The natural algebraic operations in $\mathbb{N}$ are addition and multiplication. In $\mathbb{Z}$ they are addition, subtraction, and multiplication. In $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{H}$ they are addition, subtraction, multiplication, and division by non-zero numbers. In $\mathbb{H}$ the multiplication and division are non-commutative. The number systems $\mathbb{R}, \mathbb{C}, \mathbb{H}$ have the completeness property, and so they are particularly useful for analysis.

## 2.6 The extended real number system

In analysis it is sometimes useful to have an extended real number system, consisting of $\mathbf{R}$ together with two extra points $+\infty$ and $-\infty$. The order structure is that $-\infty \leq a \leq +\infty$ for all real numbers $a$. The purpose of this system is not to clarify the notion of infinity; rather it is to have an extended real number system with a greatest element and a least element. It is possible to talk about continuity in the extended real number system, for instance by mapping it to $[-1, 1]$ via the hyperbolic tangent function.

The arithmetic in $[-\infty, +\infty]$ is worth some discussion. If $a > 0$ is an extended real number, it is natural to define $a \cdot (\pm\infty) = \pm\infty$. Similarly, if $a < 0$ is an extended real number, it is natural to define $a \cdot (\pm\infty) = \mp\infty$. For zero there is a *zero times infinity convention* that is often used in analysis:

$$0 \cdot (\pm\infty) = 0. \tag{2.32}$$

The difficult with this convention is that it makes multiplication discontinuous. To see, this, note that while $(1/n) \cdot n = 1 \to 1$ as $n \to \infty$, the limit of $1/n$ times the limit of $n$ is $0 \cdot (+\infty) = 0$.

Addition is even worse. While $a + (+\infty) = +\infty$ for all $a > -\infty$, and $a + (-\infty) = -\infty$ for all $a < +\infty$, the expressions $(-\infty) + (+\infty)$ and $(+\infty) + (-\infty)$

are undefined. This *infinity minus infinity problem* is the source of many of the most interesting phenomena of analysis. On the other hand, where addition is defined, it is continuous.

For some purposes one can get around the infinity minus infinity problem by using the systems $(-\infty, +\infty]$ or by using the system $[-\infty, +\infty)$. For either of these systems both addition and multiplication are defined.

## 2.7 Supplement: Construction of number systems

This section gives an outline of the construction of various number systems. The purpose of these constructions is merely to show that the existence of number systems follows from the assumptions of set theory. There is no claim that they explain what numbers really are.

It is assumed that there is a natural number system $\mathbb{N}$ with a zero 0 and a successor operation $s$ such that $s(n)$ is the next integer above $n$. (In the von Neumann construction 0 is the empty set and $s(n) = n \cup \{n\}$.) The characteristic feature is the induction property: If $J$ is inductive, then $\mathbb{N} \subset J$.

Addition in $\mathbb{N}$ may then be characterized inductively by $m + 0 = 0$ and $m+s(n) = s(m+n)$. Similarly, multiplication may be characterized by $m \cdot 0 = 0$ and $m \cdot s(n) = m \cdot n + m$. It is a tedious process to verify the properties of the operations by induction, but it can be done.

Next is the construction of the integers $\mathbb{Z}$. Consider the product space $\mathbb{N} \times \mathbb{N}$. The intuitive idea is that each point $(m, n)$ in this space is to define an integer $k$ with $k = n - m$. This idea leads to the following definitions. If the ordered pairs $(m, n)$ and $(m', n')$ are in this space, then their sum is the vector sum $(m, n) + (m', n') = (m + m', n + n')$. The additive inverse of $(m, n)$ is defined (somewhat unusually) as $-(m, n) = (n, m)$. The product of $(m, n)$ and $(m', n')$ has components given by inner products $((m, n) \cdot (m', n'), (n, m) \cdot (m', n'))$. This works out to be $(mm' + nn', nm' + mn')$. Two such ordered pairs $(m, n)$ and $(m', n')$ are said to be equivalent if $m+n' = n+m'$. This relation of equivalence partitions $\mathbb{N} \times \mathbb{N}$ into a disjoint union of sets. Each such set of ordered pairs defines an integer.

If we consider $\mathbb{N} \times \mathbb{N}$ geometrically, then each integer is the graph of a line with slope one. The sum of two integers is determined by taking the vector sum of any two points on the two lines. The inverse is obtained by reflecting points across the diagonal line passing through the origin. The product is obtained by taking a point to another point with coordinates given by the product formula above, which only involves inner products and the reflection across the diagonal.

Thus, for example, the integer 3 is defined as the right shift $3 = \{(0, 3), (1, 4), (2, 5), \ldots\}$ on the natural numbers. The integer $-5$ is defined as the left shift $-5 = \{(5, 0), (6, 1), (7, 2), \ldots\}$. To add the integer 3 to the integer -5, take a representative $(1, 4)$ and another representative $(7, 2)$. Add to get $(8, 6)$, which represents $-2$.

Next is the construction of the rational numbers $\mathbb{Q}$. Start with the integers $\mathbb{Z}$. Let $\mathbb{Z}^*$ be $\mathbb{Z}$ with the integer zero removed. Consider the product space $\mathbb{Z}^* \times \mathbb{Z}$. The intuitive idea is that each point $(j, k)$ with $j \neq 0$ in this space is to define a rational number $q$ with $q = k/j$. This idea leads to the following definitions. If the ordered pairs $(j, k)$ and $(j', k')$ are in this space, then their product is the pointwise product $(j \cdot j', k \cdot k')$. (Multiply the denominators; multiply the numerators.) The multiplicative inverse of $(j, k)$ for $j, k \neq 0$ is defined as the reflection $(k, j)$. (Interchange numerator and denominator.) Two such ordered pairs $(j, k)$ and $(j', k')$ are said to be equivalent if $j \cdot k' = k \cdot j'$. This relation of equivalence partitions $\mathbb{Z}^* \times \mathbb{Z}$ into a disjoint union of sets. Each such set of ordered pairs defines a rational number.

If we consider $\mathbb{Z}^* \times \mathbb{Z}$ geometrically, then each rational is the graph of a line through the origin. Thus, for example, the rational number $4/3$ is defined as $4/3 = \{(3, 4), (-3, -4), (6, 8), (-6, -8), \ldots\}$. The rational number $-5/2$ is defined as $-5/2 = \{(2, -5), (-2, 5), (4, -10), (-4, 10), \ldots\}$. To multiply integer $4/3$ by $-5/2$ , take a representative $(3, 4)$ and another representative $(4, -10)$. Use the multiplication rule to get $(12, -40)$, which represents $-10/3$. This discussion has left out addition of rational numbers; this dismal story is all too well known.

The construction of the rational numbers from the integers is quite parallel to the construction of the integers from the natural numbers. The construction of the real numbers is quite another matter. This construction is explored in detail in the chapter on ordered sets. However here is the short version.

Most real numbers are not rational, but it is tricky to prove that individual real numbers are not rational. Here is an example of a real number that is not rational. Let $s_n = 1 + 1 + 1/2 + 1/6 + \cdots 1/n!$. This is clearly a rational number, and the real number $e$ is the real number that is the supremum or least upper bound of the $s_n$. The property that makes $e$ irrational is that it can be approximated very closely by a rational number that is not equal to it. Specifically, by Taylor's theorem with remainder,

$$e = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \cdots \frac{1}{n!} + \frac{e^{c_n}}{(n+1)!}, \tag{2.33}$$

where $0 < c_n < 1$ and hence $1 < e^{c_n} < 3$. This shows that $e$ is a rational number (the partial sum) plus a very small number (the remainder). Suppose that $e$ were rational, $e = p/q$, with integer $p, q > 0$. Then

$$n!p = qn! \left[1 + 1 + \frac{1}{2} + \frac{1}{6} + \cdots \frac{1}{n!}\right] + q\frac{e^{c_n}}{n+1}, \tag{2.34}$$

The numbers $p$ and $q$ are fixed. Choose $n$ large enough so that $n+1 > 3q$. Then $0 < qe^{c_n}/(n+1) < 1$. However all the other terms are integers. This leads to a contradiction. So $e$ is not rational.

The problem is to construct the real numbers $\mathbb{R}$ from the rational numbers. This is the crucial step in the passage from algebra to analysis. In the construction each real number will be a set of rational numbers, but only certain sets are used. That is, the each real number will belong to the (huge) set $P(\mathbb{Q})$, but not all sets will be used.

If $A$ is a set of rational numbers, an upper bound for $A$ is a rational number $q$ such that each $x$ in $A$ satisfies $x \leq A$. Let $\uparrow A$ be the set of all upper bounds for $A$. If $B$ is a set of rational numbers, a lower bound for $B$ is a rational number $p$ such that each $x$ in $B$ satisfies $x \leq p$. Let $\downarrow B$ be the set of all lower bounds for $B$. Clearly $A \subset \downarrow\uparrow A$. Call a set $A$ of rational numbers a lower Dedekind cut if $\downarrow\uparrow A \subset A$. Then $\mathbb{R}$ is defined as the set of all lower Dedekind cuts $A$ such that $A \neq \emptyset$ and $\uparrow A \neq \emptyset$.

If $r$ is a rational number, then the set of lower bounds for $r$ is a lower Dedekind cut. So each rational number defines a real number. For instance, the rational number $4/3$ defines the real number consisting of all rational numbers $p$ with $p \leq 4/3$. However there are real numbers that are not rational numbers. For instance, let $S$ be the set of all rational numbers of the form $s_n = 1 + 1 + 1/2 + 1/6 + \cdots 1/n!$. Let $A = \downarrow\uparrow S$. Then $A$ is a real number that does not come from a rational number. The intuition is that $\uparrow S$ consists of the rational upper bounds for $e$, and $A = \downarrow\uparrow S$ consists of the rational lower bounds for $e$.

If $A$ and $A'$ are real numbers, regarded as lower Dedekind cuts, then $A \leq A'$ means $A \subset A'$. This defines the order structure on real numbers. Defining the additive structure is easy; the sum $A + A'$ of two lower Dedekind cuts consists of the set of all rational sums $x + x'$ with $x$ in $A$ and $x'$ in $A'$. Defining the multiplicative structure is more awkward, but it can be done.

The real number system $\mathbb{R}$ has the property that it is *boundedly complete*. That is, every non-empty subset of $\mathbb{R}$ that is bounded above has a supremum (least upper bound). It is not hard to see this from the construction. Each real number in the subset is itself a lower Dedekind cut. The union of these may or may not be a lower Dedekind cut, but there is a smallest lower Dedekind cut of which the union is a subset. This is the supremum.

# Problems

1. Say $X$ has $n$ elements. How many elements are there in $P(X)$?

2. Say $X$ has $n$ elements. Denote the number of subsets of $X$ with exactly $k$ elements by $\binom{n}{k}$. Show that $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$ and that

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{2.35}$$

   Use this to make a table of $\binom{n}{k}$ up to $n = 7$.

3. Say that $X$ has $n$ elements. Denote the number of partitions of $X$ into exactly $k$ non-empty disjoint subsets by $S(n, k)$. This is a Stirling number of the second kind. Show that $S(n, 1) = 1$ and $S(n, n) = 1$ and

$$S(n, k) = S(n-1, k-1) + kS(n-1, k). \tag{2.36}$$

   Use this to make a table of $S(n, k)$ up to $n = 5$.

4. How many functions are there from an $n$ element set to a $k$ element set?

5. How many injective functions are there from an $n$ element set to a $k$ element set?

6. How many surjective functions are there from an $n$ element set to a $k$ element set?

7. Show that $m^n = \sum_{k=0}^{m} \binom{m}{k} k! S(n,k)$.

8. Let $B_n = \sum_{k=0}^{n} S(n,k)$ be the number of partitions of an $n$ element set. Show that $B_n$ is equal to the expected number of functions from an $n$ element set to an $m$ element set, where $m$ has a Poisson probability distribution with mean one. That is, show that
$$B_n = \sum_{m=0}^{\infty} m^n \frac{1}{m!} e^{-1}. \tag{2.37}$$

9. Let $B_n$ be the number of partitions of an $n$ element set into disjoint non-empty sets. Thus $B_0 = 1$, $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$, $B_5 = 52$, and so on. One can try to write $B_{n+1}$ as a linear combination of $B_0, \ldots, B_n$. For instance,
$$B_3 = B_2 + 2B_1 + B_0 \tag{2.38}$$
and
$$B_4 = B_3 + 3B_2 + 3B_1 + B_0 \tag{2.39}$$
and
$$B_5 = B_4 + 4B_3 + 6B_2 + 4B_1 + B_0. \tag{2.40}$$
Find the general pattern. Prove that it holds for arbitrary $n$. Hint: Consider a set $S$ with $n+1$ points, with a selected point $p$ in it.

10. Is it possible to have sets $A$ and $B$, a function $f : A \to B$ that is an injection but not a surjection, and a function $g : A \to B$ that is a surjection but not an injection. Explain.

11. A totally ordered set is densely ordered if between every two distinct points there is another point. Thus $\mathbb{Q}$ is densely ordered, and also $\mathbb{R}$ is densely ordered. Show that between every two distinct points of $\mathbb{Q}$ there is a point of $\mathbb{R}$ that is irrational.

12. Is it true that between every two distinct points of $\mathbb{R}$ there is a point of $\mathbb{Q}$? Discuss.

13. Define a map from $\mathbb{R}$ to $P(\mathbb{Q})$ by $j(x) = \{r \in \mathbb{Q} \mid r \leq x\}$. Prove that $j$ is injective.

# Chapter 3

# Relations, functions, dynamical systems

## 3.1 Identity, composition, inverse, intersection

A *relation* $R$ between sets $A$ and $B$ is a subset of $A \times B$. In this context one often writes $xRy$ instead of $(x, y) \in R$, and says that $x$ is related to $y$ by the relation. Often a relation between $A$ and $A$ is called a relation on the set $A$.

There is an important relation $I_A$ on $A$, namely the *identity relation* relation consisting of all ordered pairs $(x, x)$ with $x \in A$. That is, for $x$ and $y$ in $A$, the relation $xI_Ay$ is equivalent to $x = y$.

Given an relation $R$ between $A$ and $B$ and a relation $S$ between $B$ and $C$, there is a relation $S \circ R$ between $A$ and $C$ called the *composition*. It is defined in such a way that $x(S \circ R)z$ is equivalent to the existence of some $y$ in $B$ such that $xRy$ and $ySz$. Thus if $R$ relates $A$ to $B$, and $S$ relates $B$ to $C$, then $S \circ R$ relates $A$ to $C$. In symbols,

$$S \circ R = \{(x, z) \mid \exists y\, (xRy \wedge ySz)\}. \tag{3.1}$$

Notice the order in which the factors occur, which accords with the usual convention for functions. For functions it is usual to use such a notation to indicate that $R$ acts first, and then $S$. This is perhaps not the most natural convention for relations, so in some circumstances it might be convenient to define another kind of composition in which the factors are written in the opposite order.

There are two more useful operations on relations. If $R$ is a relation between $A$ and $B$, then there is an *inverse* relation $R^{-1}$ between $B$ and $A$. It consists of all the $(y, x)$ such that $(x, y)$ is in $R$. That is, $yR^{-1}x$ is equivalent to $xRy$.

Finally, if $R$ and $S$ are relations between $A$ and $B$, then there is a relation $R \cap S$. This is also a useful operation. Notice that $R \subset S$ is equivalent to $R \cap S = R$.

Sometimes if $X \subset A$ one writes $R[X]$ for the image of $X$ under $R$, that is,

$$R[X] = \{y \mid \exists x\, (x \in X \wedge xRy)\}. \tag{3.2}$$

Also, if $a$ is in $A$, it is common to write $R[a]$ instead of $R[\{a\}]$. Thus $y$ is in $R[a]$ if $aRy$.

## 3.2 Picturing relations

There are two common ways of picturing a relation $R$ between $A$ and $B$. One way is to draw the product space $A \times B$ and sketch the set of points $(x, y)$ in $R$. This is the *graph* of the relation. The other way is to draw the disjoint union $A + B$ and for each $(x, y)$ in $R$ sketch an arrow from $x$ to $y$. This is the *cograph* of the relation.

## 3.3 Equivalence relations

Consider a relation $R$ on $A$. The relation $R$ is *reflexive* if $I_A \subset R$. The relation $R$ is *symmetric* if $R = R^{-1}$. The relation $R$ is *transitive* if $R \circ R \subset R$. A relation that is reflexive, symmetric, and transitive (RST) is called an *equivalence relation*.

**Theorem 3.1** *Consider a set $A$. Let $\Gamma$ be a partition of $A$. Then there is a corresponding equivalence relation $E$, such that $(x, y) \in E$ if and only if for some subset $U$ in $\Gamma$ both $x$ in $U$ and $y$ in $U$. Conversely, for every equivalence relation $E$ on $A$ there is a unique partition $\Gamma$ of $A$ that gives rise to the relation in this way.*

The sets in the partition defined by the equivalence relation are called the *equivalence classes* of the relation.

## 3.4 Generating relations

**Theorem 3.2** *For every relation $R$ on $A$, there is a smallest transitive relation $R^T$ such that $R \subset R^T$. This is the transitive relation generated by $R$.*

**Theorem 3.3** *For every relation $R$ on $A$, there is a smallest symmetric and transitive relation $R^{ST}$ such that $R \subset R^{ST}$. This is the symmetric and transitive relation generated by $R$.*

**Theorem 3.4** *For every relation $R$ on $A$, there is a smallest equivalence relation $E = R^{RST}$ such that $R \subset E$. This is the equivalence relation generated by $R$.*

Proof: The proofs of these theorems all follow the same pattern. Here is the proof of the last one. Let $R$ be a relation on $A$, that is, let $R$ be a subset of $A \times A$. Let $\Delta$ be the set of all equivalence relations $S$ with $R \subset S$. Then since $A \times A \in \Delta$, it follows that $\Delta$ is non-empty. Let $E = \bigcap \Delta$. Now note three facts. The intersection of a set of transitive relations is transitive. The intersection of a set of symmetric relations is symmetric. The intersection of a set of reflexive relations is reflexive. It follows that $E$ is transitive, reflexive, and symmetric. This is the required equivalence relation. □

This theorem shows that by specifying a relation $R$ one also specifies a corresponding equivalence relation $E$. This can be a convenient way of describing an equivalence relation.

## 3.5 Ordered sets

A relation $R$ on $A$ is *antisymmetric* if $R \cap R^{-1} \subset I_A$. This just says that $\forall x \forall y\, ((x \leq y \wedge y \leq x) \Rightarrow x = y)$. A *ordering* of $A$ is a relation that is reflexive, antisymmetric, and transitive (RAT). Ordered sets will merit further study. Here is one theorem about how to describe them.

**Theorem 3.5** *Consider a relation $R$ such that there exists an order relation $S$ with $R \subset S$. Then there exists a smallest order relation $P = R^{RT}$ with $R \subset P$.*

Proof: Let $R$ be a relation on $A$ that is a subset of some order relation. Let $\Delta$ be the set of all such order relations $S$ with $R \subset S$. . By assumption $\Delta \neq \emptyset$. Let $P = \bigcap \Delta$. Argue as in the case of an equivalence relation. A subset of an antisymmetric relation is antisymmetric. (Note that for an non-empty set of sets the intersection is a subset of the union.) The relation $P$ is the required order relation. □

The above theorem gives a convenient way of specifying an order relation $P$. For example, if $A$ is finite, then $P$ is generated by the successor relation $R$.

A *totally ordered* (or *linearly ordered*) set is an ordered set such that the order relation satisfies $R \cup R^{-1} = A \times A$. Thus just says that $\forall x \forall y\, (x \leq y \vee y \leq x)$. A *well-ordered set* is a linearly ordered set with the property that each non-empty subset has a least element.

A *rooted tree* is an ordered set with a least element, the root, such that for each point in the set, the elements below the point form a well-ordered set.

## 3.6 Functions

A relation $F$ from $A$ to $B$ is a *total relation* if $I_A \subset F^{-1} \circ F$. It is a *partial function* if $F \circ F^{-1} \subset I_B$. It is a *function* if it is both a total relation and a partial function (that is, it is a total function).

**Proposition 3.6** *A relation $F$ from $A$ to $B$ is total if and only if for each $S \subset A$*

$$S \subset F^{-1}[F[S]]. \tag{3.3}$$

*This is true if and only if for every $S \subset A$ and $T \subset B$ we have*

$$F[S] \subset T \Rightarrow S \subset F^{-1}[T]. \tag{3.4}$$

Proof: A relation $F$ is total if and only if for each $a$ in $A$ there exists a $b$ in $B$ with $aFb$. The first result comes from noting that $c$ is in $F^{-1}[F[S]]$ if and only if there is an $a$ in $S$ and a $b$ such that $cFb$ and $aFb$. The second result follows from the first. □

**Proposition 3.7** *A relation $F$ from $A$ to $B$ is a partial function if and only if for each $T \subset B$*

$$F[F^{-1}[T]] \subset T. \tag{3.5}$$

*This is true if and only if for every $S \subset A$ and $T \subset B$ we have*

$$S \subset F^{-1}[T] \Rightarrow F[S] \subset T. \tag{3.6}$$

Proof: A relation $F$ is a partial function if and only if for every $b$ in $B$ and $d$ in $B$ for which there exists $a$ in $A$ with $aFb$ and $aFd$ we have $d = b$. The first result comes from noting that $d$ is in $F[F^{-1}[T]$ if and only if there exists $b$ in $T$ and $a$ such that $aFb$ and $aFd$. The second result follows from the first. □

These two propositions above combine to give the following remarkable characterization of what it means for a relation to be a function. This property is used throughout analysis.

**Proposition 3.8** *A relation $F$ from $A$ to $B$ is a function if and only if for every $S \subset A$ and $T \subset B$ we have*

$$F[S] \subset T \Leftrightarrow S \subset F^{-1}[T]. \tag{3.7}$$

A function $F$ is an *injective function* if it is a function and $F^{-1}$ is a partial function. A function $F$ is a *surjective function* if it is a function and also $F^{-1}$ is a total relation. It is a *bijective function* if it is both an injective function and a surjective function. For a bijective function $F$ the inverse relation $F^{-1}$ is a function from $B$ to $A$, in fact a bijective function.

## 3.7 Relations inverse to functions

**Lemma 3.9** *Let $F$ be a relation that is a function from $A$ to $B$, and let $F^{-1}$ be the inverse relation. Then the sets $F^{-1}[b]$ for $b$ in the range of $F$ form a partition of $A$, and $F^{-1}[b] = \emptyset$ for $b$ not in the range of $F$. If $V$ is a subset of $B$, then $F^{-1}[V]$ is the union of the disjoint sets $F^{-1}[b]$ for $b$ in $V$.*

This lemma immediately gives the following remarkable and important theorem on inverse images. Contrast this theorem with the proposition on images that follows.

**Theorem 3.10** *Let $F$ be a relation that is a function from $A$ to $B$, and let $F^{-1}$ be the inverse relation. Then $F^{-1}$ respects the set operations of union, intersection, and complement. Thus:*

1. *If $\Gamma$ is a set of subsets of $B$, then $F^{-1}[\bigcup \Gamma] = \bigcup \{F^{-1}[V] \mid V \in \Gamma\}$.*

2. *If* $\Gamma$ *is a set of subsets of* $B$*, then* $F^{-1}[\bigcap \Gamma] = \bigcap\{F^{-1}[V] \mid V \in \Gamma\}$.

3. *If* $V$ *is a subset of* $B$*, then* $F^{-1}[B \setminus V] = A \setminus F^{-1}[V]$.

**Proposition 3.11** *Let* $F$ *be a relation from* $A$ *to* $B$*. Then the action of* $F$ *on subsets respects the union operation. Thus:*

1. *If* $\Gamma$ *is a set of subsets of* $B$*, then* $F[\bigcup \Gamma] = \bigcup\{F[V] \mid V \in \Gamma\}$.

One can ask why the inverse image operation on sets has better properties than the image operation. Part of the reason is as follows. Let $f : A \to B$ be a function. Suppose that $T \subset B$. To check that $x \in f^{-1}[T]$ we just have to check that $f(x) \in T$, which requires a function evaluation. On the other hand, suppose that $S \subset A$. To check that $y \in f[S]$ we need to check that $\exists x \in S\, f(x) = y$, and this requires showing that an equation has a solution.

## 3.8 Dynamical systems

Consider a function $F$ from $A$ to $A$. Such a function is often called a *dynamical system*. Thus if $a$ is the present state of the system, at the next stage the state is $F(a)$, and at the following stage after that the state is $F(F(a))$, and so on.

The *orbit* of a point $a$ in $A$ is $F^{RT}[a]$, the image of $a$ under the relation $F^{RT}$. This is the entire future history of the system (including the present), when it is started in the state $a$. Each orbit $S$ is invariant under $F$, that is, $F[S] \subset S$. If $b$ is in the orbit of $a$, then we say that $a$ leads to $b$.

The simplest way to characterize the orbit of $a$ is as the set $\{a, F(a), F(F(a)), F(F(F(a))), \ldots\}$, that is, the set of $F^{(n)}(a)$ for $n \in \mathbb{N}$, where $F^{(n)}$ is the $n$th iterate of $F$. (The $n$th iterate of $F$ is the composition of $F$ with itself $n$ times.)

**Theorem 3.12** *Let* $F : A \to A$ *be a function. Each orbit of* $a$ *under* $F$ *is either finite and consists of a sequence of points that eventually enters a periodic cycle, or it is an infinite sequence of distinct points.*

In the finite case the orbit may be described as having the form of a lasso. Special cases of the lasso are a cycle and a single point.

## 3.9 Picturing dynamical systems

Since a dynamical system is a function $F : A \to A$, there is a peculiarity that the domain and the target are the same space. However this gives a nice way of picturing orbits.

One method is to plot the graph of $F$ as a subset of $A \times A$, and use this to describe the dynamical system as acting on the diagonal. For each $x$ in the orbit, start with the point $(x, x)$ on the diagonal. Draw the vertical line from $(x, x)$ to $(x, F(x))$ on the graph, and then draw the horizontal line from $(x, F(x))$ to

$(F(x), F(x))$ back on the diagonal. This process gives a broken line curve that gives a picture of the dynamical system acting on the diagonal.

A method that is more compatible with the cograph point of view is to look at the set $A$ and draw an arrow from $x$ to $F(x)$ for each $x$ in the orbit.

## 3.10 Structure of dynamical systems

Let $F : A \to A$ be a function. Then $A$ is a disjoint union of equivalence classes under the equivalence relation $F^{RST}$ generated by $F$. The following theorem gives a more concrete way of thinking about this equivalence relation.

**Theorem 3.13** *Let $F : A \to A$ be a function. Say that $aEb$ if and only if the orbit of $a$ under $F$ has a non-empty intersection with the orbit of $b$ under $F$. Then $E$ is an equivalence relation, and it is the equivalence relation generated by $F$.*

Proof: To show that $E$ is an equivalence relation, it is enough to show that it is reflexive, symmetric, and transitive. The first two properties are obvious. To prove that it is transitive, consider points $a, b, c$ with $aEb$ and $bEc$. Then there are $m, n$ with $F^{(m)}(a) = F^{(n)}(b)$ and there are $r, s$ with $F^{(r)}(b) = F^{(s)}(c)$. Suppose that $n \leq r$. Then $F^{(m+r-n)}(a) = F^{(r)}(b) = F^{(s)}(c)$. Thus in that case $aEc$. Instead suppose that $r \leq n$. A similar argument shows that $aEc$. Thus it follows that $aEc$.

It is clear that $E$ is an equivalence relation with $F \subset E$. Let $E'$ be an arbitrary equivalence relation with $F \subset E'$. Say that $aEb$. Then there is a $c$ with $aF^{RT}c$ and $bF^{RT}c$. Then $aE'c$ and $bE'c$. Since $E'$ is an equivalence relation, it follows that $cE'b$ and hence $aE'b$. So $E \subset E'$. This shows that $E$ is the smallest equivalence relation $E'$ with $F \subset E'$. That is, $E$ is the equivalence relation generated by $F$. □

Each equivalence class of a dynamical system $F$ is invariant under $F$. Thus to study a dynamical system one needs only to look at what happens on each equivalence class.

One can think of a dynamical system as reversible if the function is bijective, as conservative if the function is injective, and as dissipative in the general case. The following theorem describes the general case. There are two possibilities. Either there is eventually stabilization at a periodic cycle. Or the dissipation goes on forever.

**Theorem 3.14** *Let $F : A \to A$ be a function. Then on each equivalence class $F$ acts in one of two possible ways. Case 1. Each point in the class has a finite orbit. In this case there is a unique cycle with some period $n \geq 1$ included in the class. Furthermore, the class itself is partitioned into $n$ trees, each rooted at a point of the cycle, such that the points in each tree lead to the root point without passing through other points of the cycle. Case 2. Each point in the class has an infinite orbit. Then the points that lead to a given point in the class form a tree rooted at the point.*

Proof: If $a$ and $b$ are equivalent, then they each lead to some point $c$. If $a$ leads to a cycle, then $c$ leads to a cycle. Thus $b$ leads to a cycle. So if one point in the equivalence class leads to a cycle, then all points lead to a cycle. There can be only one cycle in an equivalence class.

In this case, consider a point $r$ on the cycle. Say that a point leads directly to $r$ if it leads to $r$ without passing through other points on the cycle. The point $r$ together with the points that lead directly to $r$ form a set $T(r)$ with $r$ as the root. A point $q$ in $T(r)$ is said to be below a point $p$ in $T(r)$ when $p$ leads to $q$. There cannot be distinct points $p, q$ on $T(r)$ with $q$ below $p$ and $p$ below $q$, since then there would be another cycle. Therefore $T(r)$ is an ordered set. If $p$ is in $T(r)$, the part of $T(r)$ below $p$ is a finite linearly ordered set, so $T(r)$ is a tree. Each point $a$ in the equivalence class leads directly to a unique point $r$ on the cycle. It follows that the trees $T(r)$ for $r$ in the cycle form a partition of the equivalence class.

The other case is when each point in the class has an infinite orbit. There can be no cycle in the equivalence class. Consider a point $r$ in the class. The same kind of argument as in the previous case shows that the set $T(r)$ of points that lead to $r$ is a tree. □

The special case of conservative dynamical systems given by an injective function is worth special mention. In that case there can be a cycle, but no tree can lead to the cycle. In the case of infinite orbits, the tree that leads to a point has only one branch (infinite or finite).

**Corollary 3.15** *Let $F : A \to A$ be an injective function. Then on each equivalence class $F$ acts either like a shift on $\mathbf{Z}_n$ for some $n \geq 1$ (a periodic cycle) or a shift on $\mathbf{Z}$ or a right shift on $\mathbf{N}$.*

The above theorem shows exactly how an injection $F$ can fail to be a bijection. A point $p$ is not in the range of $F$ if and only if it is an initial point for one of the right shifts.

Finally, the even more special case of a reversible dynamical systems given by a bijective function is worth recording. In that case there can be a cycle, but no tree can lead to the cycle. In the case of infinite orbits, the tree that leads to a point has only one branch, and it must be infinite.

**Corollary 3.16** *Let $F : A \to A$ be a bijective function. Then on each equivalence class $F$ acts either like a shift on $\mathbf{Z}_n$ for some $n \geq 1$ (a periodic cycle) or a shift on $\mathbf{Z}$.*

A final corollary of this last result is that every permutation of a finite set is a product of disjoint cycles.

The following discussion uses the concept of cardinal number. A countable infinite set has cardinal number $\omega_0$. A set that may be placed in one-to-one correspondence with an interval of real numbers has cardinal number $c$.

Example: Consider the set $[0, 1]$ of real numbers and the function $f : [0, 1] \to [0, 1]$ given by $f(x) = (1/2)x^4 + (1/2)$. This is an injection with range $[1/2, 1]$. It has two fixed points, at 1 and at some $c$ with $1/2 < c < 1$. Each starting point in $[0, 1/2)$ defines a different $\mathbf{N}$ equivalence class. The other points in $[1/2, c)$ lie on these equivalence classs. The starting points in $(c, 1)$ define $\mathbf{Z}$ equivalence classs. Each of these $\mathbf{N}$ equivalence classs and $\mathbf{Z}$ equivalence classs are countable, that is, have cardinality $\omega_0$. The number of $\mathbf{N}$ equivalence classs has the cardinality $c$ of the continuum. The number of $\mathbf{Z}$ equivalence classs also has cardinality $c$.

## 3.11 Isomorphism of dynamical systems

The above results depend implicitly on the notion of isomorphism of dynamical systems, but this notion deserves an explicit definition. The concept of isomorphism of sets is easy: An isomorphism from $A$ to $B$ is a bijection $h : A \to B$. However a dynamical system is more than a set; it is a set $A$ together with a specified function $f : A \to A$.

Let $A, f$ and $B, g$ be dynamical systems. An mapping $h$ from the first system to the second system is a function $h : A \to B$ such that $h \circ f = g \circ h$. It follows, of course, that for each $x$ in $A$ we have $h(f^k(x)) = g^k(h(x))$ for $k = 0, 1, 2, 3, \ldots$. The mapping is an isomorphism if $h$ is a bijection. Intuitively this says that $g$ acts on $B$ in the same way that $f$ acts on $A$.

# Problems

1. Show that a relation is reflexive if and only if $\forall x \, xRx$.

2. Show that a relation is symmetric if and only if $\forall x \forall y \, (xRy \Rightarrow yRx)$.

3. Here are two possible definitions of a transitive relation. This first is $\forall x \forall y \forall z \, ((xRy \wedge yRz) \Rightarrow xRz)$. The second is $\forall x \forall z \, (\exists y (xRy \wedge yRz) \Rightarrow xRz)$. Which is correct? Discuss.

4. Let $F$ be a function. Describe $F^T[a]$ (the forward orbit of $a$ under $F$).

5. Let $F$ be a function. Describe $F^{RT}[a]$ (the orbit of $a$ under $F$).

6. Let $F$ be a function. Is it possible that $F^T[a] = F^{RT}[a]$? Discuss in detail.

7. My social security number is 539681742. This defines a function defined on 123456789. It is a bijection from a nine point set to itself. What are the cycles? How many are they? How many points in each cycle?

8. Describe the structure of the equivalence classes generated by the dynamical system $f : [0, 1] \to [0, 1]$ given by $f(x) = \sqrt{1 - x^2}$.

9. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x + 1$. What are the equivalence classes, and what type are they ($\mathbf{Z}_n, \mathbf{Z}, \mathbf{N}$)? How many are there (cardinal number) of each type?

10. Recall that for a dynamical system two points are equivalent if their orbits overlap. Let $f : [0, +\infty) \to [0, +\infty)$ be defined by $f(x) = x^2$. Then $f$ is a bijection with two fixed points and lots of $\mathbf{Z}$ equivalence classes. However instead let $h : \mathbf{R} \to \mathbf{R}$ be defined by $h(x) = x^2$. Describe the equivalence classes of $h$.

11. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 2 \arctan(x)$. (Recall that the derivative of $f(x)$ is $f'(x) = 2/(1 + x^2) > 0$, so $f$ is strictly increasing.) What is the range of $f$? How many points are there in the range of $f$ (cardinal number)? What are the equivalence classes, and what type are they $(\mathbf{Z}_n, \mathbf{Z}, \mathbf{N})$? How many are there (cardinal number) of each type? Hint: It may help to use a calculator or draw a graph.

12. Let $f : A \to A$ be an injection with range $R \subset A$. Let $R'$ be a set with $R \subset R' \subset A$. Show that there is an injection $j : A \to A$ with range $R'$. Hint: Use the structure theorem for injective functions.

13. Bernstein's theorem. Let $g : A \to B$ be an injection, and let $h : B \to A$ be an injection. Prove that there is a bijection $k : A \to B$. Hint: Use the result of the previous problem.

# Chapter 4

# Functions, cardinal number

## 4.1 Functions

A *function* $f : A \to B$ with *domain* A and *target* (or *codomain*) $B$ assigns to each element $x$ of $A$ a unique element $f(x)$ of $B$.

Here is a note about terminology. The word function is commonly used in a very general sense. However a function is often called a *map* or *mapping.* Yet another term is *transformation.* The terms map, mapping, and transformation often suggest that the function is from one set to another set of the same general kind. In particular, when the set has some extra structure, such as a topological structure, then a term such as mapping may suggest that the structure is preserved. In the case of a topological structure this would mean that the mapping was continuous. In such a context the term function sometimes takes on a more special connotation, as meaning a function whose target is $\mathbb{R}$ (or possibly $\mathbb{C}$). It is safest, however, to refer to this explicitly as a real function (or complex function).

Example: Say that $\phi : X \to Y$ is a (continuous) mapping from the topological space $X$ to the topological space $Y$. Then if $f$ is a (real) function on $Y$, then the composition $f \circ \phi$ is a (real) function on $X$.

The set of values $f(x)$ for $x$ in $A$ is called the *range* of $f$ or the *image* of $A$ under $f$. In general for $S \subset A$ the set $f[S]$ of values $f(x)$ in $B$ for $x$ in $A$ is called the *image* of $S$ under $f$. On the other hand, for $T \subset B$ the set $f^{-1}[T]$ consisting of all $x$ in $A$ with $f(x)$ in $T$ is the *inverse image* of $T$ under $f$. In this context the notation $f^{-1}$ does not imply that $f$ has an inverse function; instead it refers to the inverse relation.

The function is *injective* (or one-to-one) if $f(x)$ uniquely determines $x$, and it is *surjective* (or onto) if each element of $B$ is an $f(x)$ for some $x$, that is, the range is equal to the target. The function is *bijective* if it is both injective and surjective. In that case it has an *inverse function* $f^{-1} : B \to A$.

If $f : A \to B$ and $g : B \to C$ are functions, then the *composition* $g \circ f : A \to C$ is defined by $(g \circ f)(x) = g(f(x))$ for all $x$ in $A$.

Say that $r : A \to B$ and $s : B \to A$ are functions and that $r \circ s = I_B$, the identity function on $B$. That is, say that $r(s(b)) = b$ for all $b$ in $B$. In this situation when $r$ is a *left inverse* of $s$ and $s$ is a *right inverse* of $r$, the function $r$ is called a *retraction* and the function $s$ is called a *section*.

**Theorem 4.1** *If $r$ has a right inverse, then $r$ is a surjection.*

**Theorem 4.2** *If $s$ has a left inverse, then $s$ is an injection.*

**Theorem 4.3** *Suppose $s : B \to A$ is an injection. Assume that $B \neq \emptyset$. Then there exists a function $r : A \to B$ that is a left inverse to $s$.*

Suppose $r : A \to B$ is a surjection. The *axiom of choice* says that there is a function $s$ that is a right inverse to $r$. Thus for every $b$ in $N$ there is a set of $x$ with $r(x) = b$, and since $r$ is a surjection, each such set is non-empty. The function $s$ makes a choice $s(b)$ of an element in each set.

The notion of surjection is related to the notion of equivalence relation and equivalence classes. If $f : A \to B$ is a surjection, then the elements of $B$ are in one-to-one correspondence to the equivalence classes of $A$ that are induced by $f$. On the other hand, just giving the equivalence classes does not specify the surjection.

## 4.2 Picturing functions

Each function $f : A \to B$ has a *graph* that is a subset of the product $A \times B$. It also has a *cograph* illustrated by the disjoint union $A + B$ and an arrow from each element of $A$ to the corresponding element of $B$. The term cograph is suggested by category theory: cograph is dual to graph in the same sense that disjoint union is dual to product.

Sometimes there is a function $f : I \to B$, where $I$ is an index set or parameter set that is not particularly of interest. Then the function $f$ is called a *indexed set* or *indexed family*. Sometime a term like *parameterized set* is used. Each indexed set determines a subset $S$ of $B$, the image of $I$ under $f$. It is usually this image subset $S = f[I]$ that is of principal interest, hence the term indexed set. It is common to depict the indexed set by drawing this image. On the other hand, different indexed sets may have the same image.
Example: Consider the set $B = \{p, q, r, s\}$. Index it by $I = \{1, 2, 3\}$. Send 1 to $q$ and 2 to $s$ and 3 to $q$. Then the subset $S = \{q, s\}$ is the image whose elements have been successfully indexed. However knowing $S$ does not determine the indexing.

Another situation is when there is a function $f : A \to J$, where $J$ is an label set or index set. In that case it might be natural to call $A$ with $f$ a *classified set*. The function induces a partition $\Gamma$ of $A$, but the partition does not have labels. Thus different classified sets can induce the same partition. The elements of the partition may be called contour sets. It is common to picture a such function through its contour sets.

Example: Consider the set $A = \{a, b, c, d\}$. Label the elements by colors $J = \{R, Y, B, G\}$. Send $a$ to $G$ and $b$ to $R$ and $c$ to $B$ and $d$ to $R$. The corresponding partition is $\{\{a\}, \{c\}, \{b, d\}\}$. Knowing the partition does not determine the colors of the elements.

## 4.3 Indexed sums and products

Let $A$ be a set-valued function defined on an index set $I$. Then the union of $A$ is the union of the range of $A$ and is written $\bigcup_{t \in I} A_t$. Similarly, when $I \neq \emptyset$ the intersection of $A$ is the intersection of the range of $A$ and is written $\bigcap_{t \in I} A_t$.

Let $A$ be a set-valued function defined on an index set $I$. Let $S = \bigcup_{t \in I} A_t$. The disjoint union or sum of $A$ is

$$\sum_{t \in I} A_t = \{(t, a) \in I \times S \mid a \in A_t\}. \tag{4.1}$$

For each $j \in I$ there is a natural mapping $\{a \mapsto (j, a) : A_j \to \sum_t A_t\}$. This is the injection of the $j$th summand into the disjoint union. Notice that the disjoint union may be pictured as something like the union, but with the elements labelled to show where they come from.

Similarly, there is a natural Cartesian product of $A$ given by

$$\prod_{t \in I} A_t = \{f \in S^I \mid \forall t\, f(t) \in A_t\}. \tag{4.2}$$

For each $j$ in $I$ there is a natural mapping $\{f \mapsto f(j) : \prod_t A_t \to A_j\}$. This is the projection of the product onto the $j$th factor. The Cartesian product should be thought of as a kind of rectangular box in a high dimensional space, where the dimension is the number of points in the index set $I$. The $j$th side of the box is the set $A_j$.

**Theorem 4.4** *The product of an indexed family of non-empty sets is non-empty.*

This theorem is another version of the axiom of choice. Suppose that each $A_t \neq \emptyset$. The result says that there is a function $f$ such that for each $t$ it makes an arbitrary choice of an element $f(t) \in A_t$.

Proof: Define a function $r : \sum_{t \in I} A_t \to I$ by $r((t, a)) = t$. Thus $r$ takes each point in the disjoint union and maps it to its label. The condition that each $A_t \neq \emptyset$ guarantees that $r$ is a surjection. By the axiom of choice $r$ has a right inverse $s$ with $r(s(t)) = t$ for all $t$. Thus $s$ takes each label into some point of the disjoint union corresponding to that label. Let $f(t)$ be the second component of the ordered pair $s(t)$. Then $f(t) \in A_t$. Thus $f$ takes each label to some point in the set corresponding to that label. $\square$

Say that $f$ is a function such that $f(t) \in A_t$ for each $t \in I$. Then the function may be pictured as a single point in the product space $\prod_{t \in I} A_t$. This geometric picture of a function as a single point in a space of high dimension is a powerful conceptual tool.

## 4.4 Cartesian powers

The set of all functions from $A$ to $B$ is denoted $B^A$. In the case when $A = I$ is an index set, the set $B^I$ is called a *Cartesian power*. This is the special case of Cartesian product when the indexed family of sets always has the same value $B$. This is a common construction in mathematics. For instance, $\mathbb{R}^n$ is a Cartesian power.

Write $2 = \{0, 1\}$. Each element of $2^A$ is the *indicator function* of a subset of $A$. There is a natural bijective correspondence between the $2^A$ and $P(A)$. If $\chi$ is an element of $2^A$, then $\chi^{-1}[1]$ is a subset of $A$. On the other hand, if $X$ is a subset of $A$, then the indicator function $1_X$ that is 1 on $X$ and 0 on $A \setminus X$ is an element of $2^A$. Sometimes an indicator function is called a characteristic function, but this term has other uses.

Say that $\phi$ is a map from $A$ to $B$, and $f$ is a real function on $B$. Then the real function

$$\phi^*(f) = f \circ \phi \tag{4.3}$$

is a real function on $A$, called the *pullback* of $f$. The map $\phi^*$ sends real functions on $B$ to real functions on $A$. It is the natural mapping on real functions coming from the mapping $\phi$ on points.

Consider the special case when $f = 1_S$ is an indicator function of a subset $S$ of $B$. Then we have the identity.

$$1_S \circ \phi = 1_{\phi^{-1}[S]}. \tag{4.4}$$

This helps to explain why taking the inverse image $\phi^{-1}[S]$ of a subset $S$ is an operation with such nice properties. It is a special kind of pullback.

## 4.5 Cardinality and Cantor's theorem on power sets

Say that a set $A$ is *countable* if $A$ is empty or if there is a surjection $f : \mathbb{N} \to A$.

**Theorem 4.5** *If $A$ is countable, then there is an injection from $A \to \mathbb{N}$.*

Proof: This can be proved without the axiom of choice. For each $a \in A$, define $g(a)$ to be the least element of $\mathbb{N}$ such that $f(g(a)) = a$. Then $g$ is the required injection. □

There are sets that are not countable. For instance, $P(\mathbb{N})$ is such a set. This follows from the following theorem of Cantor.

**Theorem 4.6 (Cantor)** *Let $X$ be a set. There is no surjection from $X$ to $P(X)$.*

The proof that follows is a diagonal argument. Suppose that $f : X \to P(X)$. Form an array of ordered pairs $(a, b)$ with $a, b$ in $X$. One can ask whether

$b \in f(a)$ or $b \notin f(a)$. The trick is to look at the diagonal $a = b$ and construct the set of all $a$ where $a \notin f(a)$.

Proof: Assume that $f : X \to P(X)$. Let $S = \{x \in X \mid x \notin f(x)\}$. Suppose that $S$ were in the range of $f$. Then there would be a point $a$ in $X$ with $f(a) = S$. Suppose that $a \in S$. Then $a \notin f(a)$. But this means that $a \notin S$. This is a contradiction. Thus $a \notin S$. This means $a \notin f(a)$. Hence $a \in S$. This is a contradiction. Thus $S$ is not in the range of $f$. □

One idea of Cantor was to associate to each set $A$, finite or infinite, a cardinal number $\#A$. The important thing is that if there is a bijection between two sets, then they have the same cardinal number. If there is no bijection, then the cardinal numbers are different. That is, the statement that $\#A = \#B$ means simply that there is a bijection from $A$ to $B$.

The two most important infinite cardinal numbers are $\omega_0 = \#\mathbb{N}$ and $c = \#P(\mathbb{N})$. The Cantor theorem shows that these are different cardinal numbers.

## 4.6 Bernstein's theorem for sets

If there is an injection $f : A \to B$, then it is natural to say that $\#A \leq \#B$. Thus, for example, it is easy to see that $\omega_0 \leq c$. In fact, by Cantor's theorem $\omega_0 < c$. The following theorem was proved in an earlier chapter as an exercise.

**Theorem 4.7 (Bernstein)** *If there is an injection $f : A \to B$ and there is an injection $g : B \to A$, then there is a bijection $h : A \to B$.*

It follows from Bernstein's theorem that $\#A \leq \#B$ and $\#B \leq \#A$ together imply that $\#A = \#B$. This result gives a way of calculating the cardinalities of familiar sets.

**Theorem 4.8** *The set $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ has cardinality $\omega_0$.*

Proof: It is sufficient to construct a bijection $f : \mathbb{N}^2 \to \mathbb{N}$. Let

$$f(m, n) = \frac{r(r+1)}{2} + m, \quad r = m + n. \tag{4.5}$$

The inverse function $g(s)$ is given by finding the largest value of $r \geq 0$ with $r(r+1)/2 \leq s$. Then $m = s - r(r+1)/2$ and $n = r - m$. Clearly $0 \leq m$. Since $s < (r+1)(r+2)/2$, it follows that $m < r + 1$, that is, $m \leq r$. Thus also $0 \leq n$. □

Since the values of the inverse function run along the anti-diagonals consisting of $m, n$ with $m + n = r$, the proof could be called an "anti-diagonal" argument". There is a lovely picture that makes this obvious.

**Corollary 4.9** *A countable union of countable sets is countable.*

Proof: Let $\Gamma$ be a countable collection of countable sets. Then there exists a surjection $u : \mathbb{N} \to \Gamma$. For each $S \in \Gamma$ there is a non-empty set of surjections from $\mathbb{N}$ to $S$. By the axiom of choice, there is a function that assigns to each $S$ in $\Gamma$ a surjection $v_S : \mathbb{N} \to S$. Let $w(m, n) = v_{u(m)}(n)$. Then $v$ is a surjection from $\mathbb{N}^2$ to $\bigcup \Gamma$. It is a surjection because each element $q$ of $\bigcup \Gamma$ is an element of some $S$ in $\Gamma$. There is an $m$ such that $u(m) = S$. Furthermore, there is an $n$ such that $v_S(n) = q$. It follows that $w(m, n) = q$. However once we have the surjection $w : \mathbb{N}^2 \to \bigcup \Gamma$ we also have a surjection $\mathbb{N} \to \mathbb{N}^2 \to \bigcup \Gamma$. □

**Theorem 4.10** *The set $\mathbb{Z}$ of integers has cardinality $\omega_0$.*

Proof: There is an obvious injection from $\mathbb{N}$ to $\mathbb{Z}$. On the other hand, there is also a surjection $(m, n) \mapsto m - n$ from $\mathbb{N}^2$ to $\mathbb{Z}$. There is a bijection from $\mathbb{N}$ to $\mathbb{N}^2$ and hence a surjection from $\mathbb{N}$ to $\mathbb{Z}$. Therefore there is an injection from $\mathbb{Z}$ to $\mathbb{N}$. This proves that $\#\mathbb{Z} = \omega_0$. □

**Theorem 4.11** *The set $\mathbb{Q}$ of rational numbers has cardinality $\omega_0$.*

Proof: There is an obvious injection from $\mathbb{Z}$ to $\mathbb{Q}$. On the other hand, there is also a surjection from $\mathbb{Z}^2$ to $\mathbb{Q}$ given by $(m, n) \mapsto m/n$ when $n \neq 0$ and $(m, 0) \mapsto 0$. There is a bijection from $\mathbb{Z}$ to $\mathbb{Z}^2$. (Why?) Therefore there is a surjection from $\mathbb{Z}$ to $\mathbb{Q}$. It follows that there is an injection from $\mathbb{Q}$ to $\mathbb{Z}$. (Why?) This proves that $\#\mathbb{Q} = \omega_0$. □

**Theorem 4.12** *The set $\mathbb{R}$ of real numbers has cardinality $c$.*

Proof: First we give an injection $f : \mathbb{R} \to P(\mathbb{Q})$. In fact, we let $f(x) = \{q \in \mathbb{Q} \mid q \leq x\}$. This maps each real number $x$ to a set of rational numbers. If $x < y$ are distinct real numbers, then there is a rational number $r$ with $x < r < y$. This is enough to establish that $f$ is an injection. From this it follows that there is an injection from $\mathbb{R}$ to $P(\mathbb{N})$.

Recall that there is a natural bijection between $P(\mathbb{N})$ (all sets of natural numbers) and $2^{\mathbb{N}}$ (all sequences of zeros and ones). For the other direction, we give an injection $g : 2^{\mathbb{N}} \to \mathbb{R}$. Let

$$g(s) = \sum_{n=0}^{\infty} \frac{2s_n}{3^{n+1}}. \tag{4.6}$$

This maps $2^{\mathbb{N}}$ as an injection with range equal to the Cantor middle third set. This completes the proof that $\#\mathbb{R} = c$. □

**Theorem 4.13** *The set $\mathbb{R}^{\mathbb{N}}$ of infinite sequences of real numbers has cardinality $c$.*

Proof: Map $\mathbb{R}^{\mathbb{N}}$ to $(2^{\mathbb{N}})^{\mathbb{N}}$ to $2^{\mathbb{N} \times \mathbb{N}}$ to $2^{\mathbb{N}}$. □

# Problems

1. What is the cardinality of the set $\mathbb{N}^{\mathbb{N}}$ of all infinite sequences of natural numbers? Prove that your answer is correct.

2. What is the cardinality of the set of all finite sequences of natural numbers? Prove that your answer is correct.

3. What is the cardinality of the set of all infinite sequences of rational numbers? Justify your answer.

4. Let $C(\mathbf{R})$ be the set of continuous real functions on $\mathbf{R}$. What is the cardinality of this set? Justify your answer.

5. Define the function $g : 2^{\mathbb{N}} \to \mathbb{R}$ by

$$g(s) = \sum_{n=0}^{\infty} \frac{2s_n}{3^{n+1}}. \tag{4.7}$$

   Prove that it is an injection.

6. Define the function $g : 2^{\mathbb{N}} \to \mathbb{R}$ by

$$g(s) = \sum_{n=0}^{\infty} \frac{s_n}{2^{n+1}}. \tag{4.8}$$

   What is its range? Is it an injection?

7. Let $A$ be a set and let $f : A \to A$ be a function. Then $f$ is a relation on $A$ that generates an equivalence relation. Can there be uncountably many equivalence classes? Explain. Can there be a single equivalence class that is uncountable? Explain. What is the situation if the function is an injection? How about if it is a surjection?

8. The notion of product space comes up in elementary algebra in a natural way. Let $I$ be a finite index set and $t \mapsto A_t$ be a family of finite sets indexed by $I$. Let $S = \bigcup_t A_t$ and $z : S \to \mathbb{R}$. The claim is that

$$\prod_{t \in I} \sum_{a \in F_t} z(a) = \sum_{f \in \prod_t F_t} \prod_{t \in I} z(f(t)). \tag{4.9}$$

   The right hand side is a sum over the product space. What is this identity; what is its role in algebra? Note: The identity in this general form is highly useful in combinatorics.

# Part II

# Order and Structure

# Chapter 5

# Ordered sets and order completeness

## 5.1 Ordered sets

The main topic of this chapter is ordered sets and order completeness. A general reference for this topic is the book of Schröder [18]. Ordered sets may also be considered in the setting of category theory; this approach is explained in an encyclopedia volume contribution by Wood [22].

The motivating example is the example of the set $W$ of rational numbers $r$ such that $0 \leq r \leq 1$. Consider the subset $S$ of rational numbers $r$ that also satisfy $r^2 < 1/2$. The upper bounds of $S$ consist of rational numbers $s$ that also satisfy $s^2 > 1/2$. (There is no rational number whose square is $1/2$.) There is no least upper bound of $S$. Contrast this with the example of the set $L$ of real numbers $x$ such that $0 \leq x \leq 1$. Consider the subset $T$ of real numbers $x$ that also satisfy $x^2 < 1/2$. The upper bounds of $T$ consists of real numbers $y$ that also satisfy $y^2 \geq 1/2$. The number $\sqrt{2}$ is the least upper bound of $T$. So know whether you have an upper bound of $T$ is equivalent to knowing whether you have an upper bound of $\sqrt{2}$. As far as upper bounds are concerned, the set $T$ is represented by a single number.

Completeness is equivalent to the existence of least upper bounds. This is the property that says that there are no missing points in the ordered set. The theory applies to many other ordered sets other than the rational and real number systems. So it is worth developing in some generality.

An *pre-ordered set* is a set $W$ and a binary relation $\leq$ that is a subset of $W \times W$. The pre-order relation $\leq$ must satisfy the first two of the following properties:

1. $\forall p\, p \leq p$ (reflexivity)
2. $\forall p \forall q \forall r((p \leq q \land q \leq r) \Rightarrow p \leq r)$ (transitivity)
3. $\forall p \forall q\, ((p \leq q \land q \leq p) \Rightarrow p = q)$. (antisymmetry)

If it also satisfies the third property, then it is an *ordered set*. An ordered set is often called a *partially ordered set* or a *poset*. In an ordered set we write $p < q$ if $p \leq q$ and $p \neq q$. Once we have one ordered set, we have many related order sets, since each subset of an ordered set is an ordered set in a natural way.

In an ordered set we say that $p, q$ are *comparable* if $p \leq q$ or $q \leq p$. An ordered set is *linearly ordered* (or *totally ordered*) if each two points are comparable. (Sometime a linearly ordered set is also called a *chain*.)

Some standard examples of linearly ordered sets are obtained by looking at the ordering of number systems. Thus we shall denote by **N** a set that is ordered in the same way as $\mathbb{N}$ or $\mathbb{N}_+$. Thus it has a discrete linear order with a least element but no greatest element. Similarly, **Z** is a set ordered the same way as the integers. It has a discrete linear order but without either greatest or least element. The set $Qb$ is ordered like the rationals. It is a countable densely ordered set with no greatest or least element. Finally, **R** is a set ordered like the reals. It is an uncountable densely ordered set with no greatest or least element.

Examples:

1. The ordered sets **N**, **Z**, **Q**, and **R** are linearly ordered sets.
2. Let $I$ be a set and let $W$ be an ordered set. Then $W^I$ with the pointwise ordering is an ordered set.
3. In particular, $\mathbf{R}^I$, the set of all real functions on $I$, is an ordered set.
4. In particular, $\mathbf{R}^n$ is an ordered set.
5. If $X$ is a set, the power set $P(X)$ with the subset relation is an ordered set.
6. Since $2 = \{0, 1\}$ is an ordered set, the set $2^X$ with pointwise ordering is an ordered set. (This is the previous example in a different form.)

## 5.2 Positivity

This is a good place to record certain conventions for real numbers and real functions. We refer to a real number $x \geq 0$ as *positive*, and a number $x > 0$ as *strictly positive*. A sequence $s$ of real numbers is *increasing* if $m \leq n$ implies $s_m \leq s_n$, while it is *strictly increasing* if $m < n$ implies $s_m < s_n$. Note that many authors prefer the terminology non-negative or non-decreasing for what is here called positive or increasing. In the following we shall often write $s_n \uparrow$ to indicate that $s_n$ is increasing in our sense.

The terminology for real functions is more complicated. A function with $f(x) \geq 0$ for all $x$ is called *positive* (more specifically, pointwise positive), and we write $f \geq 0$. Correspondingly, a function $f$ with $f \geq 0$ that is not the zero function is called *positive non-zero*. While it is consistent with the conventions for ordered sets to write $f > 0$, this may risk confusion. Sometimes a term like positive semi-definite is used. In other contexts, one needs another ordering on functions. Thus the condition that either $f$ is the zero function or $f(x) > 0$ for all $x$ might be denoted $f \geq\geq 0$, though this is far from being a standard notation. The corresponding condition that $f(x) > 0$ for all $x$ is called *pointwise strictly positive*, and a suitable notation might be $f >> 0$. An alternative is to say that $f > 0$ *pointwise* or $f > 0$ *everywhere*. Sometimes a term like positive definite is used.

The main use of the term positive definite is in connection with quadratic forms. A quadratic form is always zero on the zero vector, so it is reasonable to restrict attention to non-zero vectors. Then according to the writer positive semi-definite can mean positive or positive non-zero, while positive definite would ordinarily mean pointwise strictly positive. However some authors use the word positive definite in the least restrictive sense, that is, to indicate merely that the quadratic form is positive. A reader must remain alert to the definition in use on a particular occasion.

A related notion that will be important in the following is the pointwise ordering of functions. We write $f \leq g$ to mean that for all $x$ there is an inequality $f(x) \leq g(x)$. Similarly, we write $f_n \uparrow$ to indicate an increasing sequence of functions, that is, $m \leq n$ implies $f_m \leq f_n$. Also, $f_n \uparrow f$ means that $f_n \uparrow$ and $f_n$ converges to $f$ pointwise.

## 5.3 Greatest and least; maximal and minimal

Let $W$ be an ordered set, and let $S$ be a subset of $W$. We write $p \leq S$ to mean $\forall q\, (q \in S \Rightarrow p \leq q)$. In this case we say that $p$ is a *lower bound* for $S$. Similarly, $S \leq q$ means $\forall p\, (p \in S \Rightarrow p \leq q)$. Then $q$ is an *upper bound* for $S$.

We write $\uparrow S$ for the set of all upper bounds for $S$. Similarly, we write $\downarrow S$ for the set of all lower bounds for $S$. If $S = \{r\}$ consists of just one point we write the set of upper bounds for $r$ as $\uparrow r$ and the set of lower bounds for $r$ as $\downarrow r$.

An element $p$ of $S$ is the *least* element of $S$ if $p \in S$ and $p \leq S$. Equivalently, $p \in S$ and $S \subset \uparrow p$. As a set theory identity $\downarrow S \cap S = \{p\}$. An element $q$ of $S$ is the *greatest* element of $S$ if $q \in S$ and $S \leq q$. Equivalently, $q \in S$ and $S \subset \downarrow q$. As a set theory identity $\uparrow S \cap S = \{q\}$.

An element $p$ of $S$ is a *minimal* element of $S$ if $\downarrow p \cap S = \{p\}$. An element $q$ of $S$ is a *maximal* element of $S$ if $\uparrow q \cap S = \{q\}$.

**Theorem 5.1** *If $p$ is the least element of $S$, then $p$ is a minimal element of $S$. If $q$ is the greatest element of $S$, then $a$ is a maximal element of $S$.*

In a linearly ordered set a minimal element is a least element and a maximal element is a greatest element.

## 5.4 Supremum and infimum; order completeness

A point $p$ is the *infimum* or *greatest lower bound* of $S$ if $\downarrow S = \downarrow p$. The infimum of $S$ is denoted $\inf S$ or $\bigwedge S$. A point $q$ is the *supremum* or *least upper bound* of $S$ if $\uparrow S = \uparrow q$. The supremum of $S$ is denoted $\sup S$ or $\bigvee S$. The reader should check that $p = \inf S$ if and only if $p$ is the greatest element of $\downarrow S$. Thus $p \in \downarrow S$ and $\downarrow S \leq p$. Similarly, $q = \sup S$ if and only if $q$ is the least element of $\uparrow S$. Thus $q \in \uparrow S$ and $q \leq \uparrow S$.

An ordered set $L$ is a *lattice* if every pair of points $p, q$ has an infimum $p \wedge q$ and a supremum $p \vee q$. An ordered set $L$ is a *complete lattice* if every subset $S$ of $L$ has an infimum $\bigwedge S$ and a supremum $\bigvee S$. The most important example of a linearly ordered complete lattice is the closed interval $[-\infty, +\infty]$ consisting of all extended real numbers. An example that is not linearly ordered is the set $P(X)$ of all subsets of a set $X$. In this case the infimum is the intersection and the supremum is the union.

Examples:

1. If $[a, b] \subset [-\infty, +\infty]$ is a closed interval, then $[a, b]$ is a complete lattice.
2. Let $I$ be a set and let $W$ be a complete lattice. Then $W^I$ with the pointwise ordering is a complete lattice.
3. In particular, $[a, b]^I$, the set of all extended real functions on $I$ with values in the closed interval $[a, b]$ is an complete lattice.
4. In particular, $[a, b]^n$ is a complete lattice.
5. If $X$ is a set, the power set $P(X)$ with the subset relation is a complete lattice.
6. Since $2 = \{0, 1\}$ is a complete lattice, the set $2^X$ with pointwise ordering is a complete lattice. (This is the previous example in a different form.)

## 5.5 Sequences in a complete lattice

In general a function from an ordered set to another ordered set s said to be *increasing* (or *order preserving*) if it preserves the order relation. Thus one requires that $x \leq y$ implies $f(x) \leq f(y)$. The function is *strictly increasing* if $x < y$ implies $f(x) < f(y)$.

Two ordered sets are said to be isomorphic if there is an increasing bijection from one to the other whose inverse function is also an increasing bijection. Such an isomorphism is automatically strictly increasing.

Similarly, a function is *decreasing* if it reverses the order. There is a corresponding definition of *strictly decreasing*. Sometimes it is said to be *monotone* if it is increasing or if it is decreasing.

Each of these definitions applies in particular to an ordered *sequence*, that is, a function from $\mathbf{N}$ to another ordered set.

Let $r : \mathbf{N} \to L$ be a sequence of points in a complete lattice $L$. Let $s_n = \sup_{k \geq n} r_k$. Then the decreasing sequence $s_n$ itself has an infimum. Thus there is an element

$$\limsup_{k\to\infty} r_k = \inf_n \sup_{k\geq n} r_k. \tag{5.1}$$

Similarly, the increasing sequence $s_n = \inf_{k\geq n} r_k$ has a supremum, and there is always an element

$$\liminf_{k\to\infty} r_k = \sup_n \inf_{k\geq n} r_k. \tag{5.2}$$

It is not hard to see that $\liminf_{k\to\infty} r_k \leq \limsup_{k\to\infty} r_k$.

The application of this construction to the extended real number system is discussed in a later section. However here is another situation where it is important. This situation is quite common in probability. Let $\Omega$ be a set, and let $P(\Omega)$ be the set of all subsets. Now sup and inf are union and intersection. Let $A : \mathbf{N} \to P(\Omega)$ be a sequence of subsets. Then $\liminf_{k\to\infty} A_k$ and $\limsup_{k\to\infty} A_k$ are subsets of $\Omega$, with the first a subset of the second. The interpretation of the first one is that a point $\omega \in \liminf_{k\to\infty} A_k$ if and only if $\omega$ is eventually in the sets $A_k$ as $k$ goes to infinity. The interpretation of the second one is $\omega$ is in $\limsup_{k\to\infty} A_k$ if and only if $\omega$ is in $A_k$ infinitely often as $k$ goes to infinity.

## 5.6 Order completion

Consider an ordered set. For each subset $S$ define its downward closure as $\downarrow\uparrow S$. These are the points that are below every upper bound for $S$. Thus $S \subset \downarrow\uparrow S$, that is, $S$ is a subset of its downward closure. A subset $A$ is a lower Dedekind cut if it is its own downward closure: $A = \downarrow\uparrow A$. This characterizes a lower Dedekind cut $A$ by the property that if a point is below every upper bound for $A$, then it is in $A$.

**Lemma 5.2** *For each subset $S$ the subset $\downarrow S$ is a lower Dedekind cut. In fact $\downarrow\uparrow\downarrow S = \downarrow S$.*

Proof: Since for all sets $T$ we have $T \subset \downarrow\uparrow T$, it follows by taking $T = \downarrow S$ that $\downarrow S \subset \downarrow\uparrow\downarrow S$. Since for all sets $S \subset T$ we have $\downarrow T \subset \downarrow S$, we can take $T = \uparrow\downarrow S$ and get $\downarrow\uparrow\downarrow S \subset \downarrow S$. □

**Theorem 5.3** *If $L$ is an ordered set in which each subset has a supremum, then $L$ is a complete lattice.*

Proof: Let $S$ be a subset of $L$. Then $\downarrow S$ is another subset of $L$. Let $r$ be the supremum of $\downarrow S$. This says that $\uparrow\downarrow S = \uparrow r$. It follows that $\downarrow\uparrow\downarrow S = \downarrow\uparrow r$. This is equivalent to $\downarrow S = \downarrow r$. Thus $r$ is the infimum of $S$. □

**Theorem 5.4** *An ordered set $L$ is a complete lattice if and only if for each lower Dedekind cut $A$ there exists a point $p$ with $A = \downarrow p$.*

Proof: Suppose $L$ is complete. Let $A$ be a lower Dedekind cut and $p$ be the infimum of $\uparrow A$. Then $\downarrow\uparrow A = \downarrow p$. Thus $A = \downarrow p$.

On the other hand, suppose that for every lower Dedekind cut $A$ there exists a point $p$ with $A = \downarrow p$. Let $S$ be a subset. Then $\downarrow S$ is a lower Dedekind cut. It follows that $\downarrow S = \downarrow p$. Therefore $p$ is the infimum of $S$. □

The above theorem might justify the following terminology. Call a lower Dedekind cut a virtual point. Then the theorem says that a lattice is complete if and only if every virtual point is given by a point. This is the sense in which order completeness says that there are no missing points.

**Theorem 5.5** *Let $W$ be an ordered set. Let $L$ be the ordered set of all subsets of $W$ that are lower Dedekind cuts. The ordering is set inclusion. Then $L$ is a complete lattice. Furthermore, the map $p \mapsto \downarrow p$ is an injection from $W$ to $L$ that preserves the order relation.*

Proof: To show that $L$ is a complete lattice, it is sufficient to show that every subset $\Gamma$ of $L$ has a supremum. This is not so hard: the supremum is the downward closure of $\bigcup \Gamma$. To see this, we must show that for every lower Dedekind cut $B$ we have $\downarrow\uparrow \bigcup \Gamma \subset B$ if and only if for every $A$ in $\Gamma$ we have $A \subset B$. The only if part is obvious from the fact that each $A \subset \bigcup \Gamma \subset \downarrow\uparrow \bigcup \Gamma$. For the if part, suppose that $A \subset B$ for all $A$ in $\Gamma$. Then $\bigcup A \subset B$. It follows that $\downarrow\uparrow \bigcup A \subset \downarrow\uparrow B = B$. The properties of the injection are easy to verify. □

Examples:

1. Here is a simple example of an ordered set that is not a lattice. Let $W$ be an ordered set with four points. There are elements $b, c$ each below each of $x, y$. Then $W$ is not complete. The reason is that if $S = \{b, c\}$, then $\downarrow S = \emptyset$ and $\uparrow S = \{x, y\}$.

2. Here is an example of a completion of an ordered set. Take the previous example. The Dedekind lower cuts are $A = \emptyset$, $B = \{b\}$, $C = \{c\}$, $M = \{b, c\}$, $X = \{b, c, x\}$, $Y = \{b, c, y\}$, $Z = \{b, c, x, y\}$. So the completion $L$ consists of seven points $A, B, C, M, X, Y, Z$. This lattice is complete. For example, the set $\{B, C\}$ has infimum $A$ and supremum $M$.

## 5.7 The Knaster-Tarski fixed point theorem

**Theorem 5.6 (Knaster-Tarski)** *Let $L$ be a complete lattice and $f : L \to L$ be an increasing function. Then $f$ has a fixed point $a$ with $f(a) = a$.*

Proof: Let $S = \{x \mid f(x) \leq x\}$. Let $a = \inf S$. Since $a$ is a lower bound for $S$, it follows that $a \leq x$ for all $x$ in $S$. Since $f$ is increasing, it follows that

$f(a) \leq f(x) \leq x$ for all $x$ in $S$. It follows that $f(a)$ is a lower bound for $S$. However $a$ is the greatest lower bound for $S$. Therefore $f(a) \leq a$.

Next, since $f$ is increasing, $f(f(a)) \leq f(a)$. This says that $f(a)$ is in $S$. Since $a$ is a lower bound for $S$, it follows that $a \leq f(a)$. □

## 5.8 The extended real number system

The extended real number system $[-\infty, +\infty]$ is a complete lattice. In fact, one way to construct the extended real number system is to define it as the order completion of the ordered set $\mathbf{Q}$ of rational numbers. That is, the definition of the extended real number system is as the set of all lower Dedekind cuts of rational numbers. (Note that in many treatments Dedekind cuts are defined in a slightly different way, so that they never have a greatest element. The definition used here seems most natural in the case of general lattices.)

The extended real number system is a linearly ordered set. It follows that the supremum of a set $S \subset [-\infty, +\infty]$ is the number $p$ such that $S \leq p$ and for all $a < p$ there is an element $q$ of $S$ with $a < q$. There is a similar characterization of infimum.

Let $s : \mathbf{N} \to [-\infty, +\infty]$ be a sequence of extended real numbers. Then $s$ is said to be increasing if $m \leq n$ implies $s_m \leq s_n$. For an increasing sequence the limit exists and is equal to the supremum. Similarly, for a decreasing sequence the limit exists and is equal to the infimum.

Now consider an arbitrary sequence $r : \mathbf{N} \to [-\infty, \infty]$. Then $\limsup_{k\to\infty} r_k$ and $\liminf_{k\to\infty} r_k$ are defined.

**Theorem 5.7** *If* $\liminf_{k\to\infty} r_k = \limsup_{k\to\infty} r_k = a$, *then* $\lim_{k\to\infty} r_k = a$.

**Theorem 5.8** *If* $r : \mathbf{N} \to \mathbb{R}$ *is a Cauchy sequence, then* $\liminf_{k\to\infty} r_k = \limsup_{k\to\infty} r_k = a$, *where* $a$ *is in* $\mathbb{R}$. *Hence in this case* $\lim_{k\to\infty} r_k = a$. *Every Cauchy sequence of real numbers converges to a real number.*

This result shows that the order completeness of $[-\infty, +\infty]$ implies the metric completeness of $\mathbb{R}$.

## 5.9 Supplement: The Riemann integral

The Riemann integral illustrates notions of order. Let $X$ be a set. Let $L$ be a vector lattice of real functions on $X$. That is, $L$ is a vector space of functions that is also a lattice of functions under the pointwise order.

An example to keep in mind is when $X = \mathbb{R}$ and $L$ consists of step functions. These are functions that are finite linear combinations of indicator functions of intervals $(-a, b]$, where $a$ and $b$ are each real numbers. Notice that each such function is bounded and vanishes outside of a bounded set.

Suppose that $\mu$ is a linear order-preserving function from $L$ to $\mathbb{R}$. For example, we could define $\mu$ on indicator functions $1_{(a,b]}$ by $\mu(1_{(a,b]} = b - a$. This is of course just the length of the interval. This function is extended by linearity to the step functions. So if $f$ is a step function, $\mu(f)$ is the usual sum used as a preliminary step in the definition of an integral.

Here is an abstract version of one of the standard constructions of the Riemann integral. Let $g$ be a real function on $X$. Define the *upper integral*

$$\mu^*(g) = \inf\{\mu(h) \mid h \in L, g \leq h\}. \tag{5.3}$$

Similarly, define the *lower integral*

$$\mu_*(g) = \sup\{\mu(f) \mid f \in L, f \leq g\}. \tag{5.4}$$

The upper integral is order preserving and subadditive: $\mu^*(g_1 + g_2) \leq \mu^*(g_1) + \mu^*(g_2)$. This is because if $g_1 \leq h_1$ and $g_2 \leq h_2$, with $h_1, h_2$ both in $L$, then $g_1 + g_2 \leq h_1 + h_2$ with $h_1 + h_2$ in $L$. So $\mu^*(g_1 + g_2) \leq \mu(h_1 + h_2) = \mu(h_1) + \mu(h_2)$. The subadditivity is established taking the infimum on the right hand side.

Similarly, the lower integral is order preserving and superadditive: $\mu_*(g_1 + g_2) \geq \mu_*(g_1) + \mu_*(g_2)$. Furthermore, $\mu_*(g) \leq \mu^*(g)$ for all $g$.

Define $\mathcal{R}^1(X, \mu)$ to be the set of all $g : X \to \mathbb{R}$ such that both $\mu_*(g)$ and $\mu^*(g)$ are real, and

$$\mu_*(g) = \mu^*(g). \tag{5.5}$$

Let their common value be denoted $\tilde{\mu}(g)$. This $\tilde{\mu}$ is the *Riemann integral* on the space $\mathcal{R}^1 = \mathcal{R}^1(X, \mu)$ of $\mu$ absolutely Riemann integrable functions.

Alternatively, a function $g$ is in $\mathcal{R}^1$ if for every $\epsilon > 0$ there is a function $f$ in $L$ and a function $h$ in $L$ such that $f \leq g \leq h$, $\mu(f)$ and $\mu(h)$ are finite, and $\mu(h) - \mu(f) < \epsilon$.

It is evident that the Riemann integral is order preserving, but the fact that it is linear is less obvious. However this is true. In fact, since it is both subadditive and superadditive, it must be additive.

It may be shown that every continuous real function that vanishes outside of a bounded subset is Riemann integrable. However there are also discontinuous functions that have a Riemann integral.

A somewhat more general integral, the Riemann-Stieltjes integral, may be defined by starting with a given increasing right-continuous function $F : \mathbb{R} \to \mathbb{R}$. Interpret $F(b) - F(a) \geq 0$ as the mass in the interval $(a, b]$. Then define $\mu$ on indicator functions $1_{(a,b]}$ by $\mu(1_{(a,b]} = F(b) - F(a)$. If $g$ is a real function on $\mathbb{R}$, then $g(x)$ may be interpreted as the economic value of something found at $x$. Thus if $g$ is Riemann-Stieltjes integrable, then $\mu(g)$ is the total economic value corresponding to all the mass. truction of the Lebesgue integral.

Note: Some authors extend the definition of Riemann integral to certain functions that are not absolutely integrable, but such integrals require special consideration and are not considered here. The special consideration comes from the fact that an integral that is not absolutely convergent may be rearranged to have an arbitrary value. This has nothing to do with the distinction between Riemann integral and Lebesgue integral. Conditionally convergent sums and integrals are inherently treacherous in all cases.

## 5.10 Supplement: The Bourbaki fixed point theorem

In the appendices to this chapter it is shown that the axiom of choice implies Zorn's lemma. It is quite easy to show that Zorn's lemma implies the axiom of choice.

Consider a non-empty ordered set. Suppose that every non-empty linearly ordered subset has an upper bound. Zorn's lemma is the assertion that the set must have a maximal element.

In a sense, Zorn's lemma is an obvious result. Start at some element of the ordered set. Take a strictly larger element, then another, then another, and so on. Of course it may be impossible to go on, in which case one already has a maximal element. Otherwise one can go through an infinite sequence of elements. These are linearly ordered, so there is an upper bound. Take a strictly larger element, then another, then another, and so on. Again this may generate a continuation of the linearly ordered subset, so again there is an upper bound. Continue in this way infinitely many times, if necessary. Then there is again an upper bound. This process is continued as many times as necessary. Eventually one runs out of set. Either one has reached an element from a previous element and there is not a larger element after that. In that case the element that was reached is maximal. Or one runs at some stage through an infinite sequence, and this has an upper bound, and there is nothing larger than this upper bound. In this case the upper bound is maximal.

Notice that this argument involves an incredible number of arbitrary choices. But the basic idea is simple: construct a generalized orbit that is linearly ordered. Keep the construction going until a maximal element is reached, either as the result of a previous point in the orbit, or as the result of an previous sequence in the orbit.

The key lemma that makes this rigorous is the Bourbaki fixed point theorem. This is a theorem about a dynamical system defined by a function that sends points upward in an ordered set. (The orbits of this dynamical system may be thought of as increasing functions from ordinal numbers to the ordered set.) The theorem itself does not itself depend on the axiom of choice. However together with the axiom of choice it will lead to a proof of Zorn's lemma.

**Theorem 5.9 (Bourbaki)** *Let $A$ be a non-empty ordered set. Suppose that every non-empty linearly ordered subset has a supremum. Let $f : A \to A$ be a function such that for all $x$ in $A$ we have $x \leq f(x)$. Then $f$ has a fixed point.*

Proof: The function $f : A \to A$ is a dynamical system. Since $A$ is non-empty, we can choose $a$ in $A$ as a starting point. Let $B \subset A$. We say that $B$ is admissible if $a \in B$, $f[B] \subset B$, and whenever $T \subset B$ is linearly ordered, then $\sup T \in B$. Thus $f$ restricted to $B$ is itself a dynamical system.

Let $M$ be the intersection of all admissible subsets of $A$. It is not difficult to show that $M$ is itself an admissible subset and that $a$ is the least element of $M$. Thus $f$ restricted to $M$ is a dynamical system. We want to show that there is a sense in which $M$ is a kind of generalized orbit of $f$ starting at $a$. More precisely, we want to show that $M$ is linearly ordered. The rest of the proof is to establish that this is so. Then the fixed point will just be the supremum of this linearly ordered set. In other word, the system starts at $a$ and follows this generalized orbit until forced to stop.

Let $E \subset M$ be the set of points $c \in M$ such that for all $x$ in $M$, the condition $x < c$ implies $f(x) \leq c$. Such a point $c$ will be called a "choke point,", for a reason that will be soon apparent.

Let $c \in E$. Let $M_c \subset M$ be the set of points $x$ in $M$ such that $x \leq c$ or $f(c) \leq x$. These are the points that can be compared unfavorably to $c$ or favorably to $f(c)$.

First we check that $M_c$ is admissible. First, it is clear that $a$ is in $M_c$. Second, $f$ maps the set of elements $x \leq c$ in $M_c$ to $M_c$ (since $x < c$ implies $f(x) \leq c$ and $x = c$ implies $f(c) \leq f(x)$) , and $f$ maps the set of elements $x$ in $M_c$ with $f(c) \leq x$ to itself. Third, if $T \subset M_c$ is linearly ordered with supremum $b$, then either $x \leq c$ for all $x \in T$ implies $b \leq c$ (since $b$ is the least upper bound), or $f(c) \leq x$ for some $x$ in $T$ implies $f(c) \leq b$ (since $b$ is an upper bound). Thus $b$ is also in $M_c$.

So $M \subset M_c$, in fact they are equal. This works for arbitrary $c$ in $E$. The conclusion is that $c$ in $E$, $x$ in $M$ implies $x \leq c$ or $f(c) \leq x$. Thus each choke point $c$ of $M$ splits $M$ into a part unfavorable to $c$ or favorable to $f(c)$. This justifies the term "choke point."

Next we check that the set of all choke points $E$ is admissible. First, it is vacuously true that $a$ is in $E$. Second, consider an arbitrary $c$ in $E$, so that for $x$ in $M$ we have $x < c$ implies $f(x) \leq c$. Suppose that $x$ is in $M$ and $x < f(c)$. Since $M \subset M_c$, it follows that $x \leq c$ or $f(c) \leq x$. However the latter possibility is ruled out, so $x \leq c$. If $x < c$, then $f(x) \leq c \leq f(c)$, and if $x = c$ then again $f(x) \leq f(c)$. This is enough to imply that $f(x) \leq f(c)$. Thus for $x$ in $M$ we have that $x < f(c)$ implies $f(x) \leq f(c)$. This shows that $f(c)$ is in $E$. In other words, $f$ leaves $E$ invariant. Third, let $T$ be a linearly ordered subset of $E$ with supremum $b$. Suppose $x$ is in $M$ with $x < b$. Since for all $c$ in $E$ we have $M \subset M_c$, either $f(c) \leq x$ for all $c$ in $T$, or $x \leq c$ for some $c$ in $T$. In the first case $x$ is an upper bound for $T$, and so the least upper bound $b \leq x$. This is a contradiction. In the remaining second case $x \leq c$ for some $c$ in $T$. If $x < c$, then $f(x) \leq c \leq b$, otherwise $x = c$ is in $E$ and since $b \leq x$ is ruled out, again we have $f(x) \leq b$. Thus for all $x$ in $M$ we have that $x < b$ implies $f(x) \leq b$. Hence $b$ is in $E$.

So $M \subset E$, in fact they are equal. Every point of $M$ is an choke point.

Now we are done. Suppose that $x$ and $y$ are in $M$. Since $M \subset E$, it follows that $x$ is in $E$. Since $M \subset M_x$, it follows that $y$ is in $M_x$. Thus $y \leq x$ or $f(x) \leq y$. Hence $y \leq x$ or $x \leq y$. This proves that $M$ is linearly ordered. Therefore it has a supremum $b$. However $b \leq f(b) \leq b$. So $b$ is a fixed point of $f$. □

## 5.11 Supplement: Zorn's lemma

The following is an optional topic. It is the proof that the axiom of choice applies Zorn's lemma.

**Theorem 5.10 (Hausdorff maximal principle)** *Every ordered set has a maximal linearly ordered subset.*

Proof: Let $W$ be the ordered set. Consider the set $A$ of all linearly ordered subsets of $W$. Suppose that $T$ is a non-empty linearly ordered subset of $A$. Then $\bigcup T$ is a linearly ordered subset of $A$, and it is the supremum of $T$. Suppose there is no maximal element of $A$. Then for each $x$ in $A$ the set of $U_x$ of linearly ordered subsets $y$ of $W$ with $x \subset y$ and $x \neq y$ is non-empty. By the axiom of choice there is a function $f : A \to A$ such that $f(x) \in U_x$. This $f$ does not have a fixed point. This contradicts the Bourbaki fixed point theorem. □

**Theorem 5.11 (Zorn's lemma)** *Consider a non-empty ordered set such that every non-empty linearly ordered subset has an upper bound. Then the set has a maximal element.*

Proof: Let $W$ be the ordered set. By the Hausdorff maximal principle there is a maximal linearly ordered subset $X$. Since $W$ is not empty, $X$ is not empty. Therefore there is a maximal element $m$ in $X$. Suppose there were an element $p$ with $m \neq p$ and $m < p$. Then we could adjoin $p$ to $X$ and get a strictly larger linearly ordered subset. This is a contradiction. So $m$ is maximal in $W$. □

## 5.12 Supplement: Ordinal numbers

This section is an informal supplement meant to contrast cardinal numbers with ordinal numbers. As we shall see, cardinal numbers classify sets up to isomorphism, while ordinal numbers classify well-ordered sets up to isomorphism.

A cardinal number is supposed to describe how many elements there are in a set. Two sets have the same *cardinal number* precisely when there is a bijection between the two sets. Addition of cardinal numbers corresponds to disjoint union $A+B$ of sets, while multiplication of cardinal numbers corresponds to Cartesian product $A \times B$ of sets. It may be proved using the axiom of choice that for infinite cardinal numbers $\kappa, \lambda$ we have $\kappa + \lambda = \max(\kappa, \lambda)$ and $\kappa \cdot \lambda = \max(\kappa, \lambda)$. Thus addition and multiplication are not very interesting. On the other hand, the exponential of cardinal numbers corresponds to the Cartesian power $B^A$ of sets. Cardinal exponentiation has many mysteries.

A linearly ordered set is *well-ordered* if every non-empty subset has a least element. It follows from Zorn's lemma that every non-empty set has a well-ordering. An *initial segment* of a well-ordered set $X$ is a set of the form $I_x = \{y \in X \mid y < x\}$.

**Theorem 5.12 (Transfinite induction)** *Let $X$ be a well-ordered set. Let $A$ be a subset of $X$. Suppose that for each $x$ in $A$ the condition $I_x \subset A$ implies $x \in A$. Then $A = X$.*

It may be shown [5] that given two well-ordered sets $X, Y$, either $X$ is isomorphic to $Y$, or $X$ is isomorphic to an initial segment of $Y$, or $Y$ is isomorphic to an initial segment of $X$.

The idea of *ordinal number* is that it characterizes a well-ordered set up to order isomorphism. If $\alpha, \beta$ are ordinal numbers, then $\alpha < \beta$ means that a set corresponding to $\alpha$ is an initial segment of a set corresponding to $\beta$. Thus for two ordinals, either $\alpha = \beta$, or $\alpha < \beta$, or $\beta < \alpha$.

There is also an arithmetic of ordinal numbers. This comes form corresponding operations on well-ordered sets. Suppose $A$ and $B$ are each well-ordered sets. Their disjoint sum $A+B$ can be well-ordered by taking each element of the copy of $B$ after each element of the copy of $A$. Also their Cartesian product $A \times B$ may be well-ordered by taking the $(a, b)$ pairs in the order in $(a, b) \leq (a', b')$ when $b < b'$ or when $b = b'$ and $a \leq b$. In other words, line up copies of $A$ according to the ordering of $B$.

Finally, denote the least element of $B$ by 0. Consider the space $B^{(A)}$ of all functions from $A$ to $B$ that each have the value 0 on all but finitely many points of $A$. (When $A$ is infinite this is only a small part of the Cartesian power.) Suppose $f$ and $g$ are two such functions. If $f = g$ then certainly $f \geq g$. If $f \neq g$, then there are finitely many elements $x$ of $A$ with $f(x) \neq g(x)$. Let $a$ be maximal among these. Then $f \leq g$ in $B^{(A)}$ is to hold provided that $f(a) < g(a)$ in $B$.

The ordinal numbers are supposed to classify the well-ordered sets. Thus the sum $\alpha + \beta$ is defined by the disjoint union construction, the product $\alpha \cdot \beta$ is defined by the Cartesian product construction, and the exponential $\alpha^\beta$ is defined by the function space construction. For more details see [16].

The first few ordinal numbers are $0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \omega + 3, \ldots, \omega \cdot 2, \omega \cdot 2 + 1, \omega \cdot 2 + 2, \omega \cdot 2 + 3, \ldots, \omega \cdot 2 + \omega, \omega \cdot 2 + \omega + 1, \omega \cdot 2 + \omega + 2, \ldots$. Notice that $\omega \cdot 2 = \omega + \omega$, since both represent two copies of $\omega$ lined up one after the other.

These operations are not commutative. Notice that $1+\omega = \omega$, but $\omega < \omega+1$. Also $2 \cdot \omega = \omega$, but $\omega < \omega \cdot 2$.

The examples given do not exhaust the ordinals. After $\omega \cdot 2, \ldots, \omega \cdot 3, \ldots$ and so on comes $\omega^2 = \omega \cdot \omega$. This represents countably many copies of $\omega$ lined up one after the other. Then after $\omega^2, \ldots, \omega^2 \cdot 2, \ldots, \omega^2 \cdot 3, \ldots$ comes $\omega^3$. So a typical ordinal in this range might take the form $\omega^2 \cdot 2 + \omega \cdot 7 + 4$. This represents countably many copies of $\omega$ in order, followed by the same thing, followed by seven copies of $\omega$, followed by four individual elements.

Even larger ordinals include $\omega^3 + 1, \omega^3 + 2, \ldots, \omega^4, \ldots$ and so on, up to $\omega^\omega, \omega^\omega + 1, \omega^\omega + 2, \ldots$. This is just the beginning of a long and complicated process that eventually leads to $\omega_1$, the first uncountable ordinal. Each ordinal less than this ordinal correspond to a countable well-ordered set. So while from the cardinal point of view all countable infinite sets look the same, from the ordinal point of view there is a rather complicated story.

# Problems

1. Consider the sequence of real numbers $s_n = (-1)^n \frac{n+2}{n+1}$. State the definition of $\limsup_{n\to\infty} s_n$ in terms of the concepts of supremum and infimum, and evaluate using the definition for this particular sequence.

2. Let $X = (0, +\infty)$ and for each $\epsilon > 0$ define $f_\epsilon : X \to [0, +\infty]$ by $f_\epsilon(x) = \frac{1}{\epsilon^2}(\epsilon - x) \vee 0$. Consider the complete lattice $[0, +\infty]^X$.

   (a) Find $h = \sup\{f_\epsilon \mid \epsilon > 0\}$. Hint: Maximize $f_\epsilon(x)$ for fixed $x$.

   (b) Find $\int_0^\infty f_\epsilon(x)\,dx$. Find $\int_0^\infty h(x)\,dx$.

3. Does the ordered set $\mathbf{R} \setminus \mathbf{Q}$ of irrational numbers form a boundedly complete lattice? Explain.

4. Let $L$ be a complete lattice. There is a map from the power set $P(L)$ to itself given by $S \mapsto \sup S$. There is also a map from $L$ to $P(L)$ given by $y \mapsto \downarrow y$. Show that these maps are adjoint, in the sense that $\sup S \le y \equiv S \subset \downarrow y$.

5. Show that $S \ne \emptyset$ implies $\inf S \le \sup S$.

6. Show that $\sup S \le \inf T$ implies $S \le T$ (every element of $S$ is $\le$ every element of $T$).

7. Show that $S \le T$ implies $\sup S \le \inf T$.

8. Let $L$ be a linearly ordered complete lattice. Show that $p$ is the supremum of $S$ if and only if $p$ is an upper bound for $S$ and for all $r < p$ there is an element $q$ of $S$ with $r < q$.

9. Let $L$ be a complete lattice. Suppose that $p$ is the supremum of $S$. Does it follow that for all $r < p$ there is an element $q$ of $S$ with $r < q$? Give a proof or a counterexample.

10. Let $S_n$ be the set of symmetric real $n$ by $n$ matrices. Each $A$ in $S_n$ defines a real quadratic form $x \mapsto x^T Ax : \mathbb{R}^n \to \mathbb{R}$. Here $x^T$ is the row vector that is the transpose of the column vector $x$. Since the matrix $A$ is symmetric, it is its own transpose: $A^T = A$. The order on $S_n$ is the pointwise order defined by the real quadratic forms. Show that $S_2$ is not a lattice. Hint: Let $P$ be the matrix with 1 in the upper left corner and 0 elsewhere. Let $Q$ be the matrix with 1 in the lower right corner and 0 elsewhere. Let $I = P + Q$. Show that $P \le I$ and $Q \le I$. Show that if $P \vee Q$ exists, then $P \vee Q = I$. Let $W$ be the symmetric matrix that is 4/3 on the diagonal and 2/3 off the diagonal. Show that $P \le W$ and $Q \le W$, but $I \le W$ is false.

11. Let $L = [0, 1]$ and let $f : L \to L$ be an increasing function. Can a fixed point be found by iteration? Discuss.

12. An ordered set is said to be boundedly complete if every non-empty subset that has an upper bound has a supremum (least upper bound). Prove that if an ordered set is boundedly complete, then every non-empty subset that has a lower bound has an infimum (greatest lower bound).

13. Suppose that an ordered set is boundedly complete. Show that its completion is order isomorphic to the same set with appropriate top or bottom elements adjoined (if they are missing).

14. Consider the set of natural numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$. Define an unusual order relation by taking $m <= n$ to mean that there is a $k$ in $\mathbb{N}$ with the product $m \cdot k = n$. Is this a complete lattice? If it is, give a proof. If not, describe its completion.

15. Consider the infinite binary tree. This consists of all functions $p$ such that for some $n$ in $\mathbb{N}$, $n = \{0, 1, 2, \ldots, n-1\}$ we have $p : n \to \{0, 1\}$. The ordering is extension. (In other words, the tree consists of all finite sequences of 0s and 1s; a larger point has more 0s and 1s tacked on at the end.) Is this a complete lattice? If it is, give a proof. If not, describe its completion.

16. Use the Knaster-Tarski fixed point theorem to prove that the equation $2 + \ln(x) = x$ has a solution in the interval $[2, 4]$. Where does this proof use the continuity of the logarithm function?

# Chapter 6

# Structured sets

## 6.1 Structured sets and structure maps

A *structured set* is a set together with additional information. In addition, given two such structured sets, there is a notion of *structure map* that relates the information in one structured set to an other structured set in a natural way. Such notions are dealt with in great generality in category theory [14]. However here we confine ourselves to elementary examples that may be of use in analysis.

Example: Consider a set $X$ with a distinguished subset $A$. In algebraic topology this is called a pair. A structure map from $X, A$ to $X', A'$ is a function $\phi : X \to X'$ such that $\phi[A] \subset A'$, or, what is the same, $A \subset \phi^{-1}[A']$. This is called a map of pairs.

A *structure isomorphism* between structured sets is a bijection between the underlying sets that preserves the additional information. The notion of structure map should satisfy the requirement that if $\phi : X \to X'$ is a structure map such that its inverse function $\phi^{-1}$ is also a structure map, then $\phi$ is a structure isomorphism.

Example: Consider again the example of pairs consisting of a set $X$ and a distinguished subset $A$. Then an isomorphism is a bijection $\phi : X \to X'$ such that $\phi$ restricts to a bijection from $A$ to $A'$.

## 6.2 Subset of the product space

If $X$ is a set, then we may consider its product $X^2 = X \times X$ with itself. If $X'$ is another such set, then there is a map $\phi^2 : X \times X \to X' \times X'$ given by $\phi^2(x, y) = (\phi(x), \phi(y))$.

Example: Consider a structured set consisting of a set $X$ together with a distinguished relation $R \subset X^2$. A structure map from $X, R$ to $X', R'$ is a function $\phi : X \to X'$ such that $\phi^2[R] \subset R'$. Thus if $x$ is related to $y$ by $R$, then $\phi(x)$ is related to $\phi(y)$ by $R'$.

The ordered set concept is a special case of this example. In this case the relation must satisfy the order axioms. A structure map in this case is just an order preserving map.

A dynamical system is another special case of this example. In this case the structured set is $X, f$, where $f : X \to X$. We write $(x, y) \in R$ when $y = f(x)$. Then a structure map from $(X, f)$ to $(X', f')$ is when $y = f(x)$ implies $\phi(y) = f'(\phi(x))$. This says that $\phi \circ f = f' \circ \phi$.

If $X$ is a set, then we may consider its product $X^3 = X \times X \times X$ with itself. If $X'$ is another such set, then there is a corresponding map $\phi^3 : X \times X \times X \to X' \times X' \times X'$.

Example: Consider a structured set consisting of a set $X$ together with a distinguished subset $G \subset X^3$. A structure map from $X, G$ to $X', G'$ is a function $\phi : X \to X'$ such that $\phi^3[G] \subset G'$.

This kind of example occurs in many algebraic structures. Typically $G$ is the graph of a function $f : X^2 \to X$. In this case the structured set is $X, f$, where $f : X^2 \to X$ may be thought of as some kind of multiplication. We write $(x, y, z) \in G$ when $z = f(x, y)$. Then a structure map from $(X, f)$ to $(X', f')$ is when $z = f(x, y)$ implies $\phi(z) = f'(\phi(x, y))$. This says that $\phi \circ f = f' \circ \phi^2$.

Example: Instead of taking a subset of a product space, one can take a function on the product space. In this case a structured set is of the form $X, d$, where $d : X^2 \to [0, +\infty]$ may be thought of as a distance function. A structure map from $X, d$ to $X', d'$ is a function $\phi : X \to X'$ such that $d' \circ \phi^2 \leq d$. This means that for each ordered pair $x, y$ we have $d'(\phi(x), \phi(y)) \leq d(x, y)$. In other words, the structure map is a contraction.

Notice that an isomorphism is a contraction in both directions, so it preserves distances exactly.

A metric space is of this kind. However the framework is general enough to include the example of $X, R$, where $R$ is a relation. Take $d(x, y) = 0$ for $(x, y)$ in $R$ and $d(x, y) = +\infty$ otherwise. It follows that a structure map from $X, R$ to $X', R'$ has the property that $(x, y) \in R$ implies $(\phi(x), \phi(y)) \in R'$.

## 6.3 Subset of the power set

Another class of examples come from considering a subset of the power set $P(X)$. If $\phi : X \to X'$ is a function, then $\phi^{-1} : P(X') \to P(X)$ is defined by taking inverse images under the function. This map has the wonderful property of preserving unions, intersections, and complements.

Example: Consider a structured set consisting of an set $X$ and a distinguished subset $\Gamma \subset P(X)$. A structure map from $X, \Gamma$ to $X', \Gamma'$ is a function $\phi : X \to X'$ such that $\phi^{-1}[\Gamma'] \subset \Gamma$. In other words, for each $A$ in $P(X')$ we have that $A \in \Gamma'$ implies $\phi^{-1}[A] \in \Gamma$.

Again there is a natural notion of isomorphism. In this setting an isomorphism is given by a bijection $\phi$ from $X$ to $X'$ whose action on subsets produces a bijection between $\Gamma$ and $\Gamma'$.

This example includes the usual definition of topological space and continuous map and also includes the usual definition of measurable space and measurable map.

Example: Consider a structured set consisting of an set $X$ and a distinguished subset $\mathcal{M} \subset \mathbb{R}^X$. A structure map from $X, \mathcal{M}$ to $X', \mathcal{M}'$ is a function $\phi : X \to X'$ such that for each $f : X \to \mathbb{R}$ we have that $f$ in $\mathcal{M}'$ implies $f \circ \phi \in \mathcal{M}$.

The second example gives a framework for an alternate definition of measurable space and measurable map. It is also general enough to include the previous example. Take functions with values in $\{0, 1\}$, these are indicator functions of sets. Note that $1_A \circ \phi = 1_{\phi^{-1}[A]}$. Thus a structure map from a structured set $X, \Gamma$ to a structured set $X', \Gamma'$ is a map $\phi : X \to X'$ such that $A \in \Gamma'$ implies $\phi^{-1}[A] \in \Gamma$.

## 6.4 Structured sets in analysis

As a preview of things to come, here are some important structures in analysis.

- Ordered set $X, \leq$. The relation $\leq$ satisfies the order axioms.
- Metric space $X, d$. The distance function $d$ satisfies the metric space axioms.
- Topological space $X, \mathcal{T}$. The collection $\mathcal{T}$ of subsets is closed under unions and finite intersections.
- Measurable space $X, \mathcal{F}$. The collection $\mathcal{F}$ of subsets is closed under countable unions, countable intersections, and complements.

There are also important relations between these structures. For instance, the open subsets of a metric space $X, d$ form a collection $\mathcal{T}$ such that $X, \mathcal{T}$ is a topological space. Also, the Borel subsets of a topological space $X, \mathcal{T}$ form a collection $\mathcal{F}$ such that $X, \mathcal{F}$ is a measurable space.

## Problems

1. Consider a structured set consisting of a set and an equivalence relation. Show that a structure map gives a natural way of mapping equivalence classes to equivalence classes. Show that this mapping on equivalence classes mail fail to be injective and may also fail to be surjective.
2. Consider a set labelled by the set $L$, that is, a set $X$ together with a given function $g : X \to L$. Such a set is naturally partitioned into equivalence classes. Define the structure map $\phi : X \to X'$ by requiring that $g' \circ \phi = g$. Show that a structure map gives a natural way of mapping equivalence classes to equivalence classes. Show that this mapping on equivalence classes is injective. Show that this mapping on equivalence classes may fail to be surjective.

# Part III

# Measure and Integral

# Chapter 7

# Measurable spaces

## 7.1 $\sigma$-algebras of subsets

The purpose of the first two chapters of this part is to define concepts that are fundamental in the real analysis: measurable space and measure space. The actual construction of these objects will take place in subsequent chapters. The most important point in this chapter is that there are two concepts, that of $\sigma$-algebra of subsets and of $\sigma$-algebra of real functions, but these are entirely equivalent.

Similarly, there is a concept of measurable space. It may be thought of as a set together with a $\sigma$-algebra of sets. Alternatively, it may be thought of as a set together with a $\sigma$-algebra of real functions. Again, it makes no difference. In fact, it will be convenient to keep this as a free choice and consider a measurable space as a set together with a specified $\sigma$-algebra (of either kind).

Let $X$ be a non-empty set. A *$\sigma$-algebra of subsets* of $X$ is a subcollection $\mathcal{F}$ of $P(X)$ with the following three properties:

1. If $\Gamma \subset \mathcal{F}$ is countable, then $\bigcup \Gamma \in \mathcal{F}$.

2. If $\Gamma \subset \mathcal{F}$ is countable, then $\bigcap \Gamma \in \mathcal{F}$.

3. If $A \in \mathcal{F}$, then $X \setminus A \in \mathcal{F}$.

In other words $\mathcal{F}$ must be closed under countable unions, countable intersections, and complements. The empty set $\emptyset$ and the space $X$ must also belong to the $\sigma$-algebra.

A *measurable space* $X, \mathcal{F}$ is a non-empty set $X$ together with a specified $\sigma$-algebra of subsets $\mathcal{F}$. A *measurable subset* is a subset that is in the given $\sigma$-algebra. When the particular $\sigma$-algebra under discussion is understood from context, then a measurable space $X$ is often denoted by its underlying set $X$.

Examples:

1. The first and simplest standard example is when $X$ is a countable set, and the $\sigma$-algebra consists of all subsets of $X$.

2. The second standard example is when $X = \mathbb{R}$ and the $\sigma$-algebra is the smallest $\sigma$-algebra that contains all the open intervals $(a, b)$ with $a < b$. We shall see that this example is the Borel $\sigma$-algebra. A variant of this example is when the set $\mathbb{R}$ is replaced by $[0, 1]$, and the $\sigma$-algebra is relativized to this subset.

## 7.2 Measurable maps

A map $\phi : X \to Y$ between measurable spaces is a *measurable map* if the inverse image of each measurable set is measurable.

If $\Gamma \subset P(Y)$ is an arbitrary collection of subsets of $Y$, then there is a least $\sigma$-algebra $\mathcal{F}$ of subsets of $Y$ such that $\Gamma \subset \mathcal{F}$. This is the $\sigma$-algebra of subsets generated by $\Gamma$.

For example, say that $\Gamma = \{U, V\}$, where $U \subset Y$ and $V \subset Y$. Then the $\sigma$-algebra of sets generated by $\Gamma$ can have up to 16 subsets in it.

**Theorem 7.1** *Say that $X$ and $Y$ are measurable spaces. Suppose also that $\Gamma$ generates the $\sigma$-algebra of sets $\mathcal{F}$ for $Y$. Suppose that $f : X \to Y$ and that for every set $B$ in the generating set $\Gamma$ the inverse image $f^{-1}[B]$ is a measurable set. Then $f$ is a measurable map.*

Proof: Consider the collection of all measurable subsets $B$ of $Y$ such that $\phi^{-1}[B]$ is a measurable subset of $X$. This is easily seen to be a $\sigma$-algebra. By assumption it includes every set in $\Gamma$. It follows that every measurable subset of $Y$ belongs to the collection. □

## 7.3 Metric spaces

A *metric space* is a set $M$ together with a function $d : M \times M \to [0, +\infty)$ such that for all $x, y, z$

1. $d(x, x) = 0$.

2. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality)

3. $d(x, y) = d(y, x)$ (symmetry).

4. $d(x, y) = 0$ implies $x = y$ (separatedness).

For the present purpose the example to keep in mind is $\mathbb{R}^n$ with the Euclidean distance.

In a metric space $M$ the *open ball* centered at $x$ of radius $\epsilon > 0$ is defined to be $B(x, \epsilon) = \{y \mid d(x, y) < \epsilon\}$. A subset $U$ of a metric space $M$ is *open* if $\forall x\, (x \in U \Rightarrow \exists \epsilon\, B(x, \epsilon) \subset U)$. A subset $F$ is *closed* if it is the complement of an open subset. Properties of metric spaces that are defined entirely in terms of the open and closed subsets are called *topological properties*.

We take it as known that if $X$ and $Y$ are metric spaces, then a function $f : X \to Y$ is continuous if and only if the inverse image of every open subset of $Y$ is an open subset of $X$. So continuity is a topological property.

It takes some time to get a good intuition of open and close subsets of metric spaces. However in the case of the real line $\mathbb{R}$ with the usual metric there is a particularly transparent characterization: a subset $U$ is open if and only if it is a countable union of open intervals.

Example: The *Cantor space* $X = 2^{\mathbb{N}_+}$ is the set of all infinite sequences of 0s and 1s. There is an injection $g : X \to [0, 1]$ defined by $g(x) = \sum_{n=1}^{\infty} \frac{2x_n}{3^n}$. The range of this injection is the middle third Cantor set. The Cantor space inherits its metric from this set. Thus $X$ may be thought of as a metric space with the metric $d(x, y) = |g(x) - g(y)|$.

This metric has the property that if $x_1 = y_1, \ldots, x_m = y_m$, the $d(x, y) \leq 1/3^m$. On the other hand, if $d(x, y) < 1/3^m$, then $x_1 = y_1, \ldots, x_m = y_m$. So two sequences are close in this metric if they agree on finite initial segments.

## 7.4 The Borel $\sigma$-algebra

If $X$ is a metric space, then it determines a measurable space by taking $\mathcal{F}$ as the *Borel $\sigma$-algebra of subsets*. This is the smallest $\sigma$-algebra $\mathcal{B}o_X$ that contains all the open sets of the metric space. Since it is closed under complements, it also contains all the closed sets. The subsets in this $\sigma$-algebra are said to be *Borel measurable subsets*.

Perhaps the most important example is when $X = \mathbb{R}^n$ with its usual metric. Then the Borel $\sigma$-algebra is large enough so that most subsets that one encounters in practical situations are in this $\sigma$-algebra. (However it may be shown that there are many subsets that do not belong to the Borel $\sigma$-algebra; it is just that they are somewhat complicated to construct.)

**Proposition 7.2** *Suppose $X$ is a measurable space and $Y$ is a metric space, and $Y$ has the Borel $\sigma$-algebra making it also a measurable space. Let $\phi : X \to Y$ be a map from $X$ to $Y$ such that the inverse image of every open set is measurable. Then $\phi$ is a measurable map.*

This proposition is a special case of the theorem on measurable maps. If $X$ is also a metric space with the Borel $\sigma$-algebra, then one important consequence is that every continuous map is measurable. Notice, however, that it is possible to have a map $\phi : X \to Y$ that is measurable or even continuous but such that the image $\phi[X]$ is not a Borel subset of $Y$.

Examples:

1. Recall that the Cantor space $X = 2^{\mathbb{N}_+}$ may be considered as a metric space. Let $0 \leq n$ and let $z$ be a sequence of $n$ zeros and ones. Let $F_{n;z}$

be the set of all sequences $x$ in $2^{\mathbb{N}_+}$ that agree with $z$ in the first $n$ places. This consists of all coin tosses that have a particular pattern of successes and failures in the first $n$ trials, without regard to what happens in later trials. Let $\mathcal{F}$ be the smallest $\sigma$-algebra such that each set $F_{n;z}$ is in $\mathcal{F}$. It may be shown that this is the Borel $\sigma$-algebra. The Cantor space, when regarded as a measurable space with this $\sigma$-algebra rather than as a metric space, could also be called the *coin-tossing space*.

2. Let $Y = [0, 1]$ be the closed unit interval. For this example take the Borel $\sigma$-algebra $\mathcal{B}o$.

   Let $0 \leq n$ and let $z$ be a sequence of $n$ zeros and ones. Let $t_{n;z} = \sum_{k=1}^{n} z_k/2^k$. Define the closed interval $I_{n;z} = [t_{n;z}, t_{n;z} + 1/2^n]$. This clearly belongs to the Borel $\sigma$-algebra.

3. There is a relation between the above examples. Let $\phi : 2^{\mathbb{N}_+} \to [0, 1]$ be defined by $\phi(x) = \sum_{i=1}^{\infty} x_i/2^n$. Then $\phi$ is a measurable map. In fact, it is even continuous. This is because if $d(x, y) < 1/3^m$, then $x$ and $y$ agree in their first $m$ places, and so $|\phi(x) - \phi(y)| \leq 1/2^m$.

   One useful property of the measurable map $\phi$ is that it gives a relation between the subsets $F_{n;z}$ of the coin-tossing space and the intervals $I_{n;z}$. In fact, the relation is that $\phi^{-1}[I_{n;z}] = F_{n;z}$.

Example: There are situations when it is useful to consider several $\sigma$-algebras at once. For instance, let $X = \mathbb{R}^n$. For each $k = 1, \ldots, n$ let $\sigma(x_1, \ldots, x_k)$ be the $\sigma$-algebra of subsets of $\mathbb{R}^n$ consisting of all sets of the form $A = \{(x, y) \mid x \in \mathbb{R}^k, y \in \mathbb{R}^{n-k}, x \in B\}$ for some Borel subset $B$ of $\mathbb{R}^k$. In other words, the definition of the set depends only on the first $k$ coordinates.

In probability one thinks of $\{1, \ldots, n\}$ as $n$ time steps, and $\sigma(x_1, \ldots, x_k)$ corresponds to all questions that can be answered with the information available at time $k$. That is, as the experimental unfolds, initially there is no information, then the first coordinate is revealed, then also the second, and so on.

## 7.5 Measurable functions

Let $X$ be a measurable space. Let $f : X \to \mathbb{R}$ be a real function. Then $f$ is said to be a *measurable function* if the inverse image of every Borel subset of $\mathbb{R}$ is a measurable subset of $X$. That, is, a measurable function is a measurable map where the target is $\mathbb{R}$ with its Borel $\sigma$-algebra.

**Lemma 7.3** *A real function $f$ is a measurable function if for every real $a$ the set of points where $f > a$ is a measurable subset.*

Proof: Since the subset of $X$ where $f \geq a$ is the intersection of the subsets where $f > a - 1/n$, it follows that it is also a measurable set. By taking complements, we see that the sets where $f \leq a$ and where $f < a$ are measurable sets. By taking the intersection, we see that the set where $a < f \leq b$ is a measurable set. Since every open set is a union of intervals, the inverse image of every open subset is a measurable subset. It follows that $f$ is a measurable function. □

A set of real functions $L$ is called a *vector space* of functions if the zero function is in $L$, $f$ in $L$ and $g \in L$ imply that $f + g$ is in $L$, and $a$ in $\mathbb{R}$ and $f$ in $L$ imply that $af$ is in $L$. A set of real functions $L$ is called a *lattice* of functions if $f$ in $L$ and $g$ in $L$ imply that the infimum $f \wedge g$ is in $L$ and that the supremum $f \vee g$ is in $L$. The set $L$ is called a *vector lattice* of functions if it is both a vector space and a lattice.

Notice that if $f$ is in a vector lattice $L$, then the absolute value given by the formula $|f| = f \vee 0 - f \wedge 0$ is in $L$.

**Theorem 7.4** *The collection of measurable functions forms a $\sigma$-algebra of functions. That is, it is a vector lattice of functions that contains the constant functions and is closed under pointwise monotone limits of sequences.*

Proof: First we prove that the collection of measurable functions is a lattice. That is, we prove that if $f$ and $g$ are measurable functions, then $f \vee g$ and $f \wedge g$ are measurable functions. But $f \vee g \leq a$ precisely where $f \leq a$ and $g \leq a$. This is the intersection of two measurable subsets, so it is measurable. A similar argument works for the $f \wedge g$.

Next we prove that the collection of measurable functions is a vector space. That is we prove that if $f$ and $g$ are measurable, then so are $f + g$ and $cf$. The proof for $f + g$ is not completely obvious, but there is a trick that works. It is to note that $f + g > a$ if and only if there is a rational number $r$ such that $f > a - r$ and $g > r$. Since this is a countable union, we get a measurable subset. The proof for $cf$ is easy and is left to the reader.

Next we prove that the constant functions are measurable. This is because the space $X$ and the empty set $\emptyset$ are always measurable subsets.

Finally we prove that the collection of measurable functions is closed under pointwise monotone convergence of sequences. In fact, it is closed under the operation of taking the supremum of a sequence. This is because if $f_n$ is a sequence of functions, then the set where $\sup_n f_n \leq a$ is the intersection of the sets where $f_n \leq a$. Similarly, it is closed under the taking the infimum of a sequence. □

The $\sigma$-algebra $\mathcal{F}$ of measurable subsets thus gives rise to a $\sigma$-algebra $\mathcal{F}$ of real functions. This in turn determines the original $\sigma$-algebra of subsets. In fact, a subset $A$ is measurable if and only if the indicator function $1_A$ is a measurable function.

For later use, notice that the the collection $\mathcal{F}^+$ of positive measurable functions is not a vector space, but it is a cone. This means that $\mathcal{F}^+$ is a non-empty set of functions such that if $f, g$ are in the $\mathcal{F}^+$, then $f + g$ is in $\mathcal{F}^+$, and if also $a \geq 0$, then $af$ is in $\mathcal{F}^+$.

## 7.6 $\sigma$-algebras of functions

The previous theorem suggests a variant definition of measurable space. A *measurable space* $X$ is a set $X$ together with a given*$\sigma$-algebra $\mathcal{F}$ of real functions* on $X$. That is, there is a given vector lattice $\mathcal{F}$ of real functions that contains the constant functions and that is closed under increasing pointwise limits of sequences. That is, if $f_n \uparrow f$ with pointwise convergence, and each $f_n$ is in $\mathcal{F}$, then $f$ is in $\mathcal{F}$. A function in this space is called a *measurable function.*

Of course, if the vector lattice $\mathcal{F}$ is closed under increasing pointwise limits of sequence, then it is also closed under decreasing pointwise limits of sequence. In fact, then the vector lattice $\mathcal{F}$ is closed under all pointwise limits of sequences. Suppose that $f_n \to f$ pointwise, where each $f_n$ is in $\mathcal{F}$. Fix $k$ and $m$. Since $\mathcal{F}$ is a lattice, we see that $g_{km} = \sup_{k \le n \le m} f_n$ is in $\mathcal{F}$. However $g_{km} \uparrow g_k = \sup_{k \le n} f_n$ as $m \to \infty$, so $g_k$ is also in $\mathcal{F}$. Then $g_k \downarrow \limsup_n f_n$ as $k \to \infty$. Since $f = \limsup_n f$, we are done.

Examples:

1. The first and simplest standard example is when $X$ is a countable set, and the $\sigma$-algebra consists of all real functions on $X$.

2. The second standard example is when $X = \mathbb{R}$ and the $\sigma$-algebra is the smallest $\sigma$-algebra that contains all the continuous real functions on the metric space $\mathbb{R}$. We shall see that this example is the Borel $\sigma$-algebra. A variant of this example is when $\mathbb{R}$ is replaced by$[0, 1]$.

A $\sigma$-algebra $\mathcal{F}$ of real functions gives rise to a corresponding $\sigma$-algebra of measurable sets, consisting of the sets $A$ such that $1_A$ is in $\mathcal{F}$.

**Theorem 7.5** *Consider a set $X$ and a $\sigma$-algebra $\mathcal{F}$ of real functions. Define the $\sigma$-algebra $\mathcal{F}_X$ of sets to consist of those subsets of $X$ whose indicator functions belong to $\mathcal{F}$. Then $\mathcal{F}$ consists of those functions that are measurable with respect to this $\sigma$-algebra $\mathcal{F}_X$.*

Proof: Consider a real function $f$ in $\mathcal{F}$ and a real number $a$. Then $f - f \wedge a$ is also in $\mathcal{F}$. The sequence of function $h_n = n(f - f \wedge a) \wedge 1$ converges pointwise to $1_{f>a}$. So $1_{f>a}$ is in $\mathcal{F}$. Thus $f > a$ is in the $\sigma$-algebra of sets. This is enough to prove that $f$ is measurable with respect to this $\sigma$-algebra.

Consider on the other hand a real function $f$ that is measurable with respect to the $\sigma$-algebra of subsets. Then $f > a$ is a subset in the $\sigma$-algebra. This says that $1_{f>a}$ is in $\mathcal{F}$. Then if $a < b$, then $1_{a<f\le b} = 1_{f>a} - 1_{f>b}$ is also in $\mathcal{F}$. Consider the numbers $c_{nk} = k/2^n$ for $n \in \mathbb{N}$ and $k \in \mathbf{Z}$. Then $f_n = \sum_k c_{nk} 1_{c_{nk}<f\le c_{n\,k+1}}$ is also in $\mathcal{F}$. However $f_n \uparrow f$ as $n \to \infty$. So $f$ is in $\mathcal{F}$. □

**Proposition 7.6** *Let $\mathcal{F}$ be a $\sigma$-algebra of functions. If $f$ is in $\mathcal{F}$, then so is $f^2$.*

Proof: Since $\mathcal{F}$ is a lattice, $f$ in $\mathcal{F}$ implies $|f|$ in $\mathcal{F}$. For $a \geq 0$ the condition $f^2 > a$ is equivalent to the condition $|f| > \sqrt{a}$. On the other hand, for $a < 0$ the condition$f^2 > a$ is always satisfied. □

**Theorem 7.7** *Let $\mathcal{F}$ be a $\sigma$-algebra of functions. If $f$, $g$ are in $\mathcal{F}$, then so is the pointwise product $fg$.*

Proof: Since $\mathcal{F}$ is a vector space, it follows that $f + g$ and $f - g$ are in $\mathcal{F}$. However $4fg = (f + g)^2 - (f - g)^2$. □

This last theorem shows that $\mathcal{F}$ is not only closed under addition, but also under multiplication. Thus $\mathcal{F}$ deserves to be called an algebra. It is called a $\sigma$-algebra because of the closure under pointwise sequential limits.

Example: There are situations when it is useful to consider several $\sigma$-algebras at once. For instance, let $X = \mathbb{R}^n$. For each $k = 1, \ldots, n$ let $\sigma(x_1, \ldots, x_k)$ be the $\sigma$-algebra of real functions on $\mathbb{R}^n$ consisting of all Borel functions of the coordinates $x_1, \ldots, x_k$. In other words, the function depends only on the first $k$ coordinates.

This is the same example of unfolding information as before. However then the set point of view led to thinking of $\sigma(x_1, \ldots, x_k)$ as the collection of questions that are answered by time $k$. The function point of view instead views $\sigma(x_1, \ldots, x_k)$ as the collection of experimental quantities whose values are known at time $k$.

**Proposition 7.8** *A map $\phi : X \to Y$ is a measurable map if and only if for each measurable real function $f$ on $Y$ the composition $f \circ \phi$ is a measurable real function on $X$.*

## 7.7 Borel functions

Consider the important special case when $X$ is a metric space equipped with its Borel $\sigma$-algebra $\mathcal{B}o_X$ of subsets. Then the corresponding $\sigma$-algebra $\mathcal{B}o$ of real measurable functions consists of the *Borel functions*. Such functions are said to be *Borel measurable functions*. It is clear that the space $C(X)$ of real continuous functions is a subset of $\mathcal{B}o$.

**Theorem 7.9** *If $X$ is a metric space, then the smallest $\sigma$-algebra including $C(X)$ is the Borel $\sigma$-algebra $\mathcal{B}o$.*

Proof: Consider the $\sigma$-algebra of subsets that consists of the inverse images of Borel subsets under continuous real functions. It is sufficient to show that every subset in $\mathcal{B}o_X$ is in this $\sigma$-algebra. To prove this, it is sufficient to show that every closed set is in it. Let $F$ be a closed subset. Then the function $f(x) = d(x, F)$ is a continuous function that vanishes precisely on $F$. That is, the inverse image of $\{0\}$ is $F$. □

## 7.8 Supplement: Generating sigma-algebras

This section is devoted to a fundamental fact: A Borel function of a measurable function is a measurable function. So a class of measurable functions is closed under just about any reasonable operation one can perform on the ranges of the functions.

If we are given a set $S$ of functions, then the $\sigma$-algebra of functions $\sigma(S)$ generated by this set is the smallest $\sigma$-algebra of functions that contains the original set. The Borel $\sigma$-algebra $\mathcal{B}o$ of functions on $\mathbb{R}$ is generated by the single function $x$. Similarly, the Borel $\sigma$-algebra of functions on $\mathbb{R}^k$ is generated by the coordinates $x_1, \ldots, x_k$. The following theorem shows that measurable functions are closed under nonlinear operations in a very strong sense.

**Theorem 7.10** *Let $f_1, \ldots, f_k$ be functions on $X$. Let $\mathcal{B}o$ be the $\sigma$-algebra of Borel functions on $\mathbb{R}^k$. Let*

$$\mathcal{G} = \{\phi(f_1, \ldots, f_k) \mid \phi \in \mathcal{B}o\}. \tag{7.1}$$

*The conclusion is that $\sigma(f_1, \ldots, f_k) = \mathcal{G}$. That is, the $\sigma$-algebra of functions generated by $f_1, \ldots, f_k$ consists of the Borel functions of the functions in the generating set.*

Proof: First we show that $\mathcal{G} \subset \sigma(f_1, \ldots, f_k)$. Let $\mathcal{B}'$ be the set of functions $\phi$ such that $\phi(f_1, \ldots, f_k) \in \sigma(f_1, \ldots, f_k)$. Each coordinate function $x_j$ of $\mathbb{R}^n$ is in $\mathcal{B}'$, since this just says that $f_j$ is in $\sigma(f_1, \ldots, f_k)$. Furthermore, $\mathcal{B}'$ is a $\sigma$-algebra. This is a routine verification. For instance, here is how to check upward monotone convergence. Suppose that $\phi_n$ is in $\mathcal{B}'$ for each $n$. Then $\phi_n(f_1, \ldots, f_k) \in \sigma(f_1, \ldots, f_k)$ for each $n$. Suppose that $\phi_n \uparrow \phi$ pointwise. Then $\phi_n(f_1, \ldots, f_k) \uparrow \phi(f_1, \ldots, f_k)$, so $\phi(f_1, \ldots, f_k) \in \sigma(f_1, \ldots, f_k)$. Thus $\phi$ is in $\mathcal{B}'$. Since $\mathcal{B}'$ is a $\sigma$-algebra containing the coordinate functions, it follows that $\mathcal{B}o \subset \mathcal{B}'$. This shows that $\mathcal{G} \subset \sigma(f_1, \ldots, f_k)$.

Now we show that $\sigma(f_1, \ldots, f_k) \subset \mathcal{G}$. It is enough to show that $\mathcal{G}$ contains $f_1, \ldots, f_k$ and is a $\sigma$-algebra of functions. The first fact is obvious. To show that $\mathcal{G}$ is a $\sigma$-algebra of functions, it is necessary to verify that it is a vector lattice with constants and is closed under monotone convergence. The only hard part is the monotone convergence. Suppose that $\phi_n(f_n, \ldots, f_k) \uparrow g$ pointwise. The problem is to find a Borel function $\phi$ such that $g = \phi(f_1, \ldots, f_k)$. There is no way of knowing whether the Borel functions $\phi_n$ converge on all of $\mathbb{R}^k$. However let $G$ be the subset of $\mathbb{R}^k$ on which $\phi_n$ converges. Then $G$ also consists of the subset of $\mathbb{R}^k$ on which $\phi_n$ is a Cauchy sequence. So

$$G = \bigcap_j \bigcup_N \bigcap_{m \geq N} \bigcap_{n \geq N} \{x \mid |\phi_m(x) - \phi_n(x)| < 1/j\} \tag{7.2}$$

is a Borel set. Let $\phi$ be the limit of the $\phi_n$ on $G$ and $\phi = 0$ on the complement of $G$. Then $\phi$ is a Borel function. Next note that the range of $f_1, \ldots, f_k$ is a subset of $G$. So $\phi_n(f_1, \ldots, f_k) \uparrow \phi(f_1, \ldots, f_k) = g$. $\square$

**Corollary 7.11** *Let $f_1, \ldots, f_n$ be in a $\sigma$-algebra $\mathcal{F}$ of measurable functions. Let $\phi$ be a Borel function on $\mathbb{R}^n$. Then $\phi(f_1, \ldots, f_n)$ is also in $\mathcal{F}$.*

Proof: From the theorem $\phi(f_1, \ldots, f_n) \in \sigma(f_1, \ldots, f_n)$. Since $\mathcal{F}$ is a $\sigma$-algebra and $f_1, \ldots, f_n$ are in $\mathcal{F}$, it follows that $\sigma(f_1, \ldots, f_n) \subset \mathcal{F}$. Thus $\phi(f_1, \ldots, f_n) \in \mathcal{F}$. □

This discussion illuminates the use of the term measurable for elements of a $\sigma$-algebra. The idea is that there is a starting set of functions $S$ that are regarded as those quantities that may be directly measured in some experiment. The $\sigma$-algebra $\sigma(S)$ consists of all functions that may be computed as the result of the direct measurement and other mathematical operations. Thus these are all the functions that are measurable. Notice that the idea of what is possible in mathematical computation is formalized by the concept of Borel function.

This situation plays a particularly important role in probability theory. For instance, consider the $\sigma$-algebra of functions $\sigma(S)$ generated by the functions in $S$. There is a concept of conditional expectation of a random variable $f$ given $S$. This is a numerical prediction about $f$ when the information about the values of the functions in $S$ is available. This conditional expectation will be a function in $\sigma(S)$, since it is computed by the mathematical theory of probability from the data given by the values of the functions in $S$.

# Problems

1. The most standard examples of $\sigma$-algebras are the $\sigma$-algebra of all subsets of a countable set and the $\sigma$-algebra of all Borel subsets of the line. There are also more exotic examples. Here is a relatively small one. Let $X$ be an uncountable set. Describe the smallest $\sigma$ algebra that contains all the one point subsets of $X$.

2. Let $\mathcal{B}o$ be the smallest $\sigma$-algebra of real functions on $\mathbb{R}$ containing the function $x$. This is called the $\sigma$-algebra of Borel functions. Show by a direct construction that every continuous function is a Borel function.

3. Show that every monotone function is a Borel function.

4. Can a Borel function be discontinuous at every point?

5. Let $\sigma(x^2)$ be the smallest $\sigma$-algebra of functions on $\mathbb{R}$ containing the function $x^2$. Show that $\sigma(x^2)$ is not equal to $\mathcal{B}o = \sigma(x)$. Which algebra of measurable functions is bigger (that is, which one is a subset of the other)?

6. Consider the $\sigma$-algebras of functions generated by $\cos(x)$, $\cos^2(x)$, and $\cos^4(x)$. Compare them with the $\sigma$-algebras in the previous problem and with each other. (Thus specify which ones are subsets or proper subsets of other ones.)

# Chapter 8

# Integrals

## 8.1 Measures and integrals

In this chapter the important concepts are integral and measure. An integral is defined on the positive elements of a $\sigma$-algebra of functions. A measure is defined on a $\sigma$-algebra of subsets. We shall see that an integral always determines a measure in a simple way. Conversely, there is a construction that can take us from a measure to an integral. So these are equivalent concepts.

A measure space is a set together with a given $\sigma$-algebra of functions and integral or a given $\sigma$-algebra of subsets and measure. While it is seldom that an integral is called a measure, it is quite common in many mathematical contexts to refer to an integral as a measure. Thus it is convenient to think of a measure space as consisting of a set, a $\sigma$-algebra, and a measure. This leaves a free choice of framework.

Consider a $\sigma$-algebra of measurable functions $\mathcal{F}$ on a set $X$. There is an associated cone $\mathcal{F}^+$ of positive measurable functions. An *integral* is a function

$$\mu : \mathcal{F}^+ \to [0, +\infty] \tag{8.1}$$

such that

1. $\mu(0) = 0$,
2. For each real $a > 0$ we have $\mu(af) = a\mu(f)$,
3. $\mu(f + g) = \mu(f) + \mu(g)$,
4. If $f_n \uparrow f$ pointwise, then $\mu(f_n) \uparrow \mu(f)$.

The last condition is often called *monotone convergence*. Sometimes it is also called *countable additivity*. This is because if we have $f_n = \sum_{k=1}^{n} w_k$ and $f = \sum_{k=1}^{\infty} w_k$ with each $w_k \geq 0$, then $\mu(f) = \lim_n \mu(f_n)$ says that

$$\mu(\sum_{k=1}^{\infty} w_k = \sum_{k=1}^{\infty} \mu(w_k). \tag{8.2}$$

Sometimes such a general integral is called an *abstract Lebesgue integral.* This is because it is a generalization of the usual translation-invariant Lebesgue integral defined for functions on the line or on $\mathbb{R}^n$.

Consider a $\sigma$-algebra $\mathcal{F}_X$ of subsets of $X$. Define the measure $\mu(B)$ of subsets by $\mu(B) = \mu(1_B)$. Then

1. $\mu(\emptyset) = 0$,
2. If $A \cap B = \emptyset$, then $\mu(A \cup B) = \mu(A) + \mu(B)$,
3. If $A_n \uparrow A$, then $\mu(A_n) \uparrow \mu(A)$.

Monotone convergence in this case is more often called countable additivity. This is because if we have $A_n = \bigcup_{k=1}^n B_k$ for disjoint subsets $B_k$, then

$$\mu(\bigcup_{k=1}^{\infty} B_k) = \sum_k \mu(B_k). \tag{8.3}$$

Sometimes such a general measure is called an *abstract Lebesgue measure.* This is because it is a generalization of the usual translation-invariant Lebesgue measure defined for subsets of the line or of $\mathbb{R}^n$.

**Theorem 8.1** *Consider an integral $\mu$ associated with a $\sigma$-algebra $\mathcal{F}$ of real functions on $X$. Then the restriction of this integral to indicator functions in $\mathcal{F}$ determines a measure on the associated $\sigma$-algebra $\mathcal{F}_X$ of subsets of $X$. The measure uniquely determines the integral.*

Proof: The fact that an integral determines a corresponding measure is obvious. The uniqueness follows from writing $c_{nk} = k/2^n$ for $n \in \mathbf{N}$ and $k \in \mathbf{N}$ and representing

$$f_n = \sum_k c_{nk} 1_{c_{nk} < f \leq c_{n\ k+1}}. \tag{8.4}$$

Then by countable additivity

$$\mu(f_n) = \sum_k c_{nk} \mu(c_{nk} < f \leq c_{n\ k+1}). \tag{8.5}$$

This gives an approximation to the integral in terms of the measure. Then

$$\mu_n(f) = \lim_{n \to \infty} \mu(f_n) \tag{8.6}$$

gives the integral itself. □

It is shown in many books on real analysis that there is a converse: Each measure on a $\sigma$-algebra of subsets uniquely determines an integral on the associated $\sigma$-algebra of real measurable functions. The notions of integral and measure are thus totally equivalent.

A *measure space* is a triple $X, \mathcal{F}, \mu$, where $X$ is a set, $\mathcal{F}$ is a $\sigma$-algebra of sets or of functions, and $\mu$ is a measure or integral. When the $\sigma$-algebra is clear from context, then sometimes a measure space is denoted $X, \mu$. Sometimes even the measure is taken for granted; then the measure space becomes simply $X$. In many contexts the word *measure* is used to mean integral. So one could describe a measure space as a set, a $\sigma$-algebra, and a measure. This does not prevent us from thinking of the measure as an integral acting on measurable functions, and often this point of view gives a simpler and more uniform description.

The notion of measure of subsets has an intuitive appeal, since we can think of the measure of a subset $A$ as a mass. For instance, if we have sand distributed in space, then $\mu(A)$ is the amount of sand in the region $A$ of space. This also gives a way of thinking about an integral. Say that the sand has a concentration of gold that varies from point to point. Consider $f$ as a real function that describes the economic value of the sand at each particular location. Then $\mu(f)$ is the total economic value of the sand at all locations.

Consider a set $X$ and a $\sigma$-algebra $\mathcal{F}$ of real functions on $X$. Consider an integral $\mu$ defined on $\mathcal{F}^+$, or the corresponding measure $\mu$ defined by $\mu(A) = \mu(1_A)$. The integral or measure is called *finite* if the integral $\mu(1)$ is finite. This is the same as requiring that the measure $\mu(X)$ is finite. This number is called the total mass of the measure.

In the case when the total mass of a measure is equal to one, the measure is called a *probability measure*, and the corresponding integral is called an *expectation*.

Examples:

1. The simplest standard example is when $X$ is a countable set. The $\sigma$-algebra of subsets consists of all subsets; the $\sigma$-algebra of real functions consists of all real functions. The integral is $\sum$, that is, *summation* of a function over the set. It can be written as

$$\sum f = \sum_{x \in X} f(x). \tag{8.7}$$

   The corresponding measure is called *counting measure*. The measure of a subset $A \subset X$ is the number of points $\#(A)$ in the subset.

   Another integral constructed from summation is a weighted sum

$$\mu(f) = \sum f w = \sum_{x \in X} f(x) w(x), \tag{8.8}$$

   where $w \geq 0$ is a given positive real function on $X$.

2. The second standard example is when $X = \mathbb{R}$. In this example we take the $\sigma$-algebra of subsets to consist of all Borel subsets. Correspondingly, the $\sigma$-algebra of real functions consists of all Borel functions. The integral is the Lebesgue integral $\lambda$ given by

$$\lambda(f) = \int_{-\infty}^{\infty} f(x)\, dx. \tag{8.9}$$

The corresponding measure is called Lebesgue measures.

Another integral constructed from this example would be a weighted integral

$$\mu(f) = \lambda(fw) = \int_{-\infty}^{\infty} f(x)w(x)\,dx, \tag{8.10}$$

where $w \geq 0$ is a given positive real Borel function.

3. The third standard example is when $X = [0,1]$. Again we use the Borel $\sigma$-algebra. The integral is the Lebesgue uniform distribution integral $\lambda_1$ given by

$$\lambda_1(g) = \int_0^1 g(x)\,dx. \tag{8.11}$$

The corresponding measure $\lambda_1$ defined on Borel measurable subsets of $[0,1]$ is called the uniform probability measure on the unit interval, or Lebesgue measure on the unit interval. It has the property that $\lambda(I_{n;z}) = 1/2^n$.

4. Here is a variant of the third standard example. There is a measure $\mu$ defined on measurable subsets of $2^{\mathbb{N}_+}$ with the property that $\mu(F_{n,z}) = 1/2^n$. This is called the fair coin-tossing probability measure. The corresponding integral is called the expectation for fair coin-tossing.

**Proposition 8.2** *If $0 \leq f \leq g$, then $0 \leq \mu(f) \leq \mu(g)$.*

Proof: Clearly $(g-f)+f = g$. So $\mu(g-f)+\mu(f) = \mu(g)$. But $\mu(g-f) \geq 0$. □

If $f$ in $\mathcal{F}$ is a measurable function, then its positive part $f_+ = f \vee 0 \geq 0$ and its negative part $f_- = -(f \wedge 0) = -f \vee 0 \geq 0$. So they each have integrals. If either $\mu(f_+) < +\infty$ or $\mu(f_-) < +\infty$, then we may define the integral of $f = f_+ - f_-$ to be

$$\mu(f) = \mu(f_+) - \mu(f_-). \tag{8.12}$$

In this case we say that $f$ is a *definitely integrable function.* (This is not standard terminology, but it seems helpful to have a word to fix the concept.) The possible values for this integral are real, $+\infty$, or $-\infty$. There is no ambiguity of any kind.

If both $\mu(f_+) = +\infty$ and $\mu(f_-) = +\infty$, then the integral is not defined. The expression $(+\infty) - (+\infty) = (+\infty) + (-\infty)$ is in general quite ambiguous! This is the major flaw in the theory, and it is responsible for most challenges in applying the theory of integration.

Of course in certain instances it may be possible to give a finite value to an integral by a limiting process. Thus while $\sin(x)/x$ is not definitely integrable, it is true that

$$\lim_{a\to\infty} \int_{-a}^{a} \frac{\sin(x)}{x}\,dx = \pi. \tag{8.13}$$

In such situations one has to be rather fussy about how one takes the limit.

The difficulty with the definition of integral should be contrasted with the situation in the case of measure. The measure of a measurable subset is always defined; it may be infinite. In the sand analogy, this says that the amount of sand in a region of space is a definite physical quantity. Suppose now that $f$ is a real function that represents the economic value of the sand at each point. This economic value may be positive (a valuable mineral) or negative (radioactive waste). If $f$ is not definitely integrable, it says that we are in such a heady economic context that we are trying to balance infinite positive gains with infinite negative gains. So we are no longer in the domain of physics, but instead of a variant on Pascal's wager.

**Theorem 8.3** *Suppose that* $\mu(f_-) < +\infty$ *and* $\mu(g_-) < +\infty$. *Then* $\mu(f + g) = \mu(f) + \mu(g)$.

Proof: Let $h = f + g$. Then $h_+ \leq f_+ + g_+$ and $h_- \leq f_- + g_-$. So under the hypothesis of the theorem $\mu(h_-) < +\infty$. Furthermore, from $h_+ - h_- = f_+ - f_- + g_+ - g_-$ it follows that $h_+ + f_- + g_- = h_- + f_+ + g_+$. Since these are all positive functions $\mu(h_+) + \mu(f_-) + \mu(g_-) = \mu(h_-) + \mu(f_+) + \mu(g_+)$. However then $\mu(h_+) - \mu(h_-) = \mu(f_+) - \mu(f_-) + \mu(g_+) - \mu(g_-)$. This is allowed because the terms that are subtracted are not infinite. The conclusion is that $\mu(h) = \mu(f) + \mu(g)$. $\square$

**Theorem 8.4** *If* $f$ *is in* $\mathcal{F}$ *and* $\mu(|f|) = \mu(f_+) + \mu(f_-) < \infty$, *then* $\mu(f) = \mu(f_+) - \mu(f_-)$ *is defined, and*

$$|\mu(f)| \leq \mu(|f|). \tag{8.14}$$

If $f$ is in $\mathcal{F}$, then $f$ is said to be an *absolutely integrable function* with respect to $\mu$ if

$$\mu(|f|) = \mu(f_+) + \mu(f_-) < +\infty. \tag{8.15}$$

The collection of absolutely integrable functions is called $\mathcal{L}^1(X, \mathcal{F}, \mu)$. So

$$\mu : \mathcal{L}^1(X, \mathcal{F}, \mu) \to \mathbb{R} \tag{8.16}$$

is well-defined.

An absolutely integrable function must be definitely integrable. A definitely integrable function is absolutely integrable if and only if its integral is finite.

Note: In many treatments of integration theory an absolutely integrable function is called an integrable function, and there is no term for definitely integrable function. As mentioned before, in some contexts an integral is said to exist and be finite even when it is not definitely integrable; but this situation always involves a delicate limiting definition of integral, outside the scope of the present theory.

Consider a vector lattice $\mathcal{L}$ of real functions. Suppose that there is a real function $\mu : \mathcal{L} \to \mathbb{R}$ that is linear and order preserving and satisfies the following integral bounded monotone convergence closure property: If $f_n$ is a sequence in $\mathcal{L}$ with $f_n \uparrow f$ as $n \to \infty$, and if there is a real constant $M$ such that for each $n$ we have $\mu(f_n) \leq M$, then $f$ is in $\mathcal{L}$ and $\mu(f_n) \uparrow \mu(f)$. Then this $\mu$ is called an *absolute integral.* While this term is not standard, the following theorem states that the restriction of an integral to the absolutely integrable functions is an absolute integral, and so it seems a reasonable concept.

**Theorem 8.5** *Consider a $\sigma$-algebra $\mathcal{F}$ of real functions and an integral $\mu : \mathcal{F}^+ \to [0, +\infty]$. Then the associated $\mu : \mathcal{L}^1(X, \mathcal{F}, \mu) \to \mathbb{R}$ is an absolute integral.*

## 8.2 Borel measures

For every metric space (or even topological space) there is a distinguished $\sigma$-algebra of subsets called the Borel $\sigma$-algebra, generated by the open subsets. A measure defined on the Borel $\sigma$-algebra is called a *Borel measure*.

The most obvious example is the usual Lebesgue measure on the Borel subsets of the line. It is possible for a Borel set with Lebesgue measure zero to have a subset that is not in the Borel $\sigma$-algebra. Therefore some authors like to consider a much larger $\sigma$-algebra, called the *Lebesgue $\sigma$-algebra*, and to define the *completed Lebesgue measure* on this larger domain. This has the property that every subset of a set of measure zero is also measurable.

The Borel $\sigma$-algebra has the advantage that it is defined independent of a particular measure, and so it is the potential domain for many measures. Furthermore, it is already large enough for most purposes. Therefore in most concrete applications in the present treatment the measures are defined on the Borel $\sigma$-algebra or on even smaller $\sigma$-algebras.

The *Lebesgue-Stieltjes measures* are Borel measures that generalize Lebesgue measure. Each such measure is defined by an increasing right continuous function $F : \mathbf{R} \to \mathbf{R}$. This function is called a *distribution function*. The measure associated to $F$ is $\sigma_F$, and it is characterized by the property that $\sigma_F((a, b]) = F(b) - F(a)$. Lebesgue measure is then the special case when $F(x) = x$. The corresponding Lebesgue-Stieltjes integral is denoted

$$\sigma_F(g) = \int_{-\infty}^{\infty} g(x) dF(x). \tag{8.17}$$

Notice that if $c < d$ and $F(c) = F(d)$, then the measure $\sigma_F((c, d]) = 0$. There are subsets of the interval $(c, d]$ that are not Borel measurable. It may be shown that there are even subsets of $(c, d]$ that are not Lebesgue measurable. One could try to extend the domain of the measure to such subsets, but this is usually not necessary. Borel measures suffice for most purposes.

The other extreme is that the distribution function has a jump discontinuity at some point $p$. In that case, the measure of the one point set $\{p\}$ is the size of the jump.

If $w \geq 0$ is a positive real Borel function, then it may be regarded as a weight function that defines a measure $\mu$ by $\mu(f) = \lambda(fw)$. Suppose that $w$ has a finite integral over every bounded interval. Then this measure is a Lebesgue-Stieltjes measure with distribution function $F$ satisfying by

$$F(b) - F(a) = \int_a^b w(x)\, dx. \tag{8.18}$$

The function $F$ is said to be a distribution function with *density* $w$. However it is possible to have a distribution function that is not given by integrating a density. For instance, a distribution function with a jump discontinuity has this feature. Thus not every Lebesgue-Stieltjes measure is given by a density.

## 8.3 Image measures and image integrals

Given an integral $\mu$ defined on $\mathcal{F}^+$, and given a measurable map $\phi : X \to Y$, there is an integral $\phi[\mu]$ defined on $\mathcal{G}^+$. It is given by

$$\phi[\mu](g) = \mu(g \circ \phi). \tag{8.19}$$

It is called the *image* of the integral $\mu$ under $\phi$.

The same construction works for measures. Since the measure $\mu(A)$ is the same as the integral $\mu(1_A)$, the *image measure* is given by $\phi[\mu][1_A] = \mu(1_A \circ \phi)$. However $1_A \circ \phi = 1_{\phi^{-1}[A]}$. So for measures

$$\phi[\mu](A) = \mu(\phi^{-1}[A]). \tag{8.20}$$

Example: Let $\phi : 2^{\mathbb{N}_+} \to [0,1]$ be defined by $\phi(x) = \sum_{k=1}^{\infty} x_k/2^k$. Let $\mu$ be the fair coin tossing probability measure, and let $\lambda_1$ be the uniform probability measure. Then $\phi[\mu] = \lambda_1$. This is consistent with the fact that $\lambda_1(I_{n;z}) = \mu(\phi^{-1}[I_{n,z}]) = \mu(F_{n;z}) = 1/2^n$.

## Problems

1. Let $X$ be the unit interval $[0,1]$. Consider the smallest $\sigma$ algebra of subsets that contains all the one point subsets of $X$. Restrict the Lebesgue probability measure $\lambda_1$ to this $\sigma$-algebra. Prove that every measurable subset has measure zero or has measure one.

2. Let $X$ be an uncountable set. First consider the $\sigma$-algebra of all subsets with counting measure. (The corresponding integral is just summation over $X$.) Second, consider the $\sigma$-algebra generated by the one point subsets and the restriction of counting measure to this smaller $\sigma$-algebra. Prove that every absolutely integrable function with respect to the first $\sigma$-algebra is actually measurable with respect to the second smaller $\sigma$-algebra.

# Chapter 9

# Elementary integrals

## 9.1 Stone vector lattices of functions

The purpose of this chapter is to explore the concept of a elementary integral on a Stone vector lattice of real functions. In the following chapter we shall see each such elementary integral will give to an integral defined with respect to a $\sigma$-algebra of real functions.

Recall that a set of real functions $L$ is called a vector lattice of real functions if it is both a vector space and a lattice (under the pointwise operations). Sometimes a vector lattice may have the function 1 in it, and hence have every constant function in it. However this is not required. A weaker condition often suffices. A vector lattice of real functions $L$ is a *Stone vector lattice* of functions if it is a vector lattice and satisfies the property: $f \in L$ implies $f \wedge 1 \in L$.

It the following we shall often use the notion of the *support* of a real function. This is the closure of the set on which it is non-zero.

Examples:

1. The space $L = C([0, 1])$ of real continuous functions on the unit interval is a Stone vector lattice, in fact 1 belongs to $L$.

2. The space $L = C_c(\mathbb{R})$ of real continuous functions on the line, each with compact support, is a Stone vector lattice.

3. A *rectangular function* is an indicator function of an interval $(a, b]$ of real numbers. Here $a$ and $b$ are real numbers, and the interval $(a, b]$ consists of all real numbers $x$ with $a < x \leq b$. The convention that the interval is open on the left and closed on the right is arbitrary but convenient. The nice thing about these intervals is that their intersection is an interval of the same type. Furthermore, the union of two such intervals is a finite union of such intervals. And the relative complement of two such intervals is a finite union of such intervals.

A *step function* is a finite linear combination of rectangular functions. In fact, each step function may be represented as a finite linear combination of rectangular functions that correspond to disjoint subsets. The space $L$ of real step functions (finite linear combinations of rectangle functions) defined on the line is a Stone vector lattice.

4. Consider the coin-tossing space $\Omega = 2^{\mathbf{N}_+}$ of all sequences of zeros and ones indexed by $\mathbf{N}_+ = \{1, 2, 3, \ldots\}$. This is thought of as a sequences of tails and heads, or of failures and successes. Each element $\omega$ is called an *outcome* of the coin tossing experiment. For $n \geq 0$, let $\mathcal{F}_n$ be the set of real functions on $\Omega$ that depend at most on the first $n$ coordinates.

   A *random variable* is a function $f$ from $\Omega$ to $\mathbb{R}$ (satisfying certain technical conditions, to be specified later). A random variable is a prescription for determining an experimental number, since the number $f(\omega)$ depends on the actual result $\omega$ of the experiment. Each function in $\mathcal{F}_n$ is a random variable. These are the random variables that may be determined only knowing the results of the first $n$ coin tosses.

   One important function in $\mathcal{F}_n$ is the *binary function* $f_{n;z}$ defined for $0 \leq n$ and a finite sequence $z$ of zeros and ones. Then $f_{n;z}$ is equal to one on every sequence $\omega$ that agrees with $z$ in the first $n$ places. Then the $f_{n;z}$ for $0 \leq k < 2^n$ form a basis for the $2^n$ dimensional vector space $\mathcal{F}_n$. The binary function is one precisely for those coin tosses that have a particular pattern of successes and failures in the first $n$ trials, without regard to what happens in later trials.

   An important infinite dimensional vector space is

$$L = \bigcup_{n=0}^{\infty} \mathcal{F}_n. \tag{9.1}$$

   This is the space of all random variables that each depend only on some finite number of trials. Each function in $L$ is a continuous function on $\Omega$. The space $L$ is a Stone vector lattice, in fact, the function 1 is in $L$.

Each of the vector lattices $L$ in the above examples defines a corresponding $\sigma$-algebra $\sigma(L)$.

Examples:

1. When $L = C([0, 1])$ the corresponding $\sigma$-algebra $\sigma(L)$ is the the smallest $\sigma$-algebra including the continuous functions. Since $[0, 1]$ is a metric space, this is the Borel $\sigma$-algebral $\mathcal{B}o$ of all real Borel functions on $[0, 1]$.

2. When $L = C_c(\mathbb{R})$ the corresponding $\sigma$-algebra $\sigma(L)$ is the Borel $\sigma$-algebra $\mathcal{B}o$ of all real Borel functions on $\mathbb{R}$. This is perhaps less obvious. However it is not difficult to see that each continuous real function on $\mathbb{R}$ is a pointwise limit of of a sequence of continuous functions with compact support. Since the continuous functions generate the Borel functions, the result is again the Borel $\sigma$-algebra.

3. When $L$ consists of step functions on the line the corresponding $\sigma$-algebra $\sigma(L)$ is the Borel $\sigma$-algebra $\mathcal{B}o$ of all real Borel functions on $\mathbb{R}$. This follows from the fact that each continuous real function on $\mathbb{R}$ is a pointwise limit of a sequence of step functions.

4. When $L$ consists of functions on the Cantor space that each depend on only finitely many coordinates, then the corresponding $\sigma$-algebra $\sigma(L)$ is the Borel $\sigma$-algebra of all real Borel functions on the Cantor space. In order to see this, recall that the Cantor space is a metric space with the product metric that defines pointwise convergence. Each real function on the Cantor space that depends on only finitely many coordinates is continuous. On the other hand, let $f$ be a continuous function on the Cantor space. Let $f_n(\omega) = f(\bar{\omega}_n)$, where $\bar{\omega}_n$ agrees with $\omega$ in the first $n$ coordinates and is zero in the remaining coordinates. Since $\omega_n \to \omega$ as $n \to \infty$ in the product metric, it follows that $f_n \to f$ pointwise as $n \to \infty$. This shows that the continuous functions are in $\sigma(L)$. It follows that $\sigma(L) = \mathcal{B}o$.

In each case this Borel $\sigma$ algebra $\mathcal{B}o$ of functions is huge; in fact, it is difficult to think of a real function that does not belong to $\mathcal{B}o$. However it is possible to show that the cardinality of $\mathcal{B}o$ is $c$, while the cardinality of the $\sigma$ algebra of all real functions is $2^c$.

## 9.2 Elementary integrals

Let $X$ be a non-empty set. Let $L$ be a vector lattice of real functions on $X$. Then $\mu$ is an *elementary integral* on $L$ provided that

1. $\mu : L \to \mathbb{R}$ is linear;
2. $\mu : L \to \mathbb{R}$ is order preserving;
3. $\mu$ satisfies monotone convergence within $L$.

To say that $\mu$ satisfies monotone convergence within $L$ is to say that if each $f_n$ is in $L$, and $f_n \uparrow f$, and $f$ is in $L$, then $\mu(f_n) \uparrow \mu(f)$.

**Proposition 9.1** *Suppose that $g_n$ in $L$ and $g_n \downarrow 0$ imply $\mu(g_n) \downarrow 0$. Then $\mu$ satisfies monotone convergence within $L$.*

Proof: Suppose that $f_n$ is in $L$ and $f_n \uparrow f$ and $f$ is in $L$. Since $L$ is a vector space, it follows that $g_n = f - f_n$ is in $L$. Furthermore $g_n \downarrow 0$. Therefore $\mu(g_n) \downarrow 0$. This says that $\mu(f_n) \uparrow \mu(f)$. □

Examples:

1. For the $L = C([0, 1])$, the continuous real functions on $[0, 1]$ one can take the Riemann integral (or even the regulated integral defined using extension by uniform continuity). The fact that this is an elementary integral follows from Dini's theorem in the following section.

2. For the continuous functions $C_c(\mathbb{R})$ of continuous real functions on $\mathbb{R}$, each with compact support, we can again take the Riemann integral (or the regulated integral). Because the support of each function is compact there is no nonsense about integrals that are not absolutely convergent. Again the fact that this is an elementary integral follows from Dini's theorem.

3. For the space $L$ of step functions the elementary integral is just the integral of a finite linear combination of rectangle functions. Thus all one has to do is to add the areas of rectangles, with appropriate signs. This elementary integral is indeed elementary. On the other hand, the proof that it satisfies monotone convergence requires a more complicated version of Dini's theorem, presented in a later section of this chapter.

4. Let $\Omega = \{0,1\}^{\mathbb{N}_+}$ be the set of all infinite sequences of zeros and ones indexed by $\mathbb{N}_+ = \{1,2,3,\ldots\}$. For each $k = 0,1,2,3,\ldots$ consider the set $\mathcal{F}_k$ of functions $f$ on $\Omega$ that depend only on the first $k$ elements of the sequence, that is, such that $f(\omega) = g(\omega_1,\ldots,\omega_k)$ for some function $g$ on $\mathbb{R}^k$. This is a vector lattice with dimension $2^k$. The vector lattice under consideration will be the space $L$ that is the union of all the $\mathcal{F}_k$ for $k = 0,1,2,3,\ldots$. If the function $f$ is in $\mathcal{F}_k$, let

$$\mu(f) = \sum_{\omega_1=0}^{1} \cdots \sum_{\omega_k=0}^{1} f(\omega)\frac{1}{2^k}. \tag{9.2}$$

   This is a consistent definition, since if $f$ is regarded as being in $\mathcal{F}_j$ for $k < j$, then the definition involves sums over $2^j$ sequences, but the numerical factor is $1/2^j$, and the result is the same. This example describes the expectation for a function of the result of independent tosses of a fair coin. The fact that this defines an elementary integral follows from Dini's theorem.

## 9.3 Dini's theorem

**Theorem 9.2 (Dini's theorem)** *Suppose $K$ is compact. If $f_n$ is a sequence of continuous functions on $K$ and $f_n \downarrow 0$ pointwise as $n \to \infty$, then $f_n \to 0$ uniformly.*

Proof: Let $K$ be compact. Suppose each $f_n$ is in $C(K)$ and $f_n \downarrow 0$ pointwise. Let $\epsilon > 0$. Let $A_n$ be the closed set where $f_n \geq \epsilon$. Consider arbitrary $a$ in $K$. Since $f_n(a) \downarrow 0$, it follows that for large $n$ the point $a$ is not in $A_n$. Hence it is not in the intersection $\bigcap_n A_n$. Thus $\bigcap_n A_n = \emptyset$. By the finite-intersection property, there exists an $N$ so that $n \geq N$ implies $A_n = \emptyset$. Thus for $n \geq N$ we have $f_n < \epsilon$ at each point. It follows that $f_n \downarrow 0$ uniformly. □

Dini's theorem will have several applications. The first is to continuous functions with compact support. Consider the space of continuous real functions on the real line, each with compact support. Dini's theorem says that within this space monotone pointwise convergence to zero implies uniform convergence to zero.

## 9.4 Dini's theorem for step functions

A general step function can have arbitrary values at the end points of the intervals. It is sometimes nicer to make a convention that makes the step functions left continuous (or right continuous). This will eventually make things easier when dealing with more general integrals where individual points count.

If $f$ is a step function, then its integral $\lambda(f_n)$ is defined in a completely elementary way.

**Theorem 9.3** *For each function*

$$f = \sum_{k=1}^{m} c_k 1_{(a_k, b_k]} \tag{9.3}$$

*define*

$$\lambda(f) = \sum_{k=1}^{m} c_k (b_k - a_k). \tag{9.4}$$

*If $f_n \downarrow 0$ pointwise, then $\mu(f_n) \downarrow 0$.*

Proof: Say that $f_n \to 0$, where each $f_n$ is such a function. Notice that all the $f_n$ have supports in the interior of a fixed compact interval $[p, q]$. Furthermore, they are all bounded by some fixed constant $M$. Write $f_n = \sum_{k=1}^{m_n} c_{nk} 1_{(a_{nk}, b_{nk}]}$. For each $n$ and $k$ choose an interval $I_{nk} = (a_{nk}, a'_{nk}]$ such that the corresponding length $a'_{nk}) - a_{nk}$ is bounded by $\epsilon/(2^n m_k)$. Let $I_n$ be the union of the intervals $I_{nk}$, so the total length associated with $I_n$ is $\epsilon/2^n$. For each $x$ in $[p, q]$ let $n_x$ be the first $n$ such that $f_n(x) < \epsilon$. Choose an open interval $V_x$ about $x$ such that $y$ in $V_x$ and $y$ not in $I_{n_x}$ implies $f_{n_x}(y) < \epsilon$. Now let $V_{x_1}, \ldots, V_{x_j}$ be a finite open subcover of $[p, q]$. Let $N$ be the maximum of $n_{x_1}, \ldots, n_{x_j}$. Then since the sequence of functions is monotone decreasing, $y$ in $[p, q]$ but $y$ not in the union of the $I_{n_{x_i}}$ implies $f(y) < \epsilon$. So $\lambda(f_N)$ is bounded by the total length $q - p$ times the upper bound $\epsilon$ on the values plus a length at most $\epsilon$ times the upper bound $M$ on the values. □

## 9.5 Supplement: Monotone convergence without topology

This section presents a proof of the monotone convergence property for the Cantor space (coin tossing space) that does not use topological notions. This is conceptually important, since measure and integral should be a subject that can be developed independent of topology. Once we have the measure on the Cantor space, we can get Lebesgue measure on the unit interval by sending the binary expansion of a real number to the real number. So this gives an approach to the Lebesgue integral that requires no mention of compactness.

Let $\Omega = \{0,1\}^{\mathbf{N}_+}$ be the set of all infinite sequences of zeros and ones indexed by $\mathbf{N}_+ = \{1,2,3,\ldots\}$. For each $k = 0,1,2,3,\ldots$ consider the set $\mathcal{F}_k$ of functions $f$ on $\Omega$ that depend only on the first $k$ elements of the sequence, that is, such that $f(\omega) = g(\omega_1,\ldots,\omega_k)$ for some function $g$ on $\mathbb{R}^k$. This is a vector lattice with dimension $2^k$. The vector lattice under consideration will be the space $L$ that is the union of all the $\mathcal{F}_k$ for $k = 0,1,2,3,\ldots$. In the following, we suppose that we have an elementary integral $\mu$ on $L$.

A subset $A$ of $\Omega$ is said to be an $\mathcal{F}_k$ set when its indicator function $1_A$ is in $\mathcal{F}_k$. In such a case we write $\mu(A)$ for $\mu(1_A)$ and call $\mu(A)$ the measure of $A$. Thus measure is a special case of integral.

In the following we shall need a few simple properties of measure. First, note that $\mu(\emptyset) = 0$. Second, the additivity of the integral implies the corresponding property $\mu(A\cup B)+\mu(A\cap B) = \mu(A)+\mu(B)$. In particular, if $A\cap B = \emptyset$, then $\mu(A\cup B) = \mu(A)+\mu(B)$. This is called the additivity of measure. Finally, the order preserving property implies that $A \subset B$ implies $\mu(A) \leq \mu(B)$.

Here is an example. If the function $f$ is in $\mathcal{F}_k$, let

$$\mu(f) = \sum_{\omega_1=0}^{1}\cdots\sum_{\omega_k=0}^{1} f(\omega)\frac{1}{2^k}. \tag{9.5}$$

This is a consistent definition, since if $f$ is regarded as being in $\mathcal{F}_j$ for $k < j$, then the definition involves sums over $2^j$ sequences, but the numerical factor is $1/2^j$, and the result is the same. This example describes the expectation for independent of tosses of a fair coin. Suppose $A$ is a subset of $\Omega$ whose definition depends only on finitely many coordinates. Then $A$ defines an event that happens or does not happen according to information about finitely many tosses of the coin. The measure $\mu(A) = \mu(1_A)$ is the probability of this event.

The following results shows that such an example automatically satisfies the monotone convergence property and thus gives an elementary integral. The remarkable thing about the proof that follows is that it uses no notions of topology: it is pure measure theory.

**Lemma 9.4** *Suppose that $L$ is a vector lattice consisting of bounded functions. Suppose that 1 is an element of $L$. Suppose furthermore that for each $f$ in $L$ and each real $\alpha$ the indicator function of the set where $f \geq \alpha$ is in $L$. Suppose that $\mu : L \to \mathbb{R}$ is linear and order preserving. If $\mu$ satisfies monotone convergence for sets, then $\mu$ satisfies monotone convergence for functions.*

Proof: Suppose that $\mu$ satisfies monotone convergence for sets, that is, suppose that $A_n \downarrow \emptyset$ implies $\mu(A_n) \downarrow 0$. Suppose that $f_n \downarrow 0$. Say $f_1 \leq M$. Let $\epsilon > 0$. Choose $\alpha > 0$ so that $\alpha\mu(1) < \epsilon/2$. Let $A_n$ be the set where $f_n \geq \alpha > 0$. Then $f_n \leq \alpha + M1_{A_n}$. Hence $\mu(f_n) \leq \alpha\mu(1) + M\mu(A_n)$. Since $A_n \downarrow \emptyset$, we can choose $n$ so that $M\mu(A_n) < \epsilon/2$. Then $\mu(f_n) < \epsilon$. Since $\epsilon > 0$ is arbitrary, this shows that $\mu(f_n) \downarrow 0$. Thus $\mu$ satisfies monotone convergence for functions. □

**Theorem 9.5** *Let $\Omega = \{0,1\}^{\mathbf{N}_+}$ be the set of all infinite sequences of zeros and ones. Let $L = \bigcup_{k=0}^{\infty} \mathcal{F}_k$ be the vector lattice of all functions $f$ on $\Omega$ that each depend only on the first $k$ elements of the sequence for some $k$. Suppose that $\mu : L \to \mathbb{R}$ is linear and order preserving. Then $\mu$ satisfies monotone convergence within $L$.*

Proof: By the lemma, it is enough to show that if $A_n \downarrow \emptyset$ is a sequence of sets, each of which is an $\mathcal{F}_k$ set for some $k$, then $\mu(A_n) \downarrow 0$. The idea is to prove the contrapositive. Suppose then that there is an $\epsilon > 0$ such that $\mu(A_n) \geq \epsilon$ for all $n$.

Let $\bar{\omega}[k] = (\bar{\omega}_1, \ldots, \bar{\omega}_k)$ be a finite sequence of $k$ zeros and ones. Let

$$B_{\bar{\omega}[k]} = \{\omega \mid \omega_1 = \bar{\omega}_1, \ldots, \omega_k = \bar{\omega}_k\} \tag{9.6}$$

This is the binary set of all sequences in $\Omega$ that agree with $\bar{\omega}[k]$ in the first $k$ places. It is an $\mathcal{F}_k$ set. (For $k = 0$ we may regard this as the set of all sequences in $\Omega$.)

The main step in the proof is to show that there is a consistent family of sequences $\bar{\omega}[k]$ such that for each $n$

$$\mu(A_n \cap B_{\bar{\omega}[k]}) \geq \epsilon \frac{1}{2^k}. \tag{9.7}$$

The proof is by induction. The statement is true for $k = 0$. Suppose the statement is true for $k$. By additivity

$$\mu(A_n \cap B_{\bar{\omega}[k]}) = \mu(A_n \cap B_{\bar{\omega}[k]0}) + \mu(A_n \cap B_{\bar{\omega}[k]1}). \tag{9.8}$$

Here $\bar{\omega}[k]0$ is the sequence of length $k+1$ consisting of $\bar{\omega}[k]$ followed by a 0. Similarly, $\bar{\omega}[k]1$ is the sequence of length $k+1$ consisting of $\bar{\omega}[k]$ followed by a 0. Suppose that there is an $n_1$ such that the first term on the right is less than $\epsilon/2^{k+1}$. Suppose also that there is an $n_2$ such that the second term on the right is less than $\epsilon/2^{k+1}$. Then, since the sets are decreasing with $n$, there exists an $n$ such that both terms are less than $\epsilon/2^{k+1}$. But then the measure on the left would be less than $\epsilon/2^k$ for this $n$. This is a contradiction. Thus one of the two suppositions must be false. This says that one can choose $\bar{\omega}[k+1]$ with $\bar{\omega}_{k+1}$ equal to 1 or to 0 so that for all $n$ we have $\mu(A_n \cap B_{\bar{\omega}[k+1]}) \geq \epsilon/2^{k+1}$. This completes the inductive proof of the main step.

The consistent family of finite sequences $\omega[k]$ defines an infinite sequence $\bar{\omega}$. This sequence $\bar{\omega}$ is in each $A_n$. The reason is that for each $n$ there is a $k$ such that $A_n$ is an $\mathcal{F}_k$ set. Each $\mathcal{F}_k$ set is a disjoint union of a collection of binary sets, each of which consists of the set of all sequences where the first $k$ elements have been specified in some way. The set $B_{\bar{\omega}[k]}$ is such a binary set. Hence either $A_n \cap B_{\bar{\omega}[k]} = \emptyset$ or $B_{\bar{\omega}[k]} \subset A_n$. Since $\mu(A_n \cap B_{\bar{\omega}[k]}) > 0$ the first possibility is ruled out. We conclude that

$$\bar{\omega} \in B_{\bar{\omega}[k]} \subset A_n. \tag{9.9}$$

The last argument proves that there is a sequence $\bar{\omega}$ that belongs to each $A_n$. Thus it is false that $A_n \downarrow \emptyset$. This completes the proof of the contrapositive. □

# Problems

1. Lebesgue measure on the Hilbert cube. Consider the Hilbert cube $[0,1]^{\mathbb{N}_+}$. Consider the vector lattice of all real functions $f$ on the Hilbert cube such that there exists $k$ and continuous $F : [0,1]^k \to \mathbb{R}$ with $f(x) = F(x_1, \ldots, x_k)$. Define the elementary integral by
$$\lambda(f) = \int_0^1 \cdots \int_0^1 F(x_1, \ldots, x_k)\, dx_1 \cdots dx_k. \tag{9.10}$$
Prove that in fact it satisfies monotone convergence.

2. Let $f \geq 0$. Prove the most elementary version of the Chebyshev inequality between measure and integral: For each $t > 0$
$$\mu(f \geq t) \leq \frac{\mu(f)}{t}. \tag{9.11}$$

3. Let $g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$.
   (a) Prove that
   $$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(z)g(w)\, dz\; dw = 1. \tag{9.12}$$
   Hint: Polar coordinates.
   (b) Prove the famous Gaussian integral
   $$\int_{-\infty}^{\infty} g(z)\, dz = 1. \tag{9.13}$$
   Hint: Use the previous result.
   (c) Prove that
   $$\int_{-\infty}^{\infty} z^2 g(z)\, dz = 1. \tag{9.14}$$
   Hint: Integrate by parts.
   (d) Evaluate
   $$\int_{-\infty}^{\infty} z^4 g(z)\, dz. \tag{9.15}$$

4. Gaussian measure on $\mathbb{R}^\infty$. Consider functions on $[-\infty, +\infty]^{\mathbf{N}_+}$ that each depend on finitely many coordinates through a continuous function. Define the elementary Gaussian integral as follows. Suppose that $f$ is a function such that there exists $k$ and $F : [-\infty, +\infty]^k \to \mathbb{R}$ such that $f(x) = F(x_1, \ldots, x_k)$. Define
$$\mu(f) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F(z_1, \ldots, z_k)\, g(z_1) \cdots g(z_k)\, dz_1 \cdots dz_k. \tag{9.16}$$
Verify monotone convergence to show that this defines an elementary integral.

5. Prove that this elementary Gaussian integral on functions of $n$ variables for large $n$ is mainly concentrated near the sphere of radius $\sqrt{n}$, in the sense that for each $\epsilon > 0$

$$\mu(|z_1^2 + \cdots + z_n^2 - n| \geq \epsilon n) \leq \frac{2}{\epsilon^2 n}. \tag{9.17}$$

   Hint: $\mu([\sum_{k=1}^n (z_k^2 - 1)]^2 \geq \epsilon^2 n^2) \leq \mu([\sum_{k=1}^n (z_k^2 - 1)]^2)/(\epsilon^2 n^2)$.

6. Take as known that this elementary Gaussian integral extends to an integral. Show that each cube centered at the origin has measure zero. Show that $\ell^\infty$ has measure zero. Show that $\ell^2$ has measure zero.

# Chapter 10

# Existence of integrals

## 10.1 The abstract Lebesgue integral: Daniell construction

The purpose of this chapter is to show that an elementary integral $\mu$ on a Stone vector lattice $L$ of real functions gives rise an integral $\mu$ on a $\sigma$-algebra of real functions. The $\sigma$-algebra $\mathcal{M}_\mu$ produced in the construction is depends on the integral $\mu$. However it includes the $\sigma$-algebra $\sigma(L)$ generated by the original vector lattice, which only depends on $L$.

This section is an outline of the Daniell construction of the abstract Lebesgue integral. This is a two stage process. Let $L\uparrow$ consist of the functions $h : X \to (-\infty, +\infty]$ such that there exists a sequence $h_n$ in $L$ with $h_n \uparrow h$ pointwise. These are the *upper functions*. Similarly, let $L\downarrow$ consist of the functions $f : X \to [-\infty, +\infty)$ such that there exists a sequence $f_n$ in $L$ with $f_n \downarrow f$ pointwise. These are the *lower functions*. The first stage of the construction is to extend the integral to upper functions and to lower functions.

This terminology of upper functions and lower functions is quite natural, but it may not be ideal in all respects. If $L$ is vector lattice of continuous functions, then the upper functions are lower semicontinuous, while the lower functions are upper semicontinuous.

**Lemma 10.1** *There is a unique extension of $\mu$ from $L$ to the upper functions $L\uparrow$ that satisfies the upward monotone convergence property: if $h_n$ is in $L\uparrow$ and $h_n \uparrow h$, then $h$ is in $L\uparrow$ and $\mu(h_n) \uparrow \mu(h)$. Similarly, there is a unique extension of $\mu$ from $L$ to the lower functions $L\downarrow$ that satisfies the corresponding downward monotone convergence property.*

The second stage of the process is to extend the integral to functions that are approximated by upper and lower functions in a suitable sense. Let $g$ be a real function on $X$. Define the *upper integral*

$$\mu^*(g) = \inf\{\mu(h) \mid h \in L\uparrow, g \leq h\}. \tag{10.1}$$

Similarly, define the *lower integral*

$$\mu_*(g) = \sup\{\mu(f) \mid f \in L\downarrow, f \leq g\}. \tag{10.2}$$

**Lemma 10.2** *The upper integral is order preserving and subadditive: $\mu^*(g_1 + g_2) \leq \mu^*(g_1) + \mu^*(g_2)$. Similarly, the lower integral is order preserving and superadditive: $\mu_*(g_1 + g_2) \geq \mu_*(g_1) + \mu_*(g_2)$. Furthermore, $\mu_*(g) \leq \mu^*(g)$ for all $g$.*

Define $\mathcal{L}^1(X, \mu)$ to be the set of all $g : X \to \mathbb{R}$ such that both $\mu_*(g)$ and $\mu^*(g)$ are real, and

$$\mu_*(g) = \mu^*(g). \tag{10.3}$$

Let their common value be denoted $\tilde{\mu}(g)$. This $\tilde{\mu}$ is the integral on the space $\mathcal{L}^1 = \mathcal{L}^1(X, \mu)$ of $\mu$ absolutely integrable functions.

We shall see that this extended integral is in fact an absolute integral. This says that it satisfies the integral bounded monotone convergence closure property. The upward version says that if $f_n$ is a sequence in $\mathcal{L}^1$ and $f_n \uparrow f$ pointwise and the $\tilde{\mu}(f_n)$ are bounded above, then $f$ is in $\mathcal{L}^1$ and $\tilde{\mu}(f_n) \uparrow \tilde{\mu}(f)$. There is a similar downward version. The remarkable thing is that the fact that the limiting function $f$ is in $\mathcal{L}^1$ is not a hypothesis but a conclusion.

**Theorem 10.3 (Daniell construction)** *Let $\mu$ be an elementary integral on a vector lattice $L$ of functions on $X$. Then the corresponding space $\mathcal{L}^1 = \mathcal{L}^1(X, \mu)$ of $\mu$ absolutely integrable functions is a vector lattice, and the extension $\tilde{\mu}$ is an absolute integral on it.*

If an indicator function $1_A$ is in $\mathcal{L}^1$, then $\tilde{\mu}(1_A)$ is written $\tilde{\mu}(A)$ and is called the *measure* of the set $A$. In the following we shall often write the integral of $f$ in $\mathcal{L}^1$ as $\mu(f)$ and the measure of $A$ with $1_A$ in $\mathcal{L}^1$ as $\mu(A)$.

In the following corollary we consider a vector lattice $L$. Let $L\uparrow$ consist of pointwise limits of increasing limits from $L$, and let $L\downarrow$ consist of pointwise limits of decreasing sequences from $L$. Similarly, let $L\uparrow\downarrow$ consist of pointwise limits of decreasing sequences from $L\uparrow$, and let $L\downarrow\uparrow$ consist of pointwise limits of increasing sequences from $L\downarrow$.

**Corollary 10.4** *Let $L$ be a vector lattice and let $\mu$ be an elementary integral. Consider its extension $\tilde{\mu}$ to $\mathcal{L}^1$. Then for every $g$ in $\mathcal{L}^1$ there is a $f$ in $L\downarrow\uparrow$ and an $h$ in $L\uparrow\downarrow$ with $f \leq g \leq h$ and $\tilde{\mu}(g - f) = 0$ and $\tilde{\mu}(h - g) = 0$.*

This corollary says that if we identify functions in $\tilde{L}^1$ when the integral of the absolute value of the difference is zero, then all the functions that we ever will need may be taken, for instance, from $L\uparrow\downarrow$. However this class is not closed under pointwise limits.

The proof of the theorem has a large number of routine verifications. However there are a few key steps. These will be outlined in the following sections. For more detailed accounts there are several excellent references. A classic brief account is Chapter III of the book of Loomis[12]. Chapter VIII of the recent book by Stroock[21] gives a particularly careful presentation. We shall see in the following section that it is possible to extend the integral to much larger $\sigma$-algebras. See Chapter 16 of Royden's book [17] for a related approach to this result.

## 10.2 Stage one

Begin with a vector lattice $L$ and an elementary integral $\mu$. Let $L\uparrow$ be the set of all pointwise limits of increasing sequences of elements of $L$. These functions are allowed to take on the value $+\infty$. Similarly, let $L\downarrow$ be the set of all pointwise limits of decreasing sequences of $L$. These functions are allowed to take on the value $-\infty$. Note that the functions in $L\downarrow$ are the negatives of the functions in $L\uparrow$.

For $h$ in $L\uparrow$, take $h_n \uparrow h$ with $h_n$ in $L$ and define $\mu(h) = \lim_n \mu(h_n)$. The limit of the integral exists because this is a monotone sequence of numbers. Similarly, if $f$ in $L\downarrow$, take $f_n \downarrow f$ with $f_n$ in $L$ and define $\mu(f) = \lim_n \mu(f_n)$.

**Lemma 10.5** *The definition of $\mu(h)$ for $h$ in $L\uparrow$ is independent of the sequence. There is a similar conclusion for $L\downarrow$.*

Proof: Say that $h_m$ is in $L$ with $h_m \uparrow h$ and $k_n$ is in $L$ with $k_n \uparrow k$ and $h \le k$. We will show that $\lim_m \mu(h_m) \le \lim_n \mu(k_n)$. This general fact is enough to establish the uniqueness. In fact, if $h = k$, then we can define $\mu(h)$ by either $\lim_m \mu(h_m)$ or by $\lim_n \mu(k_n)$.

Suppose that $h_m$ is in $L$ with $h_m \uparrow h$ and $k_n$ is in $L$ with $k_n \uparrow k$ and $h \le k$. All we know about $h$ and $k$ are that they are in $L\uparrow$. But

$$h_m \wedge k_n \uparrow h_m \wedge k = h_m \tag{10.4}$$

as $n \to \infty$, and $h_m$ is in $L$. By monotone convergence within $L$ it follows that

$$\mu(h_m \wedge k_n) \uparrow \mu(h_m) \tag{10.5}$$

as $n \to \infty$. But $\mu(h_m \wedge k_n) \le \mu(k_n) \le \mu(k)$. So $\mu(h_m) \le \mu(k)$. Now take $m \to \infty$; it follows that $\mu(h) \le \mu(k)$. □

**Lemma 10.6** *Upward monotone convergence holds for $L\uparrow$. Similarly, downward monotone convergence holds for $L\downarrow$.*

Proof: Here is the argument for upward monotone convergence. Say that the $h_n$ are in $L\uparrow$ and $h_n \uparrow h$ as $n \to \infty$. For each $n$, let $g_{nm}$ be a sequence of functions in $L$ such that $g_{nm} \uparrow h_n$ as $m \to \infty$. The idea is to use this to construct a single sequence $u_n$ of elements of $L$ with $u_n \uparrow h$.

Let $u_n = g_{1n} \vee g_{2n} \vee \cdots \vee g_{nn}$. Then $u_n$ is in $L$ and $u_n \le h_n \le h$, and so $u_n \uparrow u$ for some $u$ in $L^\uparrow$. There is a squeeze inequality

$$g_{in} \le u_n \le h_n \tag{10.6}$$

for $1 \leq i \leq n$. As $n \to \infty$ the $g_{in} \uparrow h_i$ and the $h_n \uparrow h$. So $h_i \leq u \leq h$. Furthermore, as $i \to \infty$ the $h_i \uparrow h$. By the squeeze inequality $h \leq u \leq h$, that is, $u_n \uparrow h$.

Again from the squeeze inequality we get

$$\mu(g_{in}) \leq \mu(u_n) \leq \mu(h_n) \tag{10.7}$$

for $1 \leq i \leq n$. Since $u_n \uparrow h$ with each $u_n$ in $L$, the preceding lemma gives $\mu(u_n) \uparrow \mu(h)$ as $n \to \infty$. So $\mu(h_i) \leq \mu(h) \leq \lim_n \mu(h_n)$. Then we can take $i \to \infty$ and get $\lim_i \mu(h_i) \leq \mu(h) \leq \lim_n \mu(h_n)$. This shows that the integrals converge to the correct value. □

## 10.3 Stage two

The integral $\mu(g)$ is the supremum of all the $\mu(f)$ for $f$ in $L \downarrow$ with $f \leq g$ and is also the infimum of all the $\mu(h)$ for $h$ in $L \uparrow$ with $g \leq h$. Alternatively, a function $g$ is in $\mathcal{L}^1$ if for every $\epsilon > 0$ there is a function $f$ in $L \downarrow$ and a function $h$ in $L \uparrow$ such that $f \leq g \leq h$, $\mu(f)$ and $\mu(h)$ are finite, and $\mu(h) - \mu(f) < \epsilon$.

It is not hard to show that the set $\mathcal{L}^1$ of absolutely integrable functions is a vector lattice and that $\mu$ is a positive linear functional on it. The crucial point is that there is also a monotone convergence theorem. This theorem says that if the $g_n$ are absolutely integrable functions with $\mu(g_n) \leq M < \infty$ and if $g_n \uparrow g$, then $g$ is absolutely integrable with $\mu(g_n) \uparrow \mu(g)$.

**Lemma 10.7** *The integral on $\mathcal{L}^1$ satisfies the monotone convergence property.*

Proof: We may suppose that $g_0 = 0$. Let $w_n = g_n - g_{n-1} \geq 0$ for $n \geq 1$ be the increment. Since the absolutely integrable functions $\mathcal{L}^1$ are a vector space, each $w_n$ is absolutely integrable. So

$$g_n = \sum_{i=1}^{n} w_i \tag{10.8}$$

is a sum of positive absolutely integrable functions.

Consider $\epsilon > 0$. Each $w_i$ may be approximated above by some $h_i$ in $L \uparrow$. In fact, we may choose $h_i$ in $L \uparrow$ for $i \geq 1$ such that $w_i \leq h_i$ and such that

$$\mu(h_i) \leq \mu(w_i) + \frac{\epsilon}{2^i}. \tag{10.9}$$

Let

$$s_n = \sum_{i=1}^{n} h_i \tag{10.10}$$

be the corresponding sum of functions in $L \uparrow$. Then $g_n \leq s_n$. Also $s_n \uparrow s$ in $L \uparrow$, and $g \leq s$. This is the $s$ in $\uparrow$ that we want to use to approximate $g$ from above.

To deal with the integrals, note that

$$\mu(s_n) \leq \mu(g_n) + \epsilon \leq M + \epsilon. \tag{10.11}$$

By monotone convergence for $L \uparrow$

$$\mu(s) \leq \lim_n \mu(g_n) + \epsilon \leq M + \epsilon. \tag{10.12}$$

Pick $m$ so large that $g_m \leq g$ satisfies $\mu(s) < \mu(g_m) + \frac{3}{2}\epsilon$. Use the fact that this $g_m$ may be approximated from below by some $r$ in $L \downarrow$. Thus pick $r$ in $L \downarrow$ with $r \leq g_m$ so that $\mu(g_m) \leq \mu(r) + \frac{1}{2}\epsilon$. Then $r \leq g \leq s$ with $\mu(s) - \mu(r) < 2\epsilon$. Since $\epsilon$ is arbitrary, this proves that $g$ is absolutely integrable.

Since $g_n \leq g$, it is clear that $\lim_n \mu(g_n) \leq \mu(g)$. On the other hand, the argument has shown that for each $\epsilon > 0$ we can find $s$ in $L \uparrow$ with $g \leq s$ and $\mu(g) \leq \mu(s) \leq \lim_n \mu(g_n) + \epsilon$. Since $\epsilon$ is arbitrary, we conclude that $\mu(g) \leq \lim_n \mu(g_n)$. This proves that $\mu(g_n) \uparrow \mu(g)$. □

The proof of the monotone convergence theorem for the functions in $L \uparrow$ and for the functions in $L \downarrow$ is natural within the context of ordered steps. However the proof of the monotone convergence theorem for the functions in $\mathcal{L}^1$ has a remarkable and deep feature: it uses in a critical way the fact that the sequence of functions is indexed by a countable set of $n$. Thus the errors in the approximations can be estimated by $\epsilon/2^n$, and these sum to the finite value $\epsilon$.

Technically a function in $L \uparrow$ or $L \downarrow$ with finite integral need not be in $\mathcal{L}^1$, because of the possible of infinite values ($+\infty$ in the case of $L \uparrow$, and $-\infty$ in the case of $L \downarrow$. This is because according to our definition functions in $\mathcal{L}^1$ have only real values. However in later developments we shall see that this technicality does not present serious problems.

## 10.4 Extension to measurable functions

**Proposition 10.8** *If $L$ is a Stone vector lattice, then the corresponding $\mathcal{L}^1$ in the Daniell construction is also a Stone vector lattice.*

Proof: Suppose that $L$ is a Stone vector lattice. Thus $f \in L$ implies $f \wedge 1 \in L$. It follows by taking monotone increasing limits that $f \in L \uparrow \cap \mathcal{L}$ implies $f \wedge 1$ in $L^{\uparrow} \cap \mathcal{L}$. Then it follows by taking monotone decreasing limits that $f \in L \uparrow\downarrow \cap\mathcal{L}$ implies $f \wedge 1 \in L \uparrow\downarrow \cap\mathcal{L}$. Now consider $g \in \mathcal{L}^1$. Find $f$ in $L^{\uparrow} \downarrow \cap\mathcal{L}^1$ with $g \leq f$ and $\mu(f - g) = 0$. Then $g \leq f$ implies $g \wedge 1 \leq f \wedge 1$. Furthermore, $0 \leq f \wedge 1 - g \wedge 1 \leq f - g$. It follows that $f \wedge 1 - g \wedge 1$ is in $\mathcal{L}^1$ with integral zero. The conclusion is that $g \wedge 1$ is in $\mathcal{L}^1$. So $\mathcal{L}^1$ is a Stone vector lattice. □

Let $\mu : \mathcal{L} \to \mathbb{R}$ be an elementary integral defined on a vector lattice $\mathcal{L}$. By definition of elementary integral, if $f_n$ is in $\mathcal{L}$ with $f_n \uparrow f$ pointwise, and if $f$ is assumed to be in $\mathcal{L}$, then $\mu(f_n) \to \mu(f)$. The monotone convergence closure property is much more powerful: it says that if $f_n$ is in $\mathcal{L}$ with $f_n \uparrow f$ pointwise, and if the $\mu(f_n)$ are bounded above, then $f$ is in $\mathcal{L}$ and $\mu(f_n) \to \mu(f)$.

Recall that an absolute integral is an integral on a vector lattice that satisfies the integral bounded monotone convergence closure property.

**Theorem 10.9** *Let $\mu : \mathcal{L} \to \mathbb{R}$ be an absolute integral defined on a Stone vector lattice $\mathcal{L}$. Then there exists a $\sigma$-algebra $\mathcal{M}$ of real functions with $\mathcal{L} \subset \mathcal{M}$ and such that $\mu$ extends to an integral on $\mathcal{M}^+$.*

Proof: For technical reasons it is best to construct first a $\sigma$-algebra of subsets $\mathcal{M}_X$. A subset $E$ of $X$ is in $\mathcal{M}_X$ provided that it is locally measurable: for each $A$ with $1_A$ in $\mathcal{L}$ it is also the case that $1_E \wedge 1_A = 1_{E\cap A}$ is in $\mathcal{L}$.

It is routine to verify that $\mathcal{M}$ is indeed an algebra of subsets. To check countable additivity, it is enough to check that if $E_n$ is in $\mathcal{M}_X$ and $E_n \uparrow E$, then $E$ is in $\mathcal{M}_X$. Suppose that each $E_n \in \mathcal{M}_X$. Consider $1_A$ in $\mathcal{L}$. Then $1_{E_n\cap A} \uparrow 1_{E\cap A}$. Since the integral of $1_{E_n\cap A}$ is bounded by the integral of $1_A$, we can conclude from the monotone convergence closure property that $1_{E\cap A}$ is indeed in $\mathcal{L}$.

A $\sigma$-algebra of subsets $\mathcal{M}_X$ always determines a $\sigma$-algebra $\mathcal{M}$ of real functions. A function $f$ is in $\mathcal{M}$ provided that for every real $a$ the set $f > a$ is in $\mathcal{M}_X$.

Next we need to establish that each $f$ in $\mathcal{L}$ is also in $\mathcal{M}$. Since every function $f$ in $\mathcal{L}$ has a decomposition $f = f_+ - f_-$ into a positive and a negative part, it is enough to verify this for positive functions. So consider a function $f \geq 0$ in $\mathcal{L}$. Consider $a \geq 0$. Since $\mathcal{L}$ is a Stone vector lattice, the function $f \wedge a$ is in $\mathcal{L}$. Hence also $f - f \wedge a$ is in $\mathcal{L}$. Consider a subset $A$ with $1_A$ in $\mathcal{L}$. Then $n(f - f \wedge a) \wedge 1_A$ is in $\mathcal{L}$. These functions all have integral bounded by the integral of $1_A$. As $n \to \infty$ they converge to $1_{f>a} \wedge 1_A$. So from the monotone convergence property this is a function in $\mathcal{L}$. This shows that the set where $f > a$ is in $\mathcal{M}_X$.

Finally, we need to extend $\mu$ to all of $\mathcal{M}_+$. One method is to define $\mu(f) = +\infty$ if $f \geq 0$ is in $\mathcal{M}^+$ but not in $\mathcal{L}$. □

**Proposition 10.10** *The elements $f \geq 0$ of $\mathcal{M}^+$ in the construction are characterized by the property that $f \in \mathcal{M}^+$ if and only if $g \in \mathcal{L}$ implies $f \wedge g \in \mathcal{L}$.*

Proof: To say that $f \geq 0$ is in $\mathcal{M}^+$ is to say that it is measurable. Clearly if it is measurable and $g$ is in $\mathcal{L}^1$, then $f \wedge g$ is measurable, and it follows easily that $f \wedge g$ is in $\mathcal{L}^1$. For the other direction, consider $f \geq 0$. Suppose that $g \in \mathcal{L}$ implies $f \wedge g \in \mathcal{L}$. Consider $c > 0$. Take $g = c1_A$ where $1_A$ is in $\mathcal{L}^1$. Then $f \wedge c1_A$ is in $\mathcal{L}$ and hence is in $\mathcal{M}$. It follows that the set where $f \wedge c1_A \geq c$ is measurable. But this is the intersection of the set where $f \geq c$ with $A$. It follows that the set where $f \geq c$ is a measurable set. Thus $f$ is a measurable function. □

Remark: The initial construction starts with an elementary integral and produces an absolute integral on a very large domain. The present construction then produces a very large $\sigma$-algebra $\mathcal{M}$ of locally measurable functions. In the following we shall often write it as $\mathcal{M}_\mu$, in order to emphasize that this $\sigma$-algebra of real functions depends on the integral $\mu$. In the following we shall most often

restrict the integral to $\mathcal{F} = \sigma(L)$, the $\sigma$-algebra generated by $L$. This is much smaller, but it already large enough for just about every practical purpose.
Example: Let $L$ consists of all continuous real functions with compact support on the line. Take the elementary integral $\lambda$ to be the Riemann integral. The $\sigma$-algebra $\mathcal{M}_\lambda$ constructed in this theorem is known as the $\sigma$-algebra of Lebesgue measurable functions. It is huge: much larger than the Borel $\sigma$-algebra $\sigma(L) = \mathcal{B}o$ generated by $L$. However there is no harm in most instances in restricting the integral to the Borel functions; there are already plenty of these.

## 10.5 Example: The Lebesgue integral

Consider the vector lattice $L$ of real step functions on the line $\mathbb{R}$. The integral $\lambda$ of such a function is given by a finite sum. By Dini's theorem for step functions this satisfies monotone convergence and hence is an elementary integral. Therefore it has an extension to an integral $\lambda$ defined for the $\sigma$-algebra $\mathcal{M}_\lambda$ of Lebesgue measurable functions. In most of the following we shall find it sufficient to regard this integral as defined for the smaller $\sigma$-algebra $\mathcal{B}o$ of Borel measurable functions. In either case, the integral $\lambda$ is written

$$\lambda(f) = \int_{-\infty}^{\infty} f(x)\,dx \tag{10.13}$$

and is called the *Lebesgue integral*. The associated measure defined on Borel subsets of the line is called *Lebesgue measure*.

One can show directly from the definition of the integral that that the Lebesgue measure of a countable set $Q$ is 0. This will involve a two-stage process. Let $q_j$, $j = 1, 2, 3, \ldots$ be an enumeration of the points in $Q$. Fix $\epsilon > 0$. For each $j$, find a n interval $B_j$ of length less than $\epsilon/2^j$ such that $q_j$ is in the interval. The indicator function $1_{B_j}$ of each such interval is in $L$. Let $h = \sum_j 1_{B_j}$. Then $h$ is in $L \uparrow$ and $\lambda(h) \leq \epsilon$. Furthermore, $0 \leq 1_Q \leq h$. This is the first stage of the approximation. Now consider a sequence of $\epsilon > 0$ values that approach zero, and construct in the same way a sequence of $h_\epsilon$ such that $0 \leq 1_Q \leq h_\epsilon$ and $\lambda(h_\epsilon) \leq \epsilon$. This is the second stage of the approximation. This shows that the integral of $1_Q$ is zero.

Notice that this could not have been done in one stage. There is no way to cover $Q$ by finitely many binary intervals of small total length. It was necessary first to find infinitely many binary intervals that cover $Q$ and have small total length, and only then let this length approach zero.

## 10.6 Example: The expectation for coin tossing

An example to which this result applies is the space $\Omega$ of the coin tossing example. Recall that the elementary integral is defined on the space $L = \bigcup_{n=0}^{\infty} \mathcal{F}_n$, where $\mathcal{F}_n$ consists of the functions that depend only on the first $n$ coordinates.

Thus $L$ consists of functions each of which depends only on finitely many coordinates. A subset $S$ of $\Omega$ is said to be an $\mathcal{F}_n$ set if its indicator function $1_S$ belongs to $\mathcal{F}_n$. This just means that the definition of the set depends only on the first $n$ coordinates. In the same way, $S$ is said to be an $L$ set if $1_S$ is in $L$.

Consider the elementary integral for fair coin tossing. The elementary integral $\mu(f)$ of a function $f$ in $\mathcal{F}_n$ may be calculated by a finite sum involving at most $2^n$ terms. It is just the sum of the values of the function for all of the $2^n$ possibilities for the first $n$ coin flips, divided by $2^n$. Similarly, the elementary measure $\mu(S)$ of an $\mathcal{F}_n$ set is the number among the $2^n$ possibilities of the first $n$ coin flips that are satisfied by $S$, again weighted by $1/2^n$.

Thus consider for example the measure of the uncountable set $S$ consisting of all $\omega$ such that $\omega_1 + \omega_2 + \omega_3 = 2$. If we think of $S$ as an $\mathcal{F}_3$ set, its measure is $3/2^3 = 3/8$. If we think of $S$ as an $\mathcal{F}_4$ set, its measure is still $6/2^4 = 3/8$.

The elementary integral on $L$ extends to an integral. The integral of a function $f$ is denoted $\mu(f)$. This is interpreted as the expectation of the random variable $f$. Consider a subset $S$ of $\Omega$. The measure $\mu(S)$ of $S$ is the integral $\mu(1_S)$ of its indicator function $1_S$. This is interpreted as the probability of the event $S$ in the coin tossing experiment.

**Proposition 10.11** *Consider the space $\Omega$ for infinitely many tosses of a coin, and the associated integral for tosses of a fair coin. Then each subset with exactly one point has measure zero.*

Proof: Consider such a set $\{\bar{\omega}\}$. Let $B_k$ be the set of all $\omega$ in $\Omega$ such that $\omega$ agrees with $\bar{\omega}$ in the first $k$ places. The indicator function of $B_k$ is in $L$. Since $\{\bar{\omega}\} \subset B_k$, we have $0 \leq \mu(\{\bar{\omega}\}) \leq \mu(B_k) = 1/2^k$ for each $k$. Hence $\mu(\{\bar{\omega}\}) = 0$. $\square$

**Proposition 10.12** *Consider the space $\Omega$ for infinitely many tosses of a coin, and the associated integral that gives the expectation for tosses of a fair coin. Let $S \subset \Omega$ be a countable subset. Then the measure of $S$ is zero.*

Proof: Here is a proof from the definition of the integral. Let $j \mapsto \omega^{(j)}$ be an enumeration of $S$. Let $\epsilon > 0$. For each $j$ let $B^{(j)}$ be a set with indicator function in $L$ such that $\omega^{(j)} \in B^{(j)}$ and $\mu(B^{(j)}) < \frac{\epsilon}{2^j}$. For instance, one can take $B^{(j)}$ to be the set of all $\omega$ that agree with $\omega^{(j)}$ in the first $k$ places, where $1/2^k \leq \epsilon/2^j$. Then

$$0 \leq 1_S \leq 1_{\bigcup_j B^{(j)}} \leq \sum_j 1_{B^{(j)}}. \tag{10.14}$$

The right hand side of this equation is in $L^\uparrow$ and has integral bounded by $\epsilon$. Hence $0 \leq \mu(S) \leq \epsilon$. It follows that $\mu(S) = 0$. $\square$

Proof: Here is a proof from the monotone convergence theorem. Let $j \mapsto \omega^{(j)}$ be an enumeration of $S$. Then

$$\sum_j 1_{\omega^{(j)}} = 1_S. \tag{10.15}$$

By the previous proposition each term in the integral has integral zero. Hence each partial sum has integral zero. By the monotone convergence theorem the sum has integral zero. Hence $\mu(S) = 0$. □

**Corollary 10.13** *Consider the space $\Omega$ for infinitely many tosses of a coin, and the associated integral that gives the expectation for tosses of a fair coin. Let $S \subset \Omega$ be the set of all sequences that are eventually either all zeros or all ones. Then the measure of $S$ is zero.*

Examples:

1. As a first practical example, consider the function $b_j$ on $\Omega$ defined by $b_j(\omega) = \omega_j$, for $j \geq 1$. This scores one for a success in the $j$th trial. It is clear that $b_j$ is in $\mathcal{F}_j$ and hence in $L$. It is easy to compute that $\mu(b_j) = 1/2$ for the fair coin $\mu$.

2. A more interesting example is $c_n = b_1 + \cdots + b_n$, for $n \geq 0$. This random variable counts the number of successes in the first $n$ trials. It is a function in $\mathcal{F}_n$ and hence in $L$. The fair coin expectation of $c_n$ is $n/2$. In $n$ coin tosses the expected number of successes is $n/2$.

3. Consider the set defined by the condition $c_n = k$ for $0 \leq k \leq n$. This is an $\mathcal{F}_n$ set, and its probability is $\mu(c_n = k) = \binom{n}{k} 1/2^n$. This is the famous binomial probability formula. These probabilities add to one:

$$\sum_{k=0}^{n} \binom{n}{k} \frac{1}{2^n} = 1. \tag{10.16}$$

   This formula has a combinatorial interpretation: the total number of subsets of an $n$ element set is $2^n$. However the number of subsets with $k$ elements is $\binom{n}{k}$. The formula for the expectation of $c_n$ gives another identity:

$$\sum_{k=0}^{n} k \binom{n}{k} \frac{1}{2^n} = \frac{1}{2} n. \tag{10.17}$$

   This also has a combinatorial interpretation: the total number of ordered pairs consisting of a subset and a point within it is the same as the number of ordered pairs consisting of a point and a subset of the complement, that is, $n 2^{n-1}$. However the number of ordered pairs consisting of a $k$ element set and a point within it is $\binom{n}{k} k$.

4. Let $u_1(\omega)$ be the first $k$ such that $\omega_k = 1$. This waiting time random variable is not in $L$, but for each $m$ with $1 \leq m < \infty$ the event $u_1 = m$ is an $\mathcal{F}^m$ set and hence an $L$ set. The probability of $u_1 = m$ is $1/2^m$. The event $u_1 = \infty$ is not an $L$ set, but it is a one point set, so it has zero probability. This is consistent with the fact that the sum of the probabilities is a geometric series with $\sum_{m=1}^{\infty} 1/2^m = 1$.

5. The random variable $u_1 = \sum_{m=1}^{\infty} m 1_{u_1=m}$ is in $L\uparrow$. Its expectation is $\mu(u_1) = \sum_{m=1}^{\infty} m/2^m = 2$. This says that the expected waiting time to get a success is two tosses.

6. Let $t_n(\omega)$ for $n \geq 0$ be the $n$th value of $k$ such that $\omega_k = 1$. (Thus $t_0 = 0$ and $t_1 = u_1$.) Look at the event that $t_n = k$ for $1 \leq k \leq n$, which is an $\mathcal{F}_k$ set. This is the same as the event $c_{k-1} = n-1, b_k = 1$ and so has probability $\binom{k-1}{n-1} 1/2^{k-1} 1/2 = \binom{k-1}{n-1} 1/2^k$. These probabilities add to one, but this is already not such an elementary fact. However the event $t_n = \infty$ is a countable set and thus has probability zero. So in fact
$$\sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{2^k} = 1. \tag{10.18}$$
This is an infinite series; a combinatorial interpretation is not apparent.

7. For $n \geq 1$ let $u_n = t_n - t_{n-1}$ be the $n$th waiting time. It is not hard to show that the event $t_{n-1} = k, u_n = m$ has probability $\mu(t_{n-1} = k) 1/2^m$, and hence that the event $u_n = m$ has probability $1/2^m$. So $u_n$ also is a geometric waiting time random variable, just like $u_1$. In particular, it has expectation 2.

8. We have $t_n = u_1 + \cdots + u_n$. Hence the expectation $\mu(t_n) = 2n$. The expected total time to wait until the $n$th success is $2n$. This gives another remarkable identity
$$\sum_{k=n}^{\infty} k \binom{k-1}{n-1} \frac{1}{2^k} = 2n. \tag{10.19}$$
It would not make much sense without the probability intuition.

# Problems

1. Let $k \to r_k$ be an enumeration of the rational points in $[0,1]$. Define $g(x) = \sum_k 2^k 1_{\{r_k\}}(x)$. Evaluate the Lebesgue integral of $g$ directly from the definition in terms of integrals of step functions, integrals of lower and upper functions, and integrals of functions squeezed between lower and upper functions.

2. The Cantor set $C$ is the subset of $[0,1]$ that is the image of $\Omega = \{0,1\}^{\mathbf{N}_+}$ under the injection
$$c(\omega) = \sum_{n=1}^{\infty} \frac{2\omega_n}{3^n}. \tag{10.20}$$
The complement of the Cantor set in $[0,1]$ is an open set obtained by removing middle thirds. Show that the indicator function of the complement of the Cantor set is a function in $L\uparrow$. Find the Lebesgue measure of the complement of the Cantor set directly from the definition. Then find the Lebesgue measure of the Cantor set.

3. Let $c$ be the cardinality of the continuum. Show that the cardinality of the set of all real functions on $[0, 1]$ is $c^c$. Show that $c^c = 2^c$.

4. Show that the cardinality of the set of real functions on $[0, 1]$ with finite Lebesgue integral is $2^c$. Hint: Think about the Cantor set.

5. The Lebesgue integral may be defined starting with the elementary integral $\lambda$ defined on $L = C([0, 1])$. Show that $L \uparrow$ consists of lower semicontinuous functions, and $L \downarrow$ consists of upper semicontinuous functions.

# Chapter 11

# Uniqueness of integrals

## 11.1 $\sigma$-rings

We have seen that there is a correspondence between the notions of $\sigma$-algebra of subsets and $\sigma$-algebra of real functions. The purpose of this section is to introduce slightly more general concepts, with ring replacing algebra. Thus there will be a correspondence between the notions of $\sigma$-ring of subsets and $\sigma$-algebra of real functions. It is possible to carry out the entire theory of measure and integration in this more general $\sigma$-ring context. However, the only use we shall have for this concept is for uniqueness results. Thus only some main fact are stated; details are in Halmos [7]. For convenient the ring and algebra cases are presented in parallel.

Let $X$ be a set. A *ring of subsets* is a collection of subsets $\mathcal{R}$ such that the empty set is in $\mathcal{R}$ and such that $\mathcal{R}$ is closed under the operations of finite union and relative complement.

A ring of sets $\mathcal{A}$ is an *algebra of subsets* if in addition the set $X$ belongs to $\mathcal{A}$. Thus the empty set belongs to $\mathcal{A}$ and it is closed under the operations of finite union and complement. To get from a ring of sets to an algebra of sets, it is enough to put in the complements of the sets in the ring.

An example of a ring of sets is the ring $\mathcal{R}$ of subsets of $\mathbb{R}$ generated by the intervals $(a, b]$ with $a < b$. This consists of the collection of sets that are finite unions of such intervals. Another example is the ring $\mathcal{R}_0$ of sets generated by the intervals $(a, b]$ such that either $a < b < 0$ or $0 < a < b$. None of the sets in this ring have the number 0 as a member.

Recall the Stone condition: If $f$ is in the vector lattice, then so is $f \wedge 1$. This does not require that 1 is in the vector lattice. However, if 1 is in the vector lattice, then it is automatically a Stone vector lattice.

**Proposition 11.1** *Let $\mathcal{R}$ be a ring of sets. Then the set of finite linear combinations of indicator functions $1_A$ with $A$ in $\mathcal{R}$ is a Stone vector lattice.*

**Proposition 11.2** *Let $\mathcal{A}$ be a algebra of sets. Then the set of finite linear* *combinations of indicator functions $1_A$ with $A$ in $\mathcal{A}$ is a vector lattice including the constant functions.*

A ring of sets is a *$\sigma$-ring of subsets* if it is closed under countable unions. Similarly, an algebra of sets is a *$\sigma$-algebra of subsets* if it is closed under countable unions.

An example of a $\sigma$-ring of sets that is not a $\sigma$-algebra of sets is the set of all countable subsets of an uncountable set $X$. The smallest $\sigma$-algebra including this $\sigma$-ring consists of all subsets that are either countable or have countable complement.

A standard example of a $\sigma$-algebra of sets is the Borel $\sigma$-algebra $\mathcal{B}o$ of subsets of $\mathbf{R}$ generated by the intervals $(a, +\infty)$ with $a \in \mathbf{R}$. A corresponding standard example of a $\sigma$-ring that is not a $\sigma$-algebra is the $\sigma$-ring $\mathcal{B}o_0$ consisting of all Borel sets $A$ such that $0 \notin A$.

A *$\sigma$-ring of real functions* $\mathcal{F}_0$ is a vector lattice that is closed under monotone convergence and that is also a Stone vector lattice.

A *$\sigma$-algebra of real functions* $\mathcal{F}$ on $X$ is a vector lattice that is closed under monotone convergence and that includes the constant functions.

Every $\sigma$-algebra of functions is a $\sigma$-ring of functions. A simple example of a $\sigma$-ring of functions that is not a $\sigma$-algebra of functions is given by the set of all real functions on $X$ that are each non-zero on a countable set. If $X$ is uncountable, then the constant functions do not belong to this $\sigma$-ring.

A $\sigma$-ring of functions or a $\sigma$-algebra of functions is automatically closed not only under the vector space and lattice operations, but also under pointwise multiplication. In addition, there is closure under pointwise limits (not necessarily monotone).

**Proposition 11.3** *Let $\mathcal{F}_0$ be a $\sigma$-ring of real functions on $X$. Then the sets $A$ such that $1_A$ are in $\mathcal{F}_0$ form a $\sigma$-ring $\mathcal{R}_0$ of subsets of $X$.*

**Proposition 11.4** *Let $\mathcal{F}$ be a $\sigma$-algebra of real functions on $X$. Then the sets $A$ such that $1_A$ are in $\mathcal{F}$ form a $\sigma$-algebra $\mathcal{R}$ of subsets of $X$.*

Let $\mathcal{R}_0$ be a $\sigma$-ring of subsets of $X$. Let $f : X \to \mathbb{R}$ be a function. Then $f$ is said to be *measurable* with respect to $\mathcal{R}_0$ if for each $B$ in $\mathcal{B}_0$ the inverse image $f^{-1}[B]$ is in $\mathcal{R}_0$.

Similarly, let $\mathcal{R}$ be a $\sigma$-algebra of subsets of $X$. Let $f : X \to \mathbb{R}$ be a function. Then $f$ is said to be *measurable* with respect to $\mathcal{R}$ if for each $B$ in $\mathcal{B}o$ the inverse image $f^{-1}[B]$ is in $\mathcal{R}$.

To check that a function is measurable, it is enough to check the inverse image property with respect to a generating class. For $\mathcal{B}o$ this could consist of the intervals $(a, +\infty)$ where $a$ is in $\mathbb{R}$. Thus to prove a function $f$ is measurable with respect to a $\sigma$-algebra $\mathcal{R}$, it would be enough to show that for each real $a$ the set where $f > a$ is in $\mathcal{R}$. For $\mathcal{B}_0$ a generating class could consist of the intervals $(a, +\infty)$ with $a > 0$ together with the intervals $(-\infty, a)$ with $a < 0$.

**Proposition 11.5** *Let $\mathcal{R}_0$ be a $\sigma$-ring of subsets of $X$. Then the collection $\mathcal{F}_0$ of real functions on $X$ that are measurable with respect to $\mathcal{R}_0$ is a $\sigma$-ring of functions on $X$.*

**Proposition 11.6** *Let $\mathcal{R}$ be a $\sigma$-algebra of subsets of $X$. Then the collection $\mathcal{F}_0$ of real functions on $X$ that are measurable with respect to $\mathcal{R}$ is a $\sigma$-algebra of functions on $X$.*

## 11.2 The uniqueness theorem

If $S$ is a set of functions, then the smallest $\sigma$-algebra including $S$ is denoted $\sigma(S)$. Correspondingly, the smallest $\sigma$-ring including $S$ is denoted $\sigma_0(S)$. (This may not be standard notation, but it seems reasonable.)

**Theorem 11.7** *Let $L$ be a Stone vector lattice. Let $m$ be an elementary integral on $L$. Let $\mathcal{F}_0 = \sigma_0(L)$ be the $\sigma$-ring of functions generated by $L$. Then the extension $\mu$ of $m$ to $\mathcal{F}_0^+$ is unique.*

The proof of this theorem is presented in a later section of this chapter.

**Corollary 11.8** *Let $L$ be a Stone vector lattice. Let $m$ be an elementary integral on $L$. Let $\mathcal{F} = \sigma(L)$ be the $\sigma$-algebra generated by $L$. Suppose that the $\sigma$-ring $\mathcal{F}_0 = \sigma_0(L)$ of functions generated by $L$ contains the constant functions, so that $\mathcal{F}_0 = \mathcal{F}$. Then the extension $\mu$ of $m$ to an integral on $\mathcal{F}^+$ is unique.*

This corollary applies in many examples. If 1 is in $L$ there is of course no problem. However even if 1 is not in $L$, it may be a pointwise limit of functions in $L$. This is the case, for example, for the real continuous functions on $\mathbb{R}$, each with compact support. It is also the case for the real step functions on $\mathbb{R}$.

This integral is not unique in every case. A trivial example is to take $L$ to consist only of the zero function, and $m$ the elementary integral that assigns the number zero to this function. Then for each $c$ with $0 \leq c \leq +\infty$ there is an integral defined by $\mu(a) = ca$.

This example might seem trivial, since the functions in $L$ do not separate points. However another example is to take $L$ to be all functions defined on a fixed uncountable set, each function having finite support. Again take the elementary integral to be zero for each of these functions. Then $\mathcal{F}_0$ consists of all functions with countable support. Each of these functions has integral zero. Again for each $c$ with $0 \leq c \leq +\infty$ there is an integral defined by $\mu(a) = ca$.

## 11.3 $\sigma$-finite integrals

An integral is *$\sigma$-finite* if there is a sequence $0 \leq u_n \uparrow 1$ of measurable functions with each $\mu(u_n) < +\infty$. If this is the case, define $E_n$ as the set where $u_n \geq 1/2$. By Chebyshev's inequality the measure $\mu(E_n) \leq 2\mu(u_n) < +\infty$. Furthermore, $E_n \uparrow X$ as $n \to \infty$. Suppose on the other hand that there exists an increasing sequence $E_n$ of measurable subsets of $X$ such that each $\mu(E_n) < +\infty$ and $X = \bigcup_n E_n$. Then it is not difficult to show that $\mu$ is $\sigma$-finite. In fact, it suffices to take $u_n$ to be the indicator function of $E_n$.

**Theorem 11.9** *Let $\mathcal{F}$ be a $\sigma$-algebra of real functions on $X$. Let $\mu : \mathcal{F}^+ \to [0, +\infty]$ be an integral. Then $\mu$ is $\sigma$-finite if and only if there exists a Stone vector lattice $L$ such that the restriction of $\mu$ to $L$ has only finite values and such that the smallest $\sigma$-ring including $L$ is $\mathcal{F}$.*

The proof of this theorem is presented in a later section of this chapter.

## 11.4 Summation

The familiar operation of *summation* is a special case of integration. Let $X$ be a set. Then there is an integral $\sum : [0, +\infty)^X \to [0, +\infty]$. It is defined for $f \geq 0$ by $\sum f = \sup_W \sum_{j \in W} f(j)$, where the supremum is over all finite subsets $W \subset X$. Since each $f(j) \geq 0$, the result is a number in $[0, +\infty]$. As usual, the sum is also defined for functions that are not positive, but only provided that there is no $(+\infty) - (+\infty)$ problem.

Suppose $f \geq 0$ and $\sum f < +\infty$. Let $S_k$ be the set of $j$ in $X$ such that $f(j) \geq 1/k$. Then $S_k$ is a finite set. Let $S$ be the set of $j$ in $X$ such that $f(j) > 0$. Then $S = \bigcup_k S_k$, so $S$ is countable. This argument proves that the sum is infinite unless $f$ vanishes off a countable set. So a finite sum is just the usual sum over a countable index set.

The $\sum$ integral is $\sigma$-finite if and only if $X$ is countable. This is because whenever $f \geq 0$ and $\sum f < +\infty$, then $f$ vanishes off a countable set $S$. So if each $f_n$ vanishes off a countable set $S_n$, and $f_n \uparrow f$, then $f$ vanishes off $S = \bigcup S_n$, which is also a countable set. This shows that $f$ cannot be a constant function $a > 0$ unless $X$ is a countable set.

One could define $\sum$ on a smaller $\sigma$-algebra of functions. The smallest one that seems natural consists of all functions of the form $f = g + a$, where the function $g$ is zero on the complement of some countable subset of $X$, and $a$ is constant. If $f \geq 0$ and $a = 0$, then $\sum f = \sum g$ is a countable sum. On the other hand, if $f \geq 0$ and $a > 0$ then $\sum f = +\infty$.

One can also look at summation from the measure point of view. The sum of an indicator function just counts the points in the associated subset. So in this perspective the measure is called *counting measure*.

## 11.5 Regularity

Recall that if $L$ is a vector lattice of real functions, then the upper functions $L\uparrow$ consist of the increasing limits of sequences functions in $L$, and the lower functions $L\downarrow$ consist of the decreasing limits of sequences of functions in $L$. It is helpful to keep in mind that if $L$ consists of continuous functions on some topological space, then $L\uparrow$ consists of lower semicontinuous functions, while $L\downarrow$ consists of upper semicontinuous functions.

**Theorem 11.10** *Let $\mathcal{F}$ be a $\sigma$-algebra of real functions and $\mu$ be an integral associated with $\mathcal{F}$. Let $L$ be a Stone vector lattice, and assume that the $\sigma$-ring generated by $L$ is $\mathcal{F}$. Suppose that $\mu$ is finite on $L$. Then $\mu$ is upper regular, in the sense that for each $g$ in $\mathcal{L}^1$ we have $\mu(g) = \inf\{\mu(h \mid g \leq h, h \in L \uparrow\}$. Also, $\mu$ is lower regular, in the sense that for each $g$ in $\mathcal{L}^1$ we have $\mu(g) = \sup\{\mu(f) \mid f \leq g, f \in L \downarrow\}$.*

Proof: This follows from the construction of the integral and the uniqueness theorem. The restriction of the integral to $L$ is an elementary integral, and so we may construct an integral on $\mathcal{L}^1$ by using upper and lower functions. By the uniqueness theorem this is the original integral. □

Notice that in the topological context the theorem above might be interpreted as saying that each absolutely integrable function is both uppper LSC regular and lower USC regular.

Consider a subset $G$ to be an outer subset if it is of the form $h > a$ for some $h$ in $L^{\uparrow}$ and some real number $a$. Similarly, consider a subset $F$ to be an inner subset if it is of the form $f \geq a$ for some $f$ in $L \downarrow$ and $a$ real. Notice that the outer subsets and the inner subsets are complements of each other. In the case when $L$ consists of continuous functions, then the subsets $G$ in $L \uparrow$ are open subsets, while the sets $F$ in $L \downarrow$ are closed subsets.

**Theorem 11.11** *Let $\mathcal{F}$ be a $\sigma$-algebra of real functions and $\mu$ be a measure associated with $\mathcal{F}$. Let $L$ be a Stone vector lattice, and assume that the $\sigma$-ring generated by $L$ is $\mathcal{F}$. Suppose that $\mu$ is finite on $L$. Then $\mu$ is outer regular, in the sense that for each subset $E$ of finite measure we have $\mu(E) = \inf\{\mu(G) \mid E \subset G, G\text{outer}\}$. Suppose now in addition that the measure space is finite. Then also $\mu$ is inner regular, in the sense that for each subset $E$ of finite measure we have $\mu(e) = \sup\{\mu(F) \mid F \subset E, F\text{inner}\}$.*

Proof: The first part is the outer regularity. Consider $\epsilon > 0$. From the previous theorem there is a function $h$ in $L^{\uparrow}$ such that $1_E \leq h$ and $\mu(h) \leq \mu(E) + \epsilon/2$. Let $G_n$ be the set where $h > 1 - 1/n$. Then $G_n$ is an inner subset, and $E \subset G_n$. Furthermore, $\mu(G_n) \leq \mu(h)/(1 - 1/n) \leq (\mu(E) + \epsilon/2)/(1 - 1/n)$. For $n$ sufficiently large we have $\mu(G_n) \leq \mu(E) + \epsilon$.

The other part is the inner regularity. When the measure space is finite this follows by applying the outer regularity to the complements. □

In the topological context the theorem above might be interpreted as saying that the subsets are both outer open regular and inner closed regular.

## 11.6 Density

**Theorem 11.12** *Let $\mathcal{F}$ be a $\sigma$-algebra of real functions and $\mu$ be an integral associated with $\mathcal{F}$. Let $L$ be a Stone vector lattice, and assume that the $\sigma$-ring generated by $L$ is $\mathcal{F}$. Suppose that $\mu$ is finite on $L$. Then $L$ is dense in the pseudo-metric space $\mathcal{L}^1$.*

Proof: Since bounded functions in $\mathcal{L}^1$ are dense in $\mathcal{L}^1$, it is enough to approximate a bounded function $g$ by an element of $L$. Consider $\epsilon > 0$. By upper regularity one can take $h$ in $L^\uparrow$ with $g \leq h$ and $\mu(h) - \mu(g) < \epsilon/2$. Furthermore, $h$ can also be taken bounded. It follows that $h$ does not assume the value $+\infty$, and so it is itself in $\mathcal{L}^1$. Then one can take $f$ in $L$ with $f \leq h$ such that $\mu(h) - \mu(f) < \epsilon/2$. Since $f, g, h$ are all in $\mathcal{L}^1$ we have

$$\mu(|g - f|) \leq \mu(|g - h|) + \mu(|h - f|). \tag{11.1}$$

Thus

$$\mu(|g - f|) \leq \mu(h) - \mu(g) + \mu(h) - \mu(f) < \epsilon. \tag{11.2}$$

This is the required approximation. □

## 11.7 Monotone classes

A set of real functions $\mathcal{F}$ is a *monotone class* if it satisfies the following two properties. Whenever $f_n \uparrow f$ is an increasing sequence of functions $f_n$ in $\mathcal{F}$ with pointwise limit $f$, then $f$ is also in $\mathcal{F}$. Whenever $f_n \downarrow f$ is a decreasing sequence of functions $f_n$ in $\mathcal{F}$ with pointwise limit $f$, then $f$ is also in $\mathcal{F}$.

**Theorem 11.13** *Let $L$ be a vector lattice of real functions. Let $\mathcal{F}$ be the smallest monotone class of which $L$ is a subset. Then $\mathcal{F}$ is a vector lattice.*

Proof: The task is to show that $\mathcal{F}$ is closed under addition, scalar multiplication, sup, and inf. Begin with addition. Let $f$ be in $L$. Consider the set $M(f)$ of functions $g$ such that $f + g$ is in $\mathcal{F}$. This set includes $L$ and is closed under monotone limits. So $\mathcal{F} \subset M(f)$. Thus $f$ in $L$ and $g$ in $\mathcal{F}$ imply $f + g \in \mathcal{F}$. Now let $g$ be in $\mathcal{F}$. Consider the set $\tilde{M}(g)$ of functions $f$ such that $f + g$ is in $\mathcal{F}$. This set includes $L$ and is closed under monotone limits. So $\mathcal{F} \subset \tilde{M}(g)$. Thus $f$ and $g$ in $\mathcal{F}$ implies $f + g$ in $\mathcal{F}$. The proof is similar for the other operations. □

**Theorem 11.14** *Let $L$ be a Stone vector lattice of real functions. Let $\mathcal{F}$ be the smallest monotone class of which $L$ is a subset. Then $\mathcal{F}$ is a Stone vector lattice.*

**Theorem 11.15** *Let $L$ be a Stone vector lattice of real functions. Let $\mathcal{F}_0$ be the smallest monotone class of which $F$ is a subset. Then $\mathcal{F}_0$ is a $\sigma$-ring of functions.*

A set of real functions $\mathcal{F}$ is a *vector lattice with constants* of functions if it is a vector lattice and each constant function belongs to $\mathcal{F}$. The following theorem is trivial, but it may be worth stating the obvious.

**Theorem 11.16** *Let $L$ be a vector lattice with constants. Let $\mathcal{F}$ be the smallest monotone class of which $L$ is a subset. Then $\mathcal{F}$ is a vector lattice with constants.*

We shall now see that a monotone class is closed under all pointwise limits.

**Theorem 11.17** *Let $\mathcal{F}$ be a monotone class of functions. Let $f_n$ be in $\mathcal{F}$ for each $n$. Suppose that* $\liminf_n f_n$ *and* $\limsup f_n$ *are finite. Then they are also in $\mathcal{F}$.*

Proof: Let $n < m$ and let $h_{nm} = f_n \wedge f_{n+1} \wedge \cdots \wedge f_m$. Then $h_{nm} \downarrow h_n$ as $m \to \infty$, where $h_n$ is the infimum of the $f_k$ for $k \geq n$. However $h_n \uparrow \liminf_n f_n$. □

The trick in this proof is to write a general limit as an increasing limit followed by a decreasing limit. We shall see in the following that this is a very important idea in integration.

## 11.8 Generating monotone classes

The following theorem says that if $L$ is a vector lattice that generates $\mathcal{F}$ by monotone limits, then the positive functions $L^+$ generate the positive functions $\mathcal{F}^+$ by monotone limits.

**Theorem 11.18** *Let $L$ be a vector lattice of real functions. Suppose that $\mathcal{F}$ is the smallest monotone class that includes $L$. Let $L^+$ be the positive elements of $L$, and let $\mathcal{F}^+$ be the positive elements of $\mathcal{F}$. Then $\mathcal{F}^+$ is the smallest monotone class that includes $L^+$.*

Proof: It is clear that $\mathcal{F}^+$ includes $L^+$. Furthermore, $\mathcal{F}^+$ is a monotone class. So all that remains to show is that if $\mathcal{G}$ is a monotone class that includes $L^+$, then $\mathcal{F}^+$ is a subset of $\mathcal{G}$. For that it is sufficient to show that for each $f$ in $\mathcal{F}$ the positive part $f \vee 0$ is in $\mathcal{G}$.

Consider the set $M$ of $f$ in $\mathcal{F}$ such that $f \vee 0$ is in $\mathcal{G}$. The set $L$ is a subset of $M$, since $f$ in $L$ implies $f \vee 0$ in $L^+$. Furthermore, $M$ is a monotone class. To check this, note that if each $f_n$ is in $M$ and $f_n \uparrow f$, then $f_n \vee 0$ is in $\mathcal{G}$ and $f_n \vee 0 \uparrow f \vee 0$, and so $f \vee 0$ is also in $\mathcal{G}$, that is, $f$ is in $M$. The argument is the same for downward convergence. Hence $\mathcal{F} \subset M$. □

A real function $f$ is said to be *$L$-bounded* if there is a function $g$ in $L^+$ with $|f| \leq g$. Say that $L$ consists of bounded functions. Then if $f$ is $L$-bounded, then $f$ is also bounded. Say on the other hand that the constant functions are in $L$. Then if $f$ is bounded, it follows that $f$ is $L$-bounded. However there are also cases when $L$ consists of bounded functions, but the constant functions are not in $L$. In such cases, being $L$-bounded is more restrictive.

A set of real functions $\mathcal{H}$ is an *$L$-bounded monotone class* if it satisfies the following two properties. Whenever $f_n \uparrow f$ is an increasing sequence of $L$-bounded functions $f_n$ in $\mathcal{H}$ with pointwise limit $f$, then $f$ is also in $\mathcal{H}$. Whenever $f_n \downarrow f$ is a decreasing sequence of $L$-bounded functions $f_n$ in $\mathcal{H}$ with pointwise limit $f$, then $f$ is also in $\mathcal{H}$. Notice that the functions in $\mathcal{H}$ do not have to be $L$-bounded.

The following theorem says that if $L^+$ generates $\mathcal{F}^+$ by monotone limits, then $L^+$ generates $\mathcal{F}^+$ using only monotone limits of $L$-bounded functions.

**Theorem 11.19** *Let $L$ be a vector lattice of bounded real functions that includes the constant functions. Let $\mathcal{F}^+$ be the smallest monotone class of which $L^+$ is a subset. Let $\mathcal{H}$ be the smallest $L$-bounded monotone class of which $L^+$ is a subset. Then $\mathcal{H} = \mathcal{F}^+$.*

Proof: It is clear that $\mathcal{H} \subset \mathcal{F}^+$. The task is to prove that $\mathcal{F}^+ \subset \mathcal{H}$.

Consider $g \geq 0$ be in $L^+$. Let $M(g)$ be the set of all $f$ in $\mathcal{F}^+$ such that $f \wedge g$ is in $\mathcal{H}$. It is clear that $L^+ \subset M(g)$. If $f_n \uparrow f$ and each $f_n$ is in $M(g)$, then $f_n \wedge g \uparrow f \wedge g$. Since each $f_n \wedge g$ is in $\mathcal{H}$ and is $L$-bounded, it follows that $f \wedge g$ is in $\mathcal{H}$. Thus $M(g)$ is closed under upward monotone convergence. Similarly, $M(g)$ is closed under downward monotone convergence. Therefore $\mathcal{F}^+ \subset M(g)$. This establishes that for each $f$ in $\mathcal{F}_+$ and $g$ in $L^+$ it follows that $f \wedge g$ is in $\mathcal{H}$.

Now consider the set of all $f$ in $\mathcal{F}$ such that there exists $h$ in $L\uparrow$ with $f \leq h$. Certainly $L$ belongs to this set. Furthermore, this set is monotone. This is obvious for downward monotone convergence. For upward monotone convergence, it follows from the fact that $L\uparrow$ is closed under upward monotone convergence. It follows that every element in $\mathcal{F}$ is in this set.

Let $f$ be in $\mathcal{F}^+$. Then there exists $h$ in $L\uparrow$ such that $f \leq h$. There exists $h_n$ in $L^+$ with $h_n \uparrow h$. Then $f \wedge h_n$ is in $\mathcal{H}$, by the first part of the proof. Furthermore, $f \wedge h_n \uparrow f$. It follows that $f$ is in $\mathcal{H}$. This completes the proof that $\mathcal{F}^+ \subset \mathcal{H}$. □

## 11.9 Proof of the uniqueness theorem

**Theorem 11.20 (improved monotone convergence)** *If $\mu(f_1) > -\infty$ and $f_n \uparrow f$, then $\mu(f_n) \uparrow \mu(f)$. Similarly, if $\mu(h_1) < +\infty$ and $h_n \downarrow h$, then $\mu(h_n) \downarrow \mu(h)$.*

Proof: For the first apply monotone convergence to $f_n - f_1$. For the second let $f_n = -h_n$. □

Proof: Let $\mu_1$ and $\mu_2$ be two integrals on $\mathcal{F}^+$ that each agree with $m$ on $L^+$. Let $\mathcal{H}$ be the smallest $L$-monotone class such that $L^+ \subset \mathcal{H}$. Let $\mathcal{G}$ be the set of all functions in $\mathcal{F}^+$ on which $\mu_1$ and $\mu_2$ agree. The main task is to show that $\mathcal{H} \subset \mathcal{G}$. It is clear that $L \subset \mathcal{G}$. Suppose that $h_n$ is in $\mathcal{G}$ and $h_n \uparrow h$. If $\mu_1(h_n) = \mu_2(h_n)$ for each $n$, then $\mu_1(h) = \mu_2(h)$. Suppose that $f_n$ is in $\mathcal{G}$ and is $L$-bounded for each $n$ and $f_n \downarrow f$. If $\mu_1(f_n) = \mu_2(f_n)$ for all $n$, then by improved monotone convergence $\mu_1(f) = \mu_2(f)$. This shows that $\mathcal{G}$ is a $L$-monotone class such that $L^+ \subset \mathcal{G}$. It follows that $\mathcal{H} \subset \mathcal{G}$. However the earlier result on $L$-monotone classes showed that $\mathcal{H} = \mathcal{F}^+$. So $\mathcal{F}^+ \subset \mathcal{G}$. □

## 11.10 Proof of the $\sigma$-finiteness theorem

Proof: Suppose that $\mu$ is $\sigma$-finite. Let $L = \mathcal{L}^1(X, \mathcal{F}, \mu)$. Consider the monotone class generated by $L$. Since $\mu$ is $\sigma$-finite, the constant functions belong to this monotone class. So it is a $\sigma$-algebra. In fact, this monotone class is equal to

$\mathcal{F}$. To see this, let $E_n$ be a family of finite measure sets that increase to $X$. Consider a function $g$ in $\mathcal{F}$. For each $n$ the function $g_n = g 1_{E_n} 1_{|g| \leq n}$ is in $L$. Then $g = \lim_n g_n$ is in the monotone class generated by $L$.

Suppose on the other hand that there exists such a vector lattice $L$. Consider the class of functions $f$ for which there exists $h$ in $L \uparrow$ with $f \leq h$. This class includes $L$ and is monotone, so it is includes all of $\mathcal{F}$.

Take $f$ in $\mathcal{F}^+$. Then there exists $h$ in $L \uparrow$ with $f \leq h$. Take $h_n \in L^+$ with $h_n \uparrow h$. Then $u_n = f \wedge h_n \uparrow f$. Thus there is a sequence of $L$-bounded functions $u_n$ in $\mathcal{F}^+$ such that $u_n \uparrow f$. Each of these functions $u_n$ has finite integral. In the present case $\mathcal{F}$ is a $\sigma$-algebra, so we may take take $f = 1$. This completes the proof that $\mu$ is $\sigma$-finite. □

The only if part of the theorem gives the existence of a vector lattice, but not necessarily the one originally used to generate the $\sigma$-algebra. Recall the example of the trivial vector lattice $L$ with only the zero function. The monotone class it generates is still trivial. However, the elementary integral on $L$ has an extension to a finite integral on the $\sigma$-algebra of constant functions.

## 11.11 Supplement: Completion of an integral

Suppose $\mu$ is an integral defined with respect to $\sigma$-algebra $\mathcal{F}$ of functions. A function $f$ in $\mathcal{F}$ is called a *null function* if $\mu(|f|) = 0$. A function $g$ is called a *null-dominated function* if $|g| \leq f$ for some null function $f$. A null-dominated function need not be in $\mathcal{F}$.

Let $\bar{\mathcal{F}}_\mu$ be the $\sigma$-algebra of functions generated by $\mathcal{F}$ together with its null-dominated functions. This is called the *completion of the $\sigma$-algebra*. It may be shown that the integral $\mu$ extends uniquely to an integral $\bar{\mu}$ with respect to $\bar{\mathcal{F}}_\mu$. This is called the *completion* of the integral.

The standard example is when $\lambda$ is the Lebesgue integral defined for Borel measurable functions $\mathcal{B}o$. The completion $\bar{\lambda}$ is defined with respect to the measurable functions $\bar{\mathcal{B}o}_\lambda$. The space $\bar{\mathcal{B}o}_\lambda$ is customarily called the space of *Lebesgue measurable functions.*

The space of Lebesgue measurable functions is much larger than the space of Borel measurable functions. In fact, the space of Borel measurable functions has cardinality $c$, while the space of Lebesgue measurable functions has cardinality $2^c = c^c$, which is a large as the cardinality of the space $\mathbb{R}^\mathbb{R}$ of all functions. So it would seem that this completed Lebesgue integral with its extremely huge domain of definition would be just the right thing.

As a matter of fact, it is seldom needed, and in fact could be somewhat of a nuisance. The Borel functions are already a huge class of functions, and it is difficult to give an example of a function that is not Borel, though such functions may be constructed. The Lebesgue measurable functions are an extremely huge class of functions, and it is impossible to give a specific example of a function that is not Lebesgue measurable, at least not without using the axiom of choice. But all those extra null-dominated functions play little role in practical problems; after all, they have extended Lebesgue integral equal to zero.

There is a good case for staying with the Borel functions. This is because they are defined independent of the integral. Suppose, as is common, that one wants to talk of two different integrals in the same context. If the common domain of definition consists of positive Borel functions, then it is easy to compare them. However their completions may have different domains, and that could lead to considerations that have nothing to do with any concrete problem.

## Problems

1. Let $X$ be a set. Let $L$ be the vector lattice of functions that are non-zero only on finite sets. The elementary integral $m$ is defined by $m(f) = \sum_{x \in S} f(x)$ if $f \neq 0$ on $S$. Find the $\sigma$-ring of functions $\mathcal{F}_0$ generated by $L$. When is it a $\sigma$-algebra? Extend $m$ to an integral $\mu$ on the smallest $\sigma$-algebra generated by $L$. Is the value of $\mu$ on the constant functions uniquely determined?

2. Consider the previous problem. The largest possible $\sigma$-algebra of functions on $X$ consists of all real functions on $X$. For $f \geq 0$ in this largest $\sigma$-algebra define the integral $\mu$ by $\mu(f) = \sum_{x \in S} f(x)$ if $f$ is non-zero on a countable set $S$. Otherwise define $\mu(f) = +\infty$. Is this an integral?

3. Let $X$ be a set. Let $A$ be a countable subset of $X$, and let $p$ be a function on $A$ with $p(x) \geq 0$ and $\sum_{x \in A} p(x) = 1$. Let $L$ be the vector lattice of functions that are non-zero only on finite sets. The probability sum is defined for $f$ in $L$ by $\mu(f) = \sum_{x \in A \cap S} f(x)p(x)$ if $f \neq 0$ on $S$. Let $\mathcal{F}_0$ be the $\sigma$-ring of functions generated by $L$. Show that if $X$ is uncountable, then $\mu$ has more than one extension to the $\sigma$-algebra $\mathcal{F}$ consisting of the sum of functions in $\mathcal{F}_0$ with constant functions. Which extension is natural for probability theory?

# Chapter 12

# Mapping integrals

## 12.1 Comparison of integrals

This chapter presents some interesting integrals. In order to compare integrals, it is useful to have a common domain. Thus, for example, let $X$ be a non-empty set and let $L$ be a Stone vector lattice of real functions on $X$. Then a suitable domain for integrals might be $\sigma(L)$, the smallest $\sigma$-algebra of real functions including $L$.

Consider, for example, the situation when $X$ is a metric space and $L$ consists of continuous functions. It may be that every continuous function on $X$ is a pointwise limit of a sequence of functions in $L$. In that case, $\sigma(L) = \mathcal{B}o$, the Borel $\sigma$-algebra of real functions on $X$.

There are many integrals that one could consider. One surprisingly useful class of examples are the integrals $\delta_p$, where $p$ is a point in $X$. This is defined by $\delta_p(f) = f(p)$. It is called the *unit point mass* at $p$, or the *Dirac delta measure* at $p$.

In general mass is a word that is used informally for measure. The idea is that the measure $\mu(S)$ of a subset $S$ is the amount of mass in the region S. Thus the point mass at $a$ describes a situation where a total mass of 1 is concentrated at the point $a$. This is because the measure $\delta_p(S)$ is 1 if $a$ is in $S$ and is 0 otherwise. In other words, all the mass is sitting at $p$.

Linear combinations of integrals with positive coefficients are also integrals. So this gives a way of generating new integrals from old. For example $\sum_j c_j \delta_{p_j}$ describes masses $c_j > 0$ sitting at the points $p_j$. The term for a single term $c\delta_p$ with $c > 0$ is *point mass* with mass $c$. A sum of point masses is called a *discrete measure*. On the other hand, a measure that assigns measure zero to each one point set is called a *continuous measure*.

## 12.2 Probability and expectation

An integral is a *probability integral* (or *expectation*) provided that $\mu(1) = 1$. This of course implies that $\mu(c) = c$ for every real constant $c$. In this context there is a special terminology. The set on which the functions are defined is called $\Omega$. A point $\omega$ in $\Omega$ is called an *outcome*.

A measurable function $f : \Omega \to \mathbb{R}$ is called a *random variable*. The value $f(\omega)$ is regarded as an experimental number, the value of the random variable when the outcome of the experiment is $\omega$. The integral $\mu(f)$ is the expectation of the random variable, provided that the integral exists. For a bounded measurable function $f$ the expectation $\mu(f)$ always exists.

A subset $A \subset \Omega$ is called an *event*. When the outcome $\omega \in A$, the event $A$ is said to happen. The measure $\mu(A)$ of an event is called the *probability* of the event. The probability $\mu(A)$ of an event $A$ is the expectation $\mu(1_A)$ of the random variable $1_A$ that is one if the event happens and is zero if the event does not happen.

**Theorem 12.1** *Let $\Omega = \{0,1\}^{\mathbf{N}_+}$ be the set of all infinite sequences of zeros and ones. Fix $p$ with $0 \leq p \leq 1$. If the function $f$ on $\Omega$ is in the space $\mathcal{F}_k$ of functions that depend only on the first $k$ values of the sequence, let $f(\omega) = h(\omega_1, \ldots, \omega_k)$ and define*

$$\mu_p(f) = \sum_{\omega_1=0}^{1} \cdots \sum_{\omega_k=0}^{1} h(\omega_1, \ldots, \omega_k) p^{\omega_1}(1-p)^{1-\omega_1} \cdots p^{\omega_k}(1-p)^{1-\omega_k}. \quad (12.1)$$

*This defines an elementary integral $\mu_p$ on the vector lattice $L$ that is the union of the $\mathcal{F}_k$ for $k = 0, 1, 2, 3, \ldots$. Let $\mathcal{F}$ be the $\sigma$-algebra generated by $L$. Then the elementary integral extends to an integral $\mu_p$ on $\mathcal{F}^+$, and this integral is uniquely defined.*

This theorem describes the expectation for a sequence of independent coin tosses where the probability of heads on each toss is $p$ and the probability of tails on each toss is $1 - p$. The special case $p = 1/2$ describes a fair coin. The proof of the theorem follows from previous considerations. It is not difficult to calculate that $\mu$ is consistently defined on $L$. It is linear and order preserving on the coin tossing vector lattice $L$, so it is automatically an elementary integral. Since $L$ contains the constant functions, the integral extends uniquely to the $\sigma$-algebra $\mathcal{F}$.

This family of integrals has a remarkable property. For each $p$ with $0 \leq p \leq 1$ let $F_p \subset \Omega$ be defined by

$$F_p = \{\omega \in \Omega \mid \lim_{n\to\infty} \frac{\omega_1 + \cdots + \omega_n}{n} = p\}. \quad (12.2)$$

It is clear that for $p \neq p'$ the sets $F_p$ and $F_{p'}$ are disjoint. This gives an uncountable family of disjoint measurable subsets of $\Omega$. The remarkable fact is that for each $p$ we have that the probability $\mu_p(F_p) = 1$. (This is the famous strong law of large numbers.) It follows that for $p' \neq p$ we have that the probability $\mu_p(F_{p'}) = 0$. Thus there are uncountably many expectations $\mu_p$. These are each defined with the same set $\Omega$ of outcomes and the same $\sigma$-algebra $\mathcal{F}$ of random variables. Yet they are concentrated on uncountably many different sets.

## 12.3 Image integrals

There are several ways of getting new integrals from old ones. One is by using a weight function. For instance, if

$$\lambda(f) = \int_{-\infty}^{\infty} f(x)\, dx \tag{12.3}$$

is the Lebesgue integral defined for Borel functions $f$, and if $w \geq 0$ is a Borel function, then

$$\mu(f) = \int_{-\infty}^{\infty} f(x) w(x)\, dx \tag{12.4}$$

is another integral. In applications $w$ can be a mass density, a probability density, or the like.

In general it is very common to denote

$$\mu(f) = \int f\, d\mu \tag{12.5}$$

or even

$$\mu(f) = \int f(x)\, d\mu(x). \tag{12.6}$$

This notation is suggestive in the case when there is more than one integral in play. Say that $\nu$ is an integral, and $w \geq 0$ is a measurable function. Then the integral $\mu(f) = \nu(fw)$ is defined. We would write this as

$$\int f(x)\, d\mu(x) = \int f(x)\, w(x) d\nu(x). \tag{12.7}$$

So the relation between the two integrals would be $d\mu(x) = w(x)d\nu(x)$. This suggests that $w(x)$ plays the role of a derivative of one integral with respect to the other.

A more important method is to map the integral forward. For instance, let $y = \phi(x) = x^2$. Then the integral $\mu$ described just above maps to an integral $\nu = \phi[\mu]$ given by

$$\nu(g) = \int_{-\infty}^{\infty} g(x^2) w(x)\, dx. \tag{12.8}$$

This is a simple and straightforward operation. Notice that the forward mapped integral lives on the range of the mapping, that is, in this case, the positive real axis. The trouble begins only when one wants to write this new integral in terms of the Lebesgue integral. Thus we may also write

$$\nu(g) = \int_0^\infty g(y)\frac{1}{2\sqrt{y}}[w(\sqrt{y}) + w(-\sqrt{y})]\,dy. \tag{12.9}$$

Here is the same idea in a general setting. Let $\mathcal{F}$ be a $\sigma$-algebra of measurable functions on $X$. Let $\mathcal{G}$ be a $\sigma$-algebra of measurable functions on $Y$. A function $\phi : X \to Y$ is called a *measurable map* if for every $g$ in $\mathcal{G}$ the composite function $g \circ \phi$ is in $\mathcal{F}$.

Given an integral $\mu$ defined on $\mathcal{F}$, and given a measurable map $\phi : X \to Y$, there is an integral $\phi[\mu]$ defined on $\mathcal{G}$. It is given by

$$\phi[\mu](g) = \mu(g \circ \phi). \tag{12.10}$$

It is called the *image integral* of the integral $\mu$ under $\phi$. Since integrals determine measures and are often even called measures, this construction is also called the *image measure*.

There is another, more abstract, way of thinking of this. Let $\phi^*$ be a map from real functions on $Y$ to real functions on $X$ defined by $\phi^*(f) = f \circ \phi$. Sometimes this is called the *pullback* map. Then define the map on measures by $\phi[\mu]$ by $\phi[\mu] = \mu \circ \phi^*$. Then $\phi[\mu](f) = \mu(\phi^*(f)) = \mu(f \circ \phi)$ as before. It might seem reasonable to call this map on measures the *pushforward* map.

This construction is important in probability theory. Let $\Omega$ be a measure space equipped with a $\sigma$-algebra of functions $\mathcal{F}$ and an expectation $\mu$ defined on $\mathcal{F}^+$. If $\phi$ is a random variable, that is, a measurable function from $\Omega$ to $\mathbb{R}$ with the Borel $\sigma$-algebra, then it may be regarded as a measurable map. The image of the expectation $\mu$ under $\phi$ is an integral $\nu = \phi[\mu]$ on the Borel $\sigma$-algebra called the *distribution* of $\phi$. We have the identity.

$$\mu(h(\phi)) = \nu(h) = \int_{-\infty}^{\infty} h(x)\,d\nu(x). \tag{12.11}$$

Sometimes the calculations do not work so smoothly. The reason is that there are really two theories of integration. The integral in real analysis acts on functions and maps forward under measurable maps. The integral in geometry and calculus pairs differential forms with oriented geometrical objects, and the differential forms maps backward under smooth maps. For instance, the differential form $g(y)\,dy$ maps backward to the differential form $g(\phi(x))\phi'(x)\,dx$. Thus a differential form calculation with oriented integrals like

$$\int_a^b g(\phi(x))\phi'(x)\,dx = \int_{\phi(a)}^{\phi(b)} g(y)\,dy \tag{12.12}$$

works very smoothly. The oriented interval maps forward; the differential form maps backward. On the other hand, the calculation of an integral in the sense of real analysis, even with a smooth change of variable with $\phi'(x) \neq 0$, gives

$$\int_{[a,b]} g(\phi(x))\,dx = \int_{\phi([a,b])} g(y)\left|\frac{1}{\phi'(\phi^{-1}(y))}\right|\,dy \tag{12.13}$$

which involves an unpleasant denominator. The problem is not with the integral, which is perfectly well defined by the left hand side with no restrictions on the function $\phi$ other than measurability. The integral maps forward; the function maps backward. The difficulty comes when one tries to express the image integral as a Lebesgue integral with a weight function. It is only at this stage that the differential form calculations play a role.

The ultimate source of this difficulty is that integrals (or measures) and differential forms are different kinds of objects. An integral assigns a number to a function. Functions map backward, so integrals map forward. Thus $g$ pulls back to $g \circ \phi$, so $\mu$ pushes forward to $\phi[\mu]$. The value of $\phi[\mu]$ on $g$ is the value of $\mu$ on $g \circ \phi$. (It makes no difference if we think instead of measures defined on subsets, since subsets map backwards and measures map forward.) A differential form assigns a number to an oriented curve. Curves map forward, so differential forms map backward. Thus a curve from $a$ to $b$ pushes forward to a curve from $\phi(a)$ to $\phi(b)$. The differential form $g(y)\,dy$ pulls back to the differential form $g(\phi(x))\phi'(x)\,dx$. The value of $g(\phi(x)\phi'(x)\,dx$ over the curve from $a$ to $b$ is the value of $g(y)\,dy$ over the curve from $\phi(a)$ to $\phi(b)$.

## 12.4 The Lebesgue integral

The image construction may be used to relate measures on the coin-tossing space to measures on the unit interval.

**Theorem 12.2** *Let $0 \leq p \leq 1$. Define the expectation $\mu_p$ for coin tossing on the set $\Omega$ of all infinite sequences $\omega : \mathbb{N}_+ \to \{0, 1\}$ as in the theorem. Here $p$ is the probability of heads on each single toss. Let*

$$\phi(\omega) = \sum_{k=1}^{\infty} \omega_k \frac{1}{2^k}. \tag{12.14}$$

*Then the image expectation $\phi[\mu_p]$ is an expectation $\nu_p$ defined for Borel functions on the unit interval $[0, 1]$.*

The function $\phi$ in this case is a random variable that rewards the $n$th coin toss by $1/2^n$ if it results in heads, and by zero if it results in tails. The random variable is the sum of all these rewards. Thus $\nu_p$ is the distribution of this random variable.

When $p = 1/2$ (the product expectation for tossing of a fair coin) the expectation $\lambda_1 = \nu_{\frac{1}{2}}$ is the Lebesgue integral for functions on $[0, 1]$. However note that there are many other integrals, for the other values of $p$. We have the following amazing fact. For each $p$ there is an integral $\nu_p$ defined for functions on the unit interval. If $p \neq p'$ are two different parameters, then there is a measurable set that has measure 1 for the $\nu_p$ measure and measure 0 for the $\nu_{p'}$ measurable. The set comes from the set of coin tosses for which the sample means converge to the number $p$. This result shows that these measures each live in a different world.

Now start with Lebesgue integral for Borel functions on the unit interval $[0, 1]$ given by

$$\lambda_1(f) = \int_0^1 f(u)\, du. \tag{12.15}$$

The image construction then gives many new integrals.

Consider the map $x = \psi(u) = \ln(u/(1-u))$ from the open interval $(0, 1)$ to $\mathbb{R}$. This is a bijection. It has derivative $dx/du = 1/(u(1-u))$. The inverse is $u = 1/(1-e^{-x})$ with derivative $u(1-u) = 1/(2+2\cosh(x))$. It is a transformation that is often used in statistics to relate problems on the unit interval $(0, 1)$ and on the line $(-\infty, +\infty)$. The image of the Lebesgue integral for $[0, 1]$ under this map is also a probability integral. It is given by

$$\psi[\lambda](f) = \int_0^1 f(\ln\left(\frac{u}{1-u}\right))\, du = \int_{-\infty}^{\infty} f(x)\,\frac{1}{2}\frac{1}{1+\cosh(x)}\, dx. \tag{12.16}$$

A variation of this idea may be used to obtain the usual Lebesgue integral for Borel functions defined on the real line $\mathbb{R}$. Let

$$\sigma(h) = \int_0^1 h(u)\frac{1}{u(1-u)}\, du. \tag{12.17}$$

This is not a probability integral. The image under $\psi$ is

$$\psi[\sigma](f) = \int_0^1 f(\ln(\frac{u}{1-u}))\frac{1}{u(1-u)}\, du = \int_{-\infty}^{\infty} f(x)\, dx = \lambda(f). \tag{12.18}$$

This calculation shows that the $dx$ integral is the image of the $1/(u(1-u))\, du$ integral under the transformation $x = \ln(u/(1-u))$. It could be taken as the final step in a multi-step construction that starts with the fair coin-tossing expectation $\mu_{\frac{1}{2}}$ and ends with the Lebesgue integral $\lambda$ for functions on the line.

## 12.5 Lebesgue-Stieltjes integrals

Once we have the Lebesgue integral defined for Borel functions on the line, we can construct a huge family of other integrals, also defined on Borel functions on the line. These are called *Lebesgue-Stieltjes integrals*. Often when several integrals are being discussed, the integrals are referred to as measures. Of course an integral defined on functions does indeed define a measure on subsets.

The class of measures under consideration consists of those measures defined on Borel functions on the line (or on Borel subsets of the line) that give finite measure to compact Borel subsets.

Examples:

1. The first example is given by taking a function $w \geq 0$ such that $w$ in absolutely integrable over each bounded Borel set. The measure is then $\mu(f) = \lambda(fw) = \int_{\infty}^{\infty} f(x)w(x)\, dx$. Such a measure is called *absolutely continuous* with respect to Lebesgue measure. Often the function $w$ is called the relative *density* (of mass or probability).

2. Another kind of example is of the form $\mu(f) = \sum_{p \in S} c_p f(p)$, where $S$ is a countable subset of the line, and each $c_p > 0$. This is called the measure that assigns point mass $c_p$ to each point $p$ in $S$. We require that $\sum_{a<p\leq b} c_p < \infty$ for each $a, b$ with $-\infty < a < b < +\infty$. Often the measure that assigns mass one to a point $p$ is denoted $\delta_p$, so $\delta_p(f) = f(p)$. With this notation the measure $\mu$ of this example is $\mu = \sum_{p \in S} c_p \delta_p$.

Suppose that $\mu$ is a Borel measure finite on compact subsets, so that the measure of each set $(a, b]$ for $a \leq b$ real is finite. Define $F(x) = \mu((0, x])$ for $x \geq 0$ and $F(x) = -\mu((x, 0])$ for $x \leq 0$. Then $F((a, b]) = F(b) - F(a)$ for all $a \leq b$. The function $F$ is increasing and right continuous. With this normalization $F(0) = 0$, but one can always add a constant to $F$ and still get the property that $F((a, b]) = F(b) - F(a)$. This nice thing about this is that the increasing right continuous function $F$ gives a rather explicit description of the measure $\mu$. It is often called the *distribution function* of the measure $\mu$.

Examples:

1. For the absolutely continuous measure $F(b) - F(a) = \int_a^b w(x)\,dx$. The function $F$ is a continuous function. However not every continuous function is absolutely continuous.

2. For the point mass measure $F(b) - F(a) = \sum_{p \in (a,b]} c_p$. The function $F$ is continuous except for jumps at the points $p$ in $S$.

**Theorem 12.3** *Let $F$ be an increasing right continuous function on $\mathbb{R}$. Then there exists a measure $\sigma_F$ defined on the Borel $\sigma$-algebra $\mathcal{B}o$, finite on compact subsets, such that*

$$\sigma_F((a, b]) = F(b) - F(a). \tag{12.19}$$

*Furthermore, this measure may be obtained as the image of Lebesgue measure on an interval under a map $G$.*

Proof: Let $m = \inf F$ and let $M = \sup F$. For $m < y < M$ let

$$G(y) = \sup\{x \mid F(x) < y\}. \tag{12.20}$$

We can compare the least upper bound $G(y)$ with an arbitrary upper bound $c$. Thus $G(y) \leq c$ is equivalent to the condition that for all $x$, $F(x) < y$ implies $x \leq c$. This in turn is equivalent to the condition that for all $x$, $c < x$ implies $y \leq F(x)$. Since $F$ is increasing and right continuous, it follows that this in turn is equivalent to the condition that $y \leq F(c)$.

It follows that $a < G(y) \leq b$ is equivalent to $F(a) < y \leq F(b)$. Thus $G$ is a kind of inverse to $F$.

Let $c = M - m$ and let $\lambda_c$ be Lebesgue measure on the interval $(m, M)$ with total mass $c$. Let $\sigma_F = G[\lambda_c]$ be the image of $\lambda_c$ under $G$. Then

$$\sigma_F((a, b]) = \lambda_c(\{y \mid a < G(y) \leq b\}) = \lambda_c(\{y \mid F(a) < y \leq F(b)\}) = F(b) - F(a), \tag{12.21}$$

so $\sigma_F$ is the desired measure. □

The above proof says that every Borel measure, finite on compact subsets, and with a certain total mass $M - m$ may be obtained from Lebesgue measure with the same mass. The forward mapping $G$ just serves to redistribute the mass.

Often the Lebesgue-Stieltges integral is written

$$\sigma_F(h) = \int_{-\infty}^{\infty} h(x)\, dF(x). \tag{12.22}$$

This is reasonable, since if $F$ were smooth with smooth inverse $G$ we would have $F(G(y)) = y$ and $F'(G(y))G'(y) = 1$ and so

$$\sigma_F(h) = \lambda(h \circ G) = \int_m^M h(G(y))\, dy = \int_m^M h(G(y))F'(G(y))G'(y)\, dy = \int_{-\infty}^{\infty} h(x)F'(x)\, dx. \tag{12.23}$$

However in general it is not required that $F$ be smooth, or that it have an inverse function.

These increasing functions $F$ give a relatively concrete description of the Borel measures, finite on compact subsets. There are three qualitatively different situations. If the function $F(x)$ is the indefinite integral of a function $w(x)$, then $F$ and $\sigma_F$ are said to be *absolutely continuous* with respect to Lebesgue measure. (In this case, it is reasonable to write $w(x) = F'(x)$. However $F(x)$ need not be differentiable at each point $x$. The example when $w(x)$ is a rectangle function provides an example.) If the function $F$ is constant except for jumps at a countable set of points, $F$ and $\sigma_F$ are said to have *point masses.* The third situation is intermediate and rather strange: the function $F$ has no jumps, but it is constant except on a set of measure zero. In this case $F$ and $\sigma_F$ are said to be *singular continuous.*

The classification into absolutely continuous, singular continuous, and point mass is very useful. In some contexts, such as the abstract characterization of measures, one wants to group the first two together and speak of continuous measure versus point mass measure. In situations where the existence of a density is crucial, the grouping is absolutely continuous measure versus singular measure.

Notice that in all cases the Lebesgue-Stieltjes measures are automatically both outer open regular and inner closed regular. In fact, they are even inner compact regular, in the sense that a measure $\mu$ is inner compact regular if for every Borel subset $E$ we have $\mu(F) = \sup\{\mu(K) \mid K \subset F, K\text{compact}\}$. This may be proved by approximating the measure of the subset by the measure of its intersection with a large bounded closed interval.

The conclusion of this discussion is that there are many regular Borel measures on Borel subsets of the line. However there is a kind of unity, since each of these is an image of Lebesgue measure on some interval.

## 12.6 The Cantor measure and the Cantor function

Here is an example of the singular continuous case. Let $\mu_{\frac{1}{2}}$ be the fair coin measure on the space $\Omega$ of all sequences of zeros and ones. Let $\chi(\omega) = \sum_{n=1}^{\infty} 2\omega_n/3^n$. Then $\chi$ maps $\Omega$ bijectively onto the Cantor set. Thus $\chi[\mu_{\frac{1}{2}}]$ is a probability measure on the line that assigns all its weight to the Cantor set. This is the *Cantor measure*. The function $F$ that goes with this measure is called the *Cantor function*. It is a continuous function that increases from zero to one, yet it is constant except on the Cantor set, which has measure zero.

The Cantor function $F$ is the distribution function of the random variable $\chi$, that is, $F(x) = \mu_{\frac{1}{2}}(\chi \leq x)$. It also has a simple non-probabilistic description. To see this, recall that $\phi$ defined by $\phi(\omega) = \sum_{n=1}^{\infty} \omega_n/2^n$ has a uniform distribution on $[0, 1]$. This says that $\mu_{\frac{1}{2}}(\phi \leq y) = y$ for all $y$ in $[0, 1]$. For $x$ in the Cantor set there is a unique $\chi^{-1}(x)$ in $\Omega$ and a corresponding $y = \phi(\chi^{-1}(x))$ in $[0, 1]$. The set where $\phi \leq y$ is the same as the set where $\chi \leq x$, up to a set of measure zero. Therefore $F(x) = \mu_{\frac{1}{2}}(\chi \leq x) = \mu_{\frac{1}{2}}(\phi \leq y) = y$. The conclusion is that $F$ restricted to the Cantor set is $\phi \circ \chi^{-1}$, and $F$ is constant elsewhere.

The map $\phi \circ \chi^{-1}$ from the the middle third Cantor set to the unit interval has a rather concrete description. Take a real number in the Cantor set and consider its base three expansion. This will have only 0 and 2 coefficients. Look at the base 2 expansion with 0 and 1 coefficients in the corresponding places. This converges to a real number in the unit interval.

Consider two such base 3 expansions that agree in the first $n - 1$ places, but such that one has a 0 in the $n$th place and the other has a 2 in the $n$th place. Then they differ by at least $1/3^n$. The interval between the largest one with a 0 (which subsequently has all 2s) and the smallest one with a 2 (which subsequently has all 0s) is an interval of constancy of the Cantor function of length $1/3^n$. The numbers in the interior of this interval all have base 3 expansions with a 1 in the $n$th place. The end points of the interval map into the same number. This number has two base 2 expansions, one with a 0 in the $n$th place followed by all 1s, and the other with a 1 in the $n$th place followed by all 0s. So it is a rational number of the form $k/2^n$. Since the Cantor function is increasing, the points in the interior of the interval also map to this same number. Such numbers are the values of the Cantor function on the intervals of constancy.

## 12.7 Change of variable

This section presents a statement of the change of variable formula in $n$ dimensions. See Folland [5] for a proof.

**Theorem 12.4 (Change of variable)** *Let $U \subset \mathbb{R}^n$ be an open subset. Let $g$ be a $C^1$ injective function from $U$ to $\mathbb{R}^n$. Then $|\det g'|$, the absolute value of* *the determinant of the matrix of partial derivatives, is a $C^1$ function from $U$ to $[0, +\infty)$. Let $\nu$ be the integral defined for functions on $U$ by $\nu(f) = \lambda(f|\det g'|)$. Thus $\nu$ has density $J$ with respect to Lebesgue measure on $U$. Then the image measure $g[\nu]$ is Lebesgue measure on $g[U]$. In other words,*

$$\lambda((f \circ g)|\det g'|) = \lambda(f). \tag{12.24}$$

Let $C$ be the set where $\det g' = 0$, and set $f = 1_C \circ g^{-1} = 1_{g[C]}$. It follows that the Lebesgue measure of $g[C]$ is zero. Also, the function $|\det g'| \circ g^{-1}$ is zero only on $g[C]$. So we may apply the change of variable formula to the function $f = h/(|\det g')| \circ g^{-1})$. This gives the following corollary.

**Corollary 12.5 (Image of Lebesgue measure)** *Let $U \subset \mathbb{R}^n$ be an open subset. Let $g$ be a $C^1$ injective function from $U$ to $\mathbb{R}^n$. Then $|\det g'|$, the absolute value of the determinant of the matrix of partial derivatives, is a $C^1$ function from $U$ to $[0, +\infty)$. Let $\mu$ be the integral defined for functions on $g[U]$ by $\mu(h) = \lambda(h/(|\det g'| \circ g^{-1})$. Then if $\lambda$ is Lebesgue measure on $U$, then the image measure $g[\lambda]$ is $\mu$. In other words,*

$$g[\lambda](h) = \lambda(h \circ g) = \lambda(h \frac{1}{|\det g'| \circ g^{-1}}. \tag{12.25}$$

## 12.8 Supplement: Direct construction of the Lebesgue-Stieltjes measure

It is also possible to give a direct construction of the Lebesgue-Stieltjes measure corresponding to an increasing right continuous function $F$. This is because the following Dini lemma for step functions has a direct self-contained proof. Once the lemma is established, then the general construction of an integral from an elementary integral applies.

**Lemma 12.6** *Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing right continuous function. For each function*

$$f = \sum_{k=1}^{m} c_k 1_{(a_k, b_k]} \tag{12.26}$$

*define*

$$\sigma_F(f) = \sum_{k=1}^{m} c_k (F(b_k) - F(a_k)). \tag{12.27}$$

*If $f_n \downarrow 0$ pointwise, then $\mu(f_n) \downarrow 0$.*

Proof: Say that $f_n \to 0$, where each $f_n$ is such a function. Notice that all the $f_n$ have supports in the interior of a fixed compact interval $[p, q]$. Furthermore, they are all bounded by some fixed constant $M$. Write $f_n = \sum_{k=1}^{m_n} c_{nk} 1_{(a_{nk}, b_{nk}]}$. For each $n$ and $k$ choose an interval $I_{nk} = (a_{nk}, a'_{nk}]$ such that the corresponding mass $F(a'_{nk}) - F(a_{nk})$ is bounded by $\epsilon/(2^n m_k)$. This can be done by using the right continuity of $F$ at each $a_{nk}$. Let $I_n$ be the union of the intervals $I_{nk}$, so the total mass associated with $I_n$ is $\epsilon/2^n$. For each $x$ in $[p,q]$ let $n_x$ be the first $n$ such that $f_n(x) < \epsilon$. Choose an open interval $V_x$ about $x$ such that $y$ in $V_x$ and $y$ not in $I_{n_x}$ implies $f_{n_x}(y) < \epsilon$. Now let $V_{x_1}, \ldots, V_{x_j}$ be a finite open subcover of $[p,q]$. Let $N$ be the maximum of $n_{x_1}, \ldots, n_{x_j}$. Then since the sequence of functions is monotone decreasing, $y$ in $[p,q]$ but $y$ not in the union of the $I_{n_{x_i}}$ implies $f(y) < \epsilon$. So $\sigma_F(f_N)$ is bounded by the total mass $F(q) - F(p)$ times the upper bound $\epsilon$ on the values plus a mass at most $\epsilon$ times the upper bound $M$ on the values. □

# Problems

1. Suppose the order preserving property $f \leq g$ implies $\mu(f) \leq \mu(g)$ is known for positive measurable functions. Show that it follows for all measurable functions, provided that the integrals exist. Hint: Decompose the functions into positive and negative parts.

2. Consider the space $\Omega = \{0,1\}^{\mathbf{N}_+}$ with the measure $\mu$ that describes fair coin tossing. Let $S_3$ be the random variable given by $S_3(\omega) = \omega_1 + \omega_2 + \omega_3$ that describes the number of heads in the first three tosses. Draw the graph of the corresponding function on the unit interval. Find the area under the graph, and check that this indeed gives the expectation of the random variable.

3. Let $\mu$ be an integral defined for Borel measurable functions on the real line. Let $\phi$ be a measurable map from the real line to itself. Then there is an image integral $\phi[\mu]$ defined by $\phi[\mu](f) = \mu(f \circ \phi)$. Let $\phi$ be defined by $\phi(x) = 3$. Find the image of Lebesgue measure $\lambda$ under $\phi$.

4. Consider the Gaussian (or normal) probability measure with expectation given by
$$\mu(f) = \int_{-\infty}^{\infty} f(z) \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} \, dz. \tag{12.28}$$
Let $c$ be in $\mathbb{R}$. Define $\phi_c(z) = e^{cz}$. For each $c$, find the forward image $\phi_c[\mu]$. (Note: This is called a log normal probability measure.)

   (a) If for some $c$ this forward image has a density with respect to Lebesgue measure, find it.

   (b) If for some $c$ it does not have a density with respect to Lebesgue measure, explain why not.

   (c) Are there distinct values of $c$ that give the same forward image? Explain.

5. An integral is $\sigma$-finite provided that there is a sequence of positive functions $f_n$ with finite integral such that $f_n \uparrow 1$. Must the image of a $\sigma$-finite integral under a measurable map $\phi$ be $\sigma$-finite? Discuss.

6. Let
$$\mu(g) = \int_{-\infty}^{\infty} g(t)w(t)\,dt \tag{12.29}$$
be a integral defined by a density $w(t) \geq 0$ with respect to Lebesgue measure $dt$. Let $s = \phi(t)$ define a smooth function that is piecewise strictly increasing or strictly decreasing or constant on open intervals separated by isolated points. Show that the image integral
$$\phi[\mu](f) = \int_{-\infty}^{\infty} f(s)h(s)\,ds + \sum_{s^*} f(s^*)c(s^*) \tag{12.30}$$
is given by a density $h(s)$ and perhaps also some point masses $c(s^*)\delta_{s^*}$. Here
$$h(s) = \sum_{\phi(t)=s} w(t)\frac{1}{|\phi'(t)|} \tag{12.31}$$
and
$$c(s^*) = \int_{\phi(t)=s^*} w(t)\,dt. \tag{12.32}$$

7. What is the increasing right continuous function that defines the integral
$$\mu(g) = \int_{-\infty}^{\infty} g(x)\frac{1}{\pi}\frac{1}{1+x^2}\,dx \tag{12.33}$$
involving the Cauchy density?

8. What is the increasing right continuous function that defines the $\delta_a$ integral given by $\delta_a(g) = g(a)$?

9. A subset of the line has Lebesgue measure zero if for every $\epsilon > 0$ it is a subset of a countable union of intervals of total length bounded by $\epsilon$. Prove that if $\phi : [0, L] \to \mathbb{R}$ is a $C^2$ function, and $S = \{x \mid \phi'(x) = 0\}$, then the image $\phi[S]$ of the subset $S$ under $\phi$ has measure zero. (This does not mean that $S$ has measure zero.) Hint: Suppose that for each $t$ we have $|\phi''(t)| \leq M$, so $|\phi(x') - \phi(x) - \phi'(x)(x' - x)| \leq (1/2)M(x' - x)^2$. Divide $[0, L]$ into $n$ closed subintervals each of length $L/n$. If $y = \phi(x)$ is in $\phi[S]$, then $x$ is in one of the intervals, and $\phi'(x) = 0$. What can you say about the size of the image of this interval?

# Chapter 13

# Convergence theorems

## 13.1 Convergence theorems

The most fundamental convergence theorem is improved monotone convergence. This was proved in the last chapter, but it is well to record it again here.

**Theorem 13.1** *(improved monotone convergence) If $\mu(f_1) > -\infty$ and $f_n \uparrow f$, then $\mu(f_n) \uparrow \mu(f)$. Similarly, if $\mu(h_1) < +\infty$ and $h_n \downarrow h$, then $\mu(h_n) \downarrow \mu(h)$.*

The next theorem is a consequence of monotone convergence that applies to a sequence of functions that is not monotone.

**Theorem 13.2 (Fatou's lemma)** *Suppose each $f_n \geq 0$. Let $f = \liminf_{n\to\infty} f_n$. Then*

$$\mu(f) \leq \liminf_{n\to\infty} \mu(f_n). \tag{13.1}$$

Proof: Let $r_n = \inf_{k\geq n} f_k$. It follows that $0 \leq r_n \leq f_k$ for each $k \geq n$. So $0 \leq \mu(r_n) \leq \mu(f_k)$ for each $k \geq n$. This gives the inequality

$$0 \leq \mu(r_n) \leq \inf_{k\geq n} \mu(f_k). \tag{13.2}$$

However $0 \leq r_n \uparrow f$. By monotone convergence $0 \leq \mu(r_n) \uparrow \mu(f)$. Therefore passing to the limit in the inequality gives the result. □

Fatou's lemma says that in the limit one can lose positive mass density, but one cannot gain it.

Examples:

1. Consider functions $f_n = n1_{(0,1/n)}$ on the line. It is clear that $\lambda(f_n) = 1$ for each $n$. On the other hand, $f_n \to 0$ pointwise, and $\lambda(0) = 0$. The density has formed a spike near the origin, and this does not produce a limiting density.

2. Consider functions $f_n = 1_{(n,n+1)}$. It is clear that $\lambda(f_n) = 1$ for each $n$. On the other hand, $f_n \to 0$ pointwise, and $\lambda(0) = 0$. The density has moved off to $+\infty$ and is lost in the limit.

It is natural to ask where the mass has gone. The only way to answer this is to reinterpret the problem as a problem about measure. Define the measure $\nu_n(\phi) = \lambda(\phi f_n)$. Take $\phi$ bounded and continuous. Then it is possible that $\nu_n(\phi) \to \nu(\phi)$ as $n$ to $\infty$. If this happens, then $\nu$ may be interpreted as a limiting measure that contains the missing mass. However this measure need not be given by a density.

Examples:

1. Consider functions $nf_n = 1_{(0,1/n)}$ on the line. In this case $\nu_n(\phi) = \lambda(\phi f_n) \to \phi(0) = \delta_0(\phi)$. The limiting measure is a point mass at the origin.

2. Consider functions $f_n = 1_{(n,n+1)}$. Suppose that we consider continuous functions with right and left hand limits at $+\infty$ and $-\infty$. In this case $\nu_n(\phi) = \lambda(\phi f_n) \to \phi(+\infty) = \delta_{+\infty}(\phi)$. The limiting measure is a point mass at $+\infty$.

**Theorem 13.3 (dominated convergence)** *Let $|f_n| \leq g$ for each $n$, where $g$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$, that is, $\mu(g) < \infty$. Suppose $f_n \to f$ pointwise as $n \to \infty$. Then $f$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$ and $\mu(f_n) \to \mu(f)$ as $n \to \infty$.*

This theorem is amazing because it requires only pointwise convergence. The only hypothesis is the existence of the dominating function

$$\forall n \forall x\, |f_n(x)| \leq g(x) \tag{13.3}$$

with

$$\int g(x)\, d\mu(x) < +\infty. \tag{13.4}$$

Then pointwise convergence

$$\forall x \lim_{n\to\infty} f_n(x) = f(x) \tag{13.5}$$

implies convergence of the integrals

$$\lim_{n\to\infty} \int f_n(x)\, d\mu(x) = \int f(x)\, d\mu(x). \tag{13.6}$$

Proof: We have $|f_k| \leq g$, so $-g \leq f_k \leq g$. Let $r_n = \inf_{k\geq n} f_k$ and $s_n = \sup_{k\geq n} f_k$. Then

$$-g \leq r_n \leq f_n \leq s_n \leq g. \tag{13.7}$$

This gives the inequality

$$-\infty < -\mu(g) \leq \mu(r_n) \leq \mu(f_n) \leq \mu(s_n) \leq \mu(g) < +\infty. \tag{13.8}$$

However $r_n \uparrow f$ and $s_n \downarrow f$. It follows from improved monotone convergence that $\mu(r_n) \uparrow \mu(f)$ and $\mu(s_n) \downarrow \mu(f)$. It follows from the inequality that $\mu(f_n) \to \mu(f)$. □

**Corollary 13.4** *Let $|f_n| \leq g$ for each $n$, where $g$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$. It follows that each $f_n$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$. Suppose $f_n \to f$ pointwise as $n \to \infty$. Then $f$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$ and $f_n \to f$ in the sense that $\mu(|f_n - f|) \to 0$ as $n \to \infty$.*

Proof: It suffices to apply the dominated convergence theorem to $|f_n - f| \leq 2g$. □

In applying the dominated convergence theorem, the function $g \geq 0$ must be independent of $n$ and have finite integral. However there is no requirement that the convergence be uniform or monotone.

Here is a simple example. Consider the sequence of functions $f_n(x) = \cos^n(x)/(1 + x^2)$. The goal is to prove that $\lambda(f_n) = \int_{-\infty}^{\infty} f_n(x)\,dx \to 0$ as $n \to \infty$. Note that $f_n \to 0$ as $n \to \infty$ pointwise, except for points that are a multiple of $\pi$. At these points one can redefine each $f_n$ to be zero, and this will not change the integrals. Apply the dominated convergence to the redefined $f_n$. For each $n$ we have $|f_n(x)| \leq g(x)$, where $g(x) = 1/(1 + x^2)$ has finite integral. Hence $\lambda(f_n) \to \lambda(0) = 0$ as $n \to \infty$.

The following examples show what goes wrong when the condition that the dominating function has finite integral is not satisfied.

Examples:

1. Consider functions $f_n = n1_{(0,1/n)}$ on the line. These are dominated by $g(x) = 1/x$ on $0 < x \leq 1$, with $g(x) = 0$ for $x \geq 0$. This is independent of $n$. However $\lambda(g) = \int_0^1 1/x\,dx = +\infty$. The dominated convergence does not apply, and the integral of the limit is not the limit of the integrals.

2. Consider functions $f_n = 1_{(n,n+1)}$. Here the obvious dominating function is $g = 1_{(0,+\infty)}$. However again $\lambda(g) = +\infty$. Thus there is nothing to prevent mass density being lost in the limit.

## 13.2 Measure

If $E$ is a subset of $X$, then $1_E$ is the indicator function of $E$. Its value is 1 for every point in $E$ and 0 for every point not in $E$. The set $E$ is said to be measurable if the function $1_E$ is measurable. The *measure* of such an $E$ is $\mu(1_E)$. This is often denoted $\mu(E)$.

**Theorem 13.5** *An integral is uniquely determined by the corresponding measure.*

Proof: Let $f \geq 0$ be a measurable function. Define

$$f_n = \sum_{k=0}^{\infty} \frac{k}{2^n} 1_{\frac{k}{2^n} < f \leq \frac{k+1}{2^n}}. \tag{13.9}$$

The integral of $f_n$ is determined by the measures of the sets where $\frac{k}{2^n} < f \leq \frac{k+1}{2^n}$. However $f_n \uparrow f$, and so the integral of $f$ is determined by the corresponding measure. □

This theorem justifies a certain amount of confusion between the notion of measure and the notion of integral. In fact, this whole subject is sometimes called measure theory.

Sometimes we denote a subset of $X$ by a condition that defines the subset. Thus, for instance, $\{x \mid f(x) > a\}$ is denoted $f > a$, and its measure is $\mu(f > a)$.

**Theorem 13.6** *If the set where $f \neq 0$ has measure zero, then $\mu(|f|) = 0$.*

Proof: For each $n$ the function $|f| \wedge n \leq n 1_{|f|>0}$ and so has integral $\mu(|f| \wedge n) \leq n \cdot 0 = 0$. However $|f| \wedge n \uparrow |f|$ as $n \to \infty$. So from monotone convergence $\mu(|f|) = 0$. □

The preceding theorem shows that changing a function on a set of measure zero does not change its integral. Thus, for instance, if we change $g_1$ to $g_2 = g_1 + f$, then $|\mu(g_2) - \mu(g_1)| = |\mu(f)| \leq \mu(|f|) = 0$, so $\mu(g_1) = \mu(g_2)$.

There is a terminology that is standard in this situation. If a property of points is true except on a subset of $\mu$ measure zero, then it is said to hold *almost everywhere* with respect to $\mu$. Thus the theorem would be stated as saying that if $f = 0$ almost everywhere, then its integral is zero. Similarly, if $g = h$ almost everywhere, then $g$ and $h$ have the same integral.

In probability the terminology is slightly different. Instead of saying that a property holds almost everywhere, on says that the event happens *almost surely* or *with probability one*.

The convergence theorems hold even when the hypotheses are violated on a set of measure zero. For instance, the dominated convergence theorem can be stated: If $|g| \leq g$ almost everywhere with respect to $\mu$ and $\mu(g) < +\infty$, then $f_n \to f$ almost everywhere with respect to $\mu$ implies $\mu(f_n) \to \mu(f)$.

**Theorem 13.7 (Chebyshev inequality)** *Let $f$ be a real measurable function and $a$ be a real number. Let $\phi$ be an increasing real function on $[a, +\infty)$ with $\phi(a) > 0$ and $\phi \geq 0$ on the range of $f$. Then*

$$\mu(f \geq a) \leq \frac{1}{\phi(a)} \mu(\phi(f)). \tag{13.10}$$

Proof: This follows from the pointwise inequality

$$1_{f \geq a} \leq 1_{\phi(f) \geq \phi(a)} \leq \frac{1}{\phi(a)} \phi(f). \tag{13.11}$$

At the points where $f \geq a$ we have $\phi(f) \geq \phi(a)$ and so the right hand size is one or greater. In any case the right hand size is positive. Integration preserves the inequality. □

The Chebyshev inequality is used in practice mainly in certain important special cases. Thus for $a > 0$ we have

$$\mu(|f| \geq a) \leq \frac{1}{a}\mu(|f|) \tag{13.12}$$

and

$$\mu(|f| \geq a) \leq \frac{1}{a^2}\mu(f^2). \tag{13.13}$$

Another important case is when $t > 0$ and

$$\mu(f \geq a) \leq \frac{1}{e^{ta}}\mu(e^{tf}). \tag{13.14}$$

**Theorem 13.8** *If $\mu(|f|) = 0$, then the set where $f \neq 0$ has measure zero.*

Proof: By the Chebyshev inequality, for each $n$ we have $\mu(1_{|f|>1/n}) \leq n\mu(|f|) = n \cdot 0 = 0$. However as $n \to \infty$, the functions $1_{|f|>1/n} \uparrow 1_{|f|>0}$. So $\mu(1_{|f|>0}) = 0$. □

The above theorem also has a statement in terms of an almost everywhere property. It says that if $|f|$ has integral zero, then $f = 0$ almost everywhere.

## 13.3 Extended real valued measurable functions

In connection with Tonelli's theorem it is natural to look at functions with values in the set $[0, +\infty]$. This system is well behaved under addition. In the context of measure theory it is useful to define $0 \cdot (+\infty) = (+\infty) \cdot 0 = 0$. It turns out that this is the most useful definition of multiplication.

Let $X$ be a non-empty set, and let $\mathcal{F}$ be a $\sigma$-algebra of real functions on $X$. A function $f : X \to [0, +\infty]$ is said to be measurable with respect to $\mathcal{F}$ if there is a sequence $f_n$ of functions in $\mathcal{F}$ with $f_n \uparrow f$ pointwise. A function is measurable in this sense if and only if there is a measurable set $A$ with $f = +\infty$ on $A$ and $f$ coinciding with a function in $\mathcal{F}$ on the complement $A^c$.

An integral $\mu : \mathcal{F}^+ \to [0, +\infty]$ is extended to such measurable functions $f$ by monotone convergence. Notice that if $A$ is the set where $f = +\infty$, then we can set $f_n = n$ on $A$ and $f$ on $A_c$. Then $\mu(f_n) = n\mu(A) + \mu(f1_{A_c})$. If we take $n \to \infty$, we get $\mu(f) = (+\infty)\mu(A) + \mu(f1_{A^c})$. For the monotone convergence theorem to hold we must interpret $(+\infty) \cdot 0 = 0$. Notice that if $\mu(f) < +\infty$, then it follows that $\mu(A) = 0$.

## 13.4 Fubini's theorem for sums and integrals

**Theorem 13.9 (Tonelli for sums of functions)** *If $w_k \geq 0$, then*

$$\mu(\sum_{k=1}^{\infty} w_k) = \sum_{k=1}^{\infty} \mu(w_k). \tag{13.15}$$

Proof: This theorem says that for positive functions integrals and sums may be interchanged. This is the monotone convergence theorem in disguise. That is, let $f_n = \sum_{k=1}^{n} w_k$. Then $f_n \uparrow f = \sum_{k=1}^{\infty} w_k$. Hence $\mu(f_n) = \sum_{k=1}^{n} \mu(w_k) \uparrow \mu(f)$. □

**Theorem 13.10 (Fubini for sums of functions)** *Suppose that the condition $\sum_{k=1}^{\infty} \mu(|w_k|) < +\infty$ is satisfied. Set $g = \sum_{k=1}^{\infty} |w_k|$. Then $g$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$ and so the set $\Lambda$ where $g < +\infty$ has $\mu(\Lambda^c) = 0$. On this set $\Lambda$ let*

$$f = \sum_{k=1}^{\infty} w_k \tag{13.16}$$

*and on $\Lambda^c$ set $f = 0$. Then $f$ is in $\mathcal{L}^1(X, \mathcal{F}, \mu)$ and*

$$\mu(f) = \sum_{k=1}^{\infty} \mu(w_k). \tag{13.17}$$

*In other words,*

$$\int_{\Lambda} \sum_{k=1}^{\infty} w_k \, d\mu = \sum_{k=1}^{\infty} \int w_k \, d\mu. \tag{13.18}$$

Proof: This theorem says that absolute convergence implies that integrals and sums may be interchanged. Here is a first proof. By the hypothesis and Tonelli s theorem $\mu(g) < +\infty$. It follows that $g < +\infty$ on a set $\Lambda$ whose complement has measure zero. Let $f_n = \sum_{k=1}^{n} 1_{\Lambda} w_k$. Then $|f_n| \leq g$ for each $n$. Furthermore, the series defining $f$ is absolutely convergent on $\Lambda$ and hence convergent on $\Lambda$. Thus $f_n \to f$ as $n \to \infty$. Furthermore $\mu(f_n) = \sum_{k=1}^{n} \mu(1_{\Lambda} w_k) = \sum_{k=1}^{n} \mu(w_k)$. The conclusion follows by the dominated convergence theorem. □

Proof: Here is a second proof. Decompose each $w_j = w_j^+ - w_j^-$ into a positive and negative part. Then by Tonelli's theorem $\mu(\sum_{j=1}^{\infty} w_j^{\pm}) < +\infty$. Let $\Lambda$ be the set where both sums $\sum_{j=1}^{\infty} w_j^{\pm} < +\infty$. Then $\mu(\Lambda^c) = 0$. Let $f = \sum_{j=1}^{\infty} w_j$ on $\Lambda$ and $f = 0$ on $\Lambda^c$. Then $f = \sum_{j=1}^{\infty} 1_{\Lambda} w_j^+ - \sum_{j=1}^{\infty} 1_{\Lambda} w_j^-$. Therefore $\mu(f) = \mu(\sum_{j=1}^{\infty} 1_{\Lambda} w_j^+) - \mu(\sum_{j=1}^{\infty} 1_{\Lambda} w_j^-) = \sum_{j=1}^{\infty} \mu(w_j^+) - \sum_{j=1}^{\infty} \mu(w_j^-) = \sum_{j=1}^{\infty}(\mu(w_j^+) - \mu(w_j^-)) = \sum_{j=1}^{\infty} \mu(w_j)$. The hypothesis guarantees that there is never a problem with $(+\infty) - (+\infty)$. □

## 13.5 Fubini's theorem for sums

The following two theorems give conditions for when sums may be interchanged. Usually these results are applied when the sums are both over countable sets. However the case when one of the sums is uncountable also follows from the corresponding theorems in the preceding section.

**Theorem 13.11 (Tonelli for sums)** *If $w_k(x) \geq 0$, then*

$$\sum_x \sum_{k=1}^{\infty} w_k(x) = \sum_{k=1}^{\infty} \sum_x w_k(x). \tag{13.19}$$

**Theorem 13.12 (Fubini for sums)** *Suppose that the condition $\sum_{k=1}^{\infty} \sum_x |w_k(x)| < +\infty$ is satisfied. Then for each $x$ the series $\sum_{k=1}^{\infty} w_k(x)$ is absolutely summable, and*

$$\sum_x \sum_{k=1}^{\infty} w_k(x) = \sum_{k=1}^{\infty} \sum_x w_k(x). \tag{13.20}$$

Here is an example that shows why absolute convergence is essential. Let $g : \mathbf{N} \times \mathbf{N} \to \mathbb{R}$ be defined by $g(m, n) = 1$ if $m = n$ and $g(m, n) = -1$ if $m = n + 1$. Then

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} g(m, n) = 0 \neq 1 = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} g(m, n). \tag{13.21}$$

# Problems

1. Can the plane be represented as a countable union of circles (of varying radii)? Justify your answer.

2. This problem is to show that one can get convergence theorems when the family of functions is indexed by real numbers. Prove that if $f_t \to f$ pointwise as $t \to t_0$, $|f_t| \leq g$ pointwise, and $\mu(g) < \infty$, then $\mu(f_t) \to \mu(f)$ as $t \to t_0$.

3. Show that if $f$ is a Borel function and $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$, then $F(b) = \int_{-\infty}^{b} f(x)\, dx$ is continuous.

4. Must the function $F$ in the preceding problem be differentiable at every point? Discuss.

5. Show that
$$\int_0^{\infty} \frac{\sin(e^x)}{1 + nx^2}\, dx \to 0 \tag{13.22}$$
as $n \to \infty$.

6. Show that
$$\int_0^{1} \frac{n \cos(x)}{1 + n^2 x^{\frac{3}{2}}}\, dx \to 0 \tag{13.23}$$
as $n \to \infty$.

7. Evaluate
$$\lim_{n \to \infty} \int_a^{\infty} \frac{n}{1 + n^2 x^2}\, dx \tag{13.24}$$
as a function of $a$.

8. Consider the integral

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{1+nx^2}}\,dx. \tag{13.25}$$

Show that the integrand is monotone decreasing and converges pointwise as $n \to \infty$, but the integral of the limit is not equal to the limit of the integrals. How does this relate to the monotone convergence theorem?

9. Let $f \geq 0$ satisfy $\int_0^\infty f(x)\,dx < +\infty$. Evaluate

$$\lim_{n\to\infty} \int_0^\infty x^n f(x)\,dx. \tag{13.26}$$

There are several possible answers: discuss all cases.

10. Let $g$ be a Borel function with

$$\int_{-\infty}^{\infty} |g(x)|\,dx < \infty \tag{13.27}$$

and

$$\int_{-\infty}^{\infty} g(x)\,dx = 1 \tag{13.28}$$

Let

$$g_\epsilon(x) = g(\frac{x}{\epsilon})\frac{1}{\epsilon}. \tag{13.29}$$

Let $\phi$ be bounded and continuous. Show that

$$\int_{-\infty}^{\infty} g_\epsilon(y)\phi(y)\,dy \to \phi(0) \tag{13.30}$$

as $\epsilon \to 0$. This problem gives a very general class of functions $g_\epsilon(x)$ such that integration with $g_\epsilon(x)\,dx$ converges to the Dirac delta integral $\delta_0$ given by $\delta_0(\phi) = \phi(0)$.

11. Let $f$ be bounded and continuous. Show that for each $x$ the convolution

$$\int_{-\infty}^{\infty} g_\epsilon(x-z)f(z)\,dz \to f(x) \tag{13.31}$$

as $\epsilon \to 0$.

12. Prove *countable subadditivity*:

$$\mu(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n). \tag{13.32}$$

Show that if the $A_n$ are disjoint this is an equality (countable additivity). Hint: $1_{\bigcup_{n=1}^\infty A_n} \leq \sum_{n=1}^\infty 1_{A_n}$.

13. Egoroff's theorem. Let $n \mapsto f_n$ be a sequence of measurable functions defined on a finite measure space $X, \mu$. Suppose that $f_n \to f$ pointwise as $n \to \infty$. Show that for every $a > 0$ there is a set $E$ with measure $\mu(E) \geq \mu(X) - a$ and a sequence $k \mapsto N_k$ defined for $k = 1, 2, 3, \ldots$ such that for each $x$ in $E$ we have $\forall k \, \forall n \geq N_k \, |f_n(x) - f(x)| < 1/k$. Since $N_k$ depends only on $k$, the result says that $f_n \to f$ uniformly on $E$.

    Hint: Let $E_{kN} = \{x \mid \forall n \geq N \, |f_n(x) - f(x)| < 1/k$. Consider $a > 0$. Prove that $\lim_{N \to \infty} \mu(E_{kN_k} = 1$; be explicit about what convergence theorem you use. Then for $N$ sufficiently large $\mu(E_{kN}) \geq 1 - a/2^k$. Define $N_k$ with $\mu(E_{kN_k}) \geq a/2^k$.

# Chapter 14

# Fubini's theorem

## 14.1 Introduction

As an introduction, consider the Tonelli and Fubini theorems for Borel functions of two variables.

**Theorem 14.1 (Tonelli)** *If $f(x, y) \geq 0$, then*

$$\int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(x, y)\, dx \right] dy = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(x, y)\, dy \right] dx. \tag{14.1}$$

**Theorem 14.2 (Fubini)** *If*

$$\int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} |f(x, y)|\, dx \right] dy < +\infty, \tag{14.2}$$

*then*

$$\int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(x, y)\, dx \right] dy = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} f(x, y)\, dy \right] dx. \tag{14.3}$$

A slightly more careful statement of Fubini's theorem would acknowledge that the inner integrals may not be defined. However let

$$\Lambda_1 = \{x \mid \int_{-\infty}^{\infty} |f(x, y)|\, dy < +\infty\} \tag{14.4}$$

and

$$\Lambda_2 = \{y \mid \int_{-\infty}^{\infty} |f(x, y)|\, dx < +\infty\} \tag{14.5}$$

Then the inner integrals are well-defined on these sets. Furthermore, by the hypothesis of Fubini's theorem and by Tonelli's theorem, the complements of these sets have measure zero. So a more precise statement of the conclusion of Fubini's theorem is that

$$\int_{\Lambda_2} \left[ \int_{-\infty}^{\infty} f(x, y)\, dx \right] dy = \int_{\Lambda_1} \left[ \int_{-\infty}^{\infty} f(x, y)\, dy \right] dx. \tag{14.6}$$

This just amounts to replacing the undefined inner integrals by zero on the troublesome sets that are the complements of $\Lambda_1$ and $\Lambda_2$. It is quite fortunate that these sets are of measure zero.

The Tonelli and Fubini theorems may be formulated in a way that does not depend on writing the variables of integration explicitly. Consider for example Tonelli's theorem, which applies to a positive measurable function $f$ on the plane. Let $f^{|1}$ be the function on the line whose value at a real number is obtained by holding the first variable fixed at this number and looking at $f$ as a function of the second variable. Thus the value $f^{|1}(x)$ is the function $y \mapsto f(x, y)$. Similarly, let $f^{|2}$ be the function on the line whose value at a real number is obtained by holding the second variable fixed at this number and looking at $f$ as a function of the first variable. The value $f^{|2}(y)$ is the function $x \mapsto f(x, y)$. Then the inner integrals are $(\lambda \circ f^{|2})(y) = \lambda(f^{|2}(y)) = \int_{-\infty}^{\infty} f(x, y)\, dx$ and $(\lambda \circ f^{|1})(x) = \lambda(f^{|1}(x)) = \int_{-\infty}^{\infty} f(x, y)\, dy$. So $\lambda \circ f^{|2}$ and $\lambda \circ f^{|1}$ are each a positive measurable function on the line. The conclusion of Tonelli's theorem may then be stated as the equality $\lambda(\lambda \circ f^{|2}) = \lambda(\lambda \circ f^{|1})$.

Here is rather interesting example where the hypothesis and conclusion of Fubini's theorem are both violated. Let $\sigma^2 > 0$ be a fixed diffusion constant. Let

$$u(x, t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp(-\frac{x^2}{2\sigma^2 t}). \tag{14.7}$$

This describes the diffusion of a substance that has been created at time zero at the origin. For instance, it might be a broken bottle of perfume, and the molecules of perfume each perform a kind of random walk, moving in an irregular way. The motion is so irregular that the average squared distance that a particle moves in time $t$ is only $x^2 = \sigma^2 t$.

As time increases the density gets more and more spread out. Then $u$ satisfies

$$\frac{\partial u}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 u}{\partial x^2}. \tag{14.8}$$

Note that

$$\frac{\partial u}{\partial x} = -\frac{x}{\sigma^2 t} u \tag{14.9}$$

and

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{\sigma^2 t}(\frac{x^2}{\sigma^2 t} - 1)u. \tag{14.10}$$

This says that $u$ is increasing in the space time region $x^2 > \sigma^2 t$ and decreasing in the space-time region $x^2 < \sigma^2 t$.

Fix $s > 0$. It is easy to compute that

$$\int_s^\infty \int_{-\infty}^\infty \frac{\partial u}{\partial t}\, dx\, dt = \frac{\sigma^2}{2} \int_s^\infty \int_{-\infty}^\infty \frac{\partial^2 u}{\partial x^2}\, dx\, dt = 0 \tag{14.11}$$

and

$$\int_{-\infty}^\infty \int_s^\infty \frac{\partial u}{\partial t}\, dt\, dx = -\int_{-\infty}^\infty u(x, s)\, dx = -1. \tag{14.12}$$

One can stop at this point, but it is interesting to look at the mechanism of the failure of the Fubini theorem. It comes from the fact that the time integral is extended to infinity, and in this limit the density spreads out more and more and approaches zero pointwise. So mass is lost in this limit, at least if one tries to describe it as a density. A description of the mass as a measure might lead instead to the conclusion that the mass is sitting at $x = \pm\infty$ in the limit $t \to \infty$. Even this does not capture the essence of the situation, since the diffusing particles do not go to infinity in any systematic sense; they just wander more and more.

The Tonelli and Fubini theorems are true for the Lebesgue integral defined for Borel functions on the line. However they are not true for arbitrary integrals that are not required to be $\sigma$-finite. Here is an example based on the example of summation over an uncountable set.

Let $\lambda(g) = \int_0^1 g(x)\,dx$ be the usual uniform Lebesgue integral on the interval $[0, 1]$. Let $\sum h = \sum_y h(y)$ be summation indexed by the points in the interval $[0, 1]$. The measure $\sum$ is not $\sigma$-finite, since there are uncountably many points in $[0, 1]$. Finally, let $\delta_{xy} = 1$ if $x = y$, and $\delta_{xy} = 0$ for $x \neq y$. Now for each $x$, the sum $\sum_y \delta_{xy} = 1$. So the integral over $x$ is also 1. On the other hand, for each $y$ the integral $\int_0^1 \delta_{xy}\,dx = 0$, since the integrand is zero except for a single point of $\lambda$ measure zero, where it has the value one. So the sum over $y$ is also zero. Thus the two orders of integration give different results.

## 14.2 Product sigma-algebras

This section defines the product $\sigma$-algebra. Let $X_1$ and $X_2$ be non-empty sets. Then their product $X_1 \times X_2$ is another non-empty set. There are projections $\pi_1 : X_1 \times X_2 \to X_1$ and $\pi_2 : X_1 \times X_2 \to X_2$. These are of course define by $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$.

Suppose that $\mathcal{F}_1$ is a $\sigma$-algebra of real functions on $X_1$ and $\mathcal{F}_2$ is a $\sigma$-algebra of real functions on $X_2$. Then there is a *product* $\sigma$-algebra $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ of real functions on $X_1 \times X_2$. This is the smallest $\sigma$-algebra of functions on $\mathcal{F}_1 \times \mathcal{F}_2$ such that the projections $\pi_1$ and $\pi_2$ are measurable maps.

The condition that the projections $\pi_1$ and $\pi_2$ are measurable maps is the same as saying that for each $g$ in $\mathcal{F}_1$ the function $g \circ \pi_1$ is measurable and for each $h$ in $\mathcal{F}_2$ the function $h \circ \pi_2$ is measurable. In other words, the functions $(x, y) \mapsto g(x)$ and $(x, y) \mapsto h(y)$ are required to be measurable functions. This condition determines the $\sigma$-algebra of measurable functions $\mathcal{F}_1 \bigotimes \mathcal{F}_2$.

If $g$ is a real function on $X_1$ and $h$ is a real function on $X_2$, then there is a real function $g \otimes h$ on $X$ defined by

$$(g \otimes h)(x, y) = g(x)h(y). \tag{14.13}$$

This is sometimes called the tensor product of the two functions. Such functions are called *decomposable*. Another term is *separable*, as in "separation of variables." The function $g \otimes h$ could be define more abstractly as $g \otimes h =$

$(g \circ \pi_1)(h \circ \pi_2)$. This identity could also be stated as $g \otimes h = (g \otimes 1)(1 \otimes h)$. It is easy to see that $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ may also be characterized as the $\sigma$-algebra generated by the functions $g \otimes h$ with $g$ in $\mathcal{F}_1$ and $h$ in $\mathcal{F}_2$.

Examples:

1. If $\mathcal{B}o$ is the Borel $\sigma$-algebra of functions on the line, then $\mathcal{B}o \bigotimes \mathcal{B}o$ is the Borel $\sigma$-algebra of functions on the plane.

2. Take the two sigma-algebras to be the Borel $\sigma$-algebra of real functions on $[0, 1]$ and the $\sigma$-algebra $\mathbb{R}^{[0,1]}$ of all real functions on $[0, 1]$. These are the $\sigma$-algebras relevant to the counterexample with $\lambda$ and $\sum$. The product $\sigma$-algebra then consists of all functions $f$ on the square such that $x \mapsto f(x, y)$ is a Borel function for each $y$. The diagonal function $\delta$ is measurable, but $\sum$ is not $\sigma$-finite, so Tonelli's theorem does not apply.

3. Take the two sigma-algebras to be the Borel $\sigma$-algebra of real functions on $[0, 1]$ and the $\sigma$-algebra consisting of all real functions $y \mapsto a+h(y)$ on $[0, 1]$ that differ from a constant function $a$ on a countable set. These are the $\sigma$-algebras relevant to the counterexample with $\lambda$ and $\sum$, but in the case when we restrict $\sum$ to the smallest $\sigma$-algebra for which it makes sense. The product $\sigma$-algebra is generated by functions of the form $(x, y) \mapsto g(x)$ and $(x, y) \mapsto a+h(y)$, where $h$ vanishes off a countable set. This is a rather small $\sigma$-algebra; the diagonal function $\delta$ used in the counterexample does not belong to it. Already for this reason Tonelli's theorem cannot be used.

**Lemma 14.3** *Let $X_1$ be a set with $\sigma$-algebra $\mathcal{F}_1$ of functions and $\sigma$-finite integral $\mu_1$. Let $X_2$ be another set with a $\sigma$-algebra $\mathcal{F}_2$ of functions and $\sigma$-finite integral $\mu_2$. Let $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ be the product $\sigma$-algebra of functions on $X_1 \times X_2$. Let $L$ consist of finite linear combinations of indicator functions of products of sets of finite measure. Then $L$ is a vector lattice, and the smallest monotone class including $L$ is $\mathcal{F}_1 \bigotimes \mathcal{F}_2$.*

Proof: Let $L \subset \mathcal{F}_1 \bigotimes \mathcal{F}_2$ be the set of all finite linear combinations

$$f = \sum_i c_i 1_{A_i \times B_i} = \sum_i c_i 1_{A_i} \otimes 1_{B_i}, \tag{14.14}$$

where $A_i$ and $B_i$ each have finite measure. The space $L$ is obviously a vector space. The proof that it is a lattice is found in the last section of the chapter.

Let $E_n$ be a sequence of sets of finite measure that increase to $X_1$. Let $F_n$ be a sequence of sets of finite measure that increase to $X_2$. Then the $E_n \times F_n$ increase to $X_1 \times X_2$. This is enough to show that the the constant functions belong to the monotone class generated by $L$. Since $L$ is a vector lattice and the monotone class generated by $L$ has all constant functions, it follows that the monotone class generated by $L$ is a $\sigma$-algebra. To show that this $\sigma$-algebra is equal to all of $\mathcal{F}_1 \bigotimes \mathcal{F}_2$, it is sufficient to show that each $g \otimes h$ is in the $\sigma$-algebra generated by $L$. Let $g_n = g1_{E_n}$ and $h_n = h1_{F_n}$. It is sufficient to show that each $g_n \otimes h_n$ is in this $\sigma$-algebra. However $g_n$ may be approximated by functions of the form $\sum_i a_i 1_{A_i}$ with $A_i$ of finite measure, and $h_n$ may also be approximated by functions of the form $\sum_j b_j 1_{B_j}$ with $B_j$ of finite measure. So $g_n \otimes h_n$ is approximated by $\sum_i \sum_j a_i b_j 1_{A_i} \otimes 1_{B_j} = \sum_i \sum_j a_i b_j 1_{A_i \times B_j}$. These are indeed functions in $L$. □

## 14.3 The product integral

This section gives a proof of the uniqueness of the product of two $\sigma$-finite integrals.

**Theorem 14.4** *Let $\mathcal{F}_1$ be a $\sigma$-algebra of measurable functions on $X_1$, and let $\mathcal{F}_2$ be a $\sigma$-algebra of measurable functions on $X_2$. Let $\mu_1 : \mathcal{F}_1^+ \to [0, +\infty]$ and $\mu_2 : \mathcal{F}_2^+ \to [0, +\infty]$ be corresponding $\sigma$-finite integrals. Consider the product space $X_1 \times X_2$ and the product $\sigma$-algebra of functions $\mathcal{F}_1 \bigotimes \mathcal{F}_2$. Then there exists at most one $\sigma$-finite integral $\nu : (\mathcal{F}_1 \bigotimes \mathcal{F}_2)^+ \to [0, +\infty]$ with the property that if $A$ and $B$ each have finite measure, then $\nu(A \times B) = \mu_1(A)\mu_2(B)$.*

Proof: Let $L$ be the vector lattice of the preceding lemma. The integral $\nu$ is uniquely defined on $L$ by the explicit formula. Since the smallest monotone class including $L$ is $\mathcal{F}_1 \bigotimes \mathcal{F}_2$, it follows that the smallest $L$-monotone class including $L^+$ is $(\mathcal{F}_1 \bigotimes \mathcal{F}_2)^+$. Say that $\nu$ and $\nu'$ were two such integrals. Then they agree on $L$, since they are given by an explicit formula. However the set of functions on which they agree is an $L$-monotone class. Therefore the integral is uniquely determined on all of $\mathcal{F}^+$. □

The integral $\nu$ described in the above theorem is called the *product integral* and denoted $\mu_1 \times \mu_2$. The corresponding measure is called the *product measure*. The existence of the product of $\sigma$-finite integrals will be a byproduct of the Tonelli theorem. This product integral $\nu$ has the more general property that if $g \geq 0$ is in $\mathcal{F}_1$ and $h \geq 0$ is in $\mathcal{F}_2$, then

$$\nu(g \otimes h) = \mu_1(g)\mu_2(h). \tag{14.15}$$

The product of integrals may be of the form $0 \cdot (+\infty)$ or $(+\infty) \cdot 0$. In that case the multiplication is performed using $0 \cdot (+\infty) = (+\infty) \cdot 0 = 0$. The characteristic property $(\mu_1 \times \mu_2)(g \otimes h) = \mu_1(g)\mu_2(h)$ may also be written in the more explicit form

$$\int g(x)h(y)\, d(\mu_1 \times \mu_2)(x, y) = \int g(x)\, d\mu_1(x) \int h(y)\, d\mu_2(y). \tag{14.16}$$

The definition of product integral does not immediately give a useful way to compute the integral of functions that are not written as sums of decomposable functions. For this we need Tonelli's theorem and Fubini's theorem.

## 14.4 Tonelli's theorem

Let $X_1$ and $X_2$ be two sets. Let $f : X_1 \times X_2 \to \mathbb{R}$ be a function on the product space. For each $x$ in $X_1$ there is a *section* (or slice) function $y \mapsto f(x, y)$. Then there is a function $f^{|1}$ from $X_1$ to $\mathbb{R}^{X_2}$ defined by saying that the value $f^{|1}(x)$ is the function $y \mapsto f(x, y)$. In other words, $f^{|1}$ is $f$ with the first variable temporarily held constant.

Similarly, for each $y$ in $X_2$ there is a section function $x \mapsto f(x, y)$. Hence there is a function $f^{|2}$ from $X_2$ to $\mathbb{R}^{X_1}$ defined by saying that the value $f^{|2}(y)$ is the function $x \mapsto f(x, y)$. In other words, $f^{|2}$ is $f$ with the second variable temporarily held constant.

**Lemma 14.5** *Let $f : X_1 \times X_2 \to [0, +\infty]$ be a $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ measurable function. Then for each $x$ the function $f^{|1}(x)$ is a $\mathcal{F}_2$ measurable function on $X_2$. Also, for each $y$ the function $f^{|2}(y)$ is a $\mathcal{F}_1$ measurable function on $X_1$.*

Explicitly, this lemma says that the functions

$$y \mapsto f(x, y) \tag{14.17}$$

with fixed $x$ and

$$x \mapsto f(x, y) \tag{14.18}$$

with fixed $y$ are measurable functions.

Proof: Let $L$ be the space of finite linear combinations of indicator functions of products of sets of finite measure. Consider the class $S$ of functions $f$ for which the lemma holds. If $f$ is in $L$, then $f = \sum_i c_i 1_{A_i \times B_i}$, where each $A_i$ is an $\mathcal{F}_1$ set and each $B_i$ is a $\mathcal{F}_2$ set. Then for fixed $x$ consider the function $y \mapsto \sum_i c_i 1_{A_i}(x) 1_{B_i}(y)$. This is clearly in $\mathcal{F}_2$. This shows that $L \subset S$. Now suppose that $f_n \uparrow f$ and each $f_n$ is in $S$. Then for each $x$ we have that $f_n(x, y)$ is measurable in $y$ and increases to $f(x, y)$ pointwise in $y$. Therefore $f(x, y)$ is measurable in $y$. This proves $S$ is closed under upward monotone convergence. The argument for downward monotone convergence is the same. Thus $S$ is a monotone class. Since $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ is the smallest monotone class including $L$, this establishes the result. $\square$

**Lemma 14.6** *Let $\mu_1$ be a $\sigma$-finite integral defined on $\mathcal{F}_1^+$. Also let $\mu_2$ be a $\sigma$-finite integral defined on $\mathcal{F}_2^+$. Let $f : X_1 \times X_2 \to [0, +\infty]$ be a $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ measurable function. Then the function $\mu_2 \circ f^{|1}$ is an $\mathcal{F}_1$ measurable function on $X_1$ with values in $[0, +\infty]$. Also the function $\mu_1 \circ f^{|2}$ is an $\mathcal{F}_2$ measurable function on $X_2$ with values in $[0, +\infty]$.*

Explicitly, this lemma says that the functions

$$x \mapsto \int f(x, y)\, d\mu_2(y) \tag{14.19}$$

and

$$y \mapsto \int f(x, y)\, d\mu_1(y) \tag{14.20}$$

are measurable functions.

Proof: The previous lemma shows that the integrals are well defined. Consider the class $S$ of functions $f$ for which the first assertion of the lemma holds. If $f$ is in $L^+$, then $f = \sum_i c_i 1_{A_i \times B_i}$, where each $A_i$ is an $\mathcal{F}_1$ set and each $B_i$ is a $\mathcal{F}_2$ set. Then for fixed $x$ consider the function $y \mapsto \sum_i c_i 1_{A_i}(x) 1_{B_i}(y)$. Its $\mu_2$ integral is $\sum_i c_i 1_{A_i}(x)\mu(B_i)$. This is clearly in $\mathcal{F}_1$ as a function of $x$. This shows that $L \subset S$. Now suppose that $f_n$ is a sequence of $L$-bounded functions, that $f_n \uparrow f$, and each $f_n$ is in $S$. Then we have that $\int f_n(x,y)\, d\mu_2(y)$ is measurable in $x$. Furthermore, for each $x$ it increases to $\int f(x,y)\, d\mu_2(y)$, by the monotone convergence theorem. Therefore $\int f(x,y)\, d\mu_2(y)$ is measurable in $x$. This proves $S$ is closed under upward monotone convergence of $L$-bounded functions. The argument for downward monotone convergence uses the improved monotone convergence theorem; here it is essential that each $f_n$ be an $L$-bounded function. Thus $S$ is an $L$-bounded monotone class including $L^+$. It follows that $(\mathcal{F}_1 \bigotimes \mathcal{F}_2)^+ \subset S$. □

**Lemma 14.7** *Let $f : X_1 \times X_2 \to [0, +\infty]$ be a $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ measurable function . Then $\nu_{12}(f) = \mu_2(\mu_1 \circ f^{|2})$ defines an integral $\nu_{12}$. Also $\nu_{21}(f) = \mu_1(\mu_2 \circ f^{|1})$ defines an integral $\nu_{21}$.*

Explicitly, this lemma says that the iterated integrals

$$\nu_{12}(f) = \int \left( \int f(x,y)\, d\mu_1(x) \right) d\mu_2(y) \tag{14.21}$$

and

$$\nu_{21}(f) = \int \left( \int f(x,y)\, d\mu_2(y) \right) d\mu_1(x) \tag{14.22}$$

are defined.

Proof: The previous lemma shows that the integral $\nu_{12}$ is well defined. It is easy to see that $\nu_{12}$ is linear and order preserving. The remaining task is to prove upward monotone convergence. Say that $f_n \uparrow f$ pointwise. Then by the monotone convergence theorem for $\mu_1$ we have that for each $y$ the integral $\int f_n(x,y)\, d\mu_1(x) \uparrow \int f(x,y)\, d\mu_1(x)$. Hence by the monotone convergence theorem for $\mu_2$ we have that $\int \int f_n(x,y)\, d\mu_1(x)\, d\mu_2(y) \uparrow \int f(x,y)\, d\mu_1(x)\, d\mu_2(y)$. This is the same as saying that $\nu_{12}(f_n) \uparrow \mu_{12}(f)$. □

**Theorem 14.8 (Tonelli's theorem)** . *Let $\mathcal{F}_1$ be a $\sigma$-algebra of real functions on $X_1$, and let $\mathcal{F}_2$ be a $\sigma$-algebra of real functions on $X_2$. Let $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ be the product $\sigma$-algebra of real functions on $X_1 \times X_2$. Let $\mu_1 : \mathcal{F}_1^+ \to [0, +\infty]$ and $\mu_2 : \mathcal{F}_2^+ \to [0, +\infty]$ be $\sigma$-finite integrals. Then there is a unique $\sigma$-finite integral*

$$\mu_1 \times \mu_2 : (\mathcal{F}_1 \bigotimes \mathcal{F}_2)^+ \to [0, +\infty] \tag{14.23}$$

*such that $(\mu_1 \times \mu_2)(g \otimes h) = \mu_1(g)\mu_2(h)$ for each $g$ in $\mathcal{F}_1^+$ and $h$ in $\mathcal{F}_2^+$. Furthermore, for $f$ in $(\mathcal{F}_1 \bigotimes \mathcal{F}_2)^+$ we have*

$$(\mu_1 \times \mu_2)(f) = \mu_2(\mu_1 \circ f^{|2}) = \mu_1(\mu_2 \circ f^{|1}). \tag{14.24}$$

In this statement of the theorem $f^{|2}$ is regarded as a function on $X_2$ with values that are functions on $X_1$. Similarly, $f^{|1}$ is regarded as a function on $X_1$ with values that are functions on $X_2$. Thus the composition $\mu_1 \circ f^{|2}$ is a function on $X_2$, and the composition $\mu_2 \circ f^{|1}$ is a function on $X_1$.

The theorem may be also be stated in a version with bound variables:

$$\int f(x,y)\, d(\mu_1 \times \mu_2)(x,y) = \int \left[ \int f(x,y)\, d\mu_1(x) \right] d\mu_2(y) = \int \left[ \int f(x,y)\, d\mu_2(y) \right] d\mu_1(x). \tag{14.25}$$

Proof: The integrals $\nu_{12}$ and $\nu_{21}$ agree on $L^+$. Consider the set $S$ of $f \in (\mathcal{F}_1 \bigotimes \mathcal{F}_2)^+$ such that $\nu_{12}(f) = \nu_{21}(f)$. The argument of the previous lemma shows that this is an $L$-monotone class. Hence $S$ is all of $(\mathcal{F}_1 \bigotimes \mathcal{F}_2)^+$. Define $\nu(f)$ to be the common value $\nu_{12}(f) = \nu_{21}(f)$. Then $\nu$ is uniquely defined by its values on $L^+$. This $\nu$ is the desired product measure $\mu_1 \times \mu_2$. □

The integral $\nu$ is called the product integral and is denoted by $\mu_1 \times \mu_2$. Let $F^2 : \mathbb{R}^{X_1 \times X_2} \to (\mathbb{R}^{X_1})^{X_2}$ be given by $f \mapsto f^{|2}$, that is, $F_2$ says to hold the second second variable constant. Similarly, let $F^1 : \mathbb{R}^{X_1 \times X_2} \to (\mathbb{R}^{X_2})^{X_1}$ be given by $f \mapsto f^{|1}$, that is, $F^1$ says to hold the first variable constant. Then the Tonelli theorem says that the product integral $\mu_1 \times \mu_2 : (\mathcal{F}_1 \times \mathcal{F}_2)^+ \to [0, +\infty]$ satisfies

$$\mu_1 \times \mu_2 = \mu_2 \circ \mu_1 \circ F^2 = \mu_1 \circ \mu_2 \circ F^1. \tag{14.26}$$

## 14.5 Fubini's theorem

Recall that for an arbitrary non-empty set $X$, $\sigma$-algebra of functions $\mathcal{F}$, and integral $\mu$, the space $\mathcal{L}^1(X, \mathcal{F}, \mu)$ consists of all real functions $f$ in $\mathcal{F}$ such that $\mu(|f|) < +\infty$. For such a function $\mu(|f|) = \mu(f_+) + \mu(f_-)$, and $\mu(f) = \mu(f_+) - \mu(f_-)$ is a well-defined real number.

Let $f$ be in $\mathcal{L}^1(X \times Y, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \times \mu_2)$. Let $\Lambda_1$ be the set of all $x$ with $f^{|1}(x)$ in $\mathcal{L}^1(X_2, \mathcal{F}_2, \mu_2)$ and let $\Lambda_2$ be the set of all $y$ with $f^{|2}(y)$ in $\mathcal{L}^1(X_1, \mathcal{F}_1, \mu_1)$. Then $\mu_1(\Lambda_1^c) = 0$ and $\mu_2(\Lambda_2^c) = 0$. Define the *partial integral* $\mu_2(f \mid 1)$ by $\mu_2(f \mid 1)(x) = \mu_2(f^{|1}(x))$ for $x \in \Lambda_1$ and $\mu_2(f \mid 1)(x) = 0$ for $x \in \Lambda_1^c$. Define the partial integral $\mu_1(f \mid 2)$ by $\mu_1(f \mid 2)(y) = \mu_1(f^{|2}(y))$ for $y \in \Lambda_2$ and $\mu_1(f \mid 2)(y) = 0$ for $y \in \Lambda_2^c$.

**Theorem 14.9 (Fubini's theorem)** *Let $\mathcal{F}_1$ be a $\sigma$-algebra of real functions on $X_1$, and let $\mathcal{F}_2$ be a $\sigma$-algebra of real functions on $X_2$. Let $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ be the product $\sigma$-algebra of real functions on $X_1 \times X_2$. Let $\mu_1$ and $\mu_2$ be $\sigma$-finite integrals, and consider the corresponding functions*

$$\mu_1 : \mathcal{L}^1(X, \mathcal{F}_1, \mu_1) \to \mathbb{R} \tag{14.27}$$

*and*

$$\mu_2 : \mathcal{L}^1(X_2, \mathcal{F}_2, \mu_2) \to \mathbb{R}. \tag{14.28}$$

*The product integral $\mu_1 \times \mu_2$ defines a function*

$$\mu_1 \times \mu_2 : \mathcal{L}^1(X_1 \times X_2, \mathcal{F}_1 \bigotimes F_2, \mu_1 \times \mu_2) \to \mathbb{R}. \tag{14.29}$$

*Let $f$ be in $\mathcal{L}^1(X \times Y, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu_1 \times \mu_2)$. Then the partial integral $\mu_2(f \mid 1)$ is in $\mathcal{L}^1(X_1, \mathcal{F}_1, \mu_1)$, and the partial integral $\mu_1(f \mid 2)$ is in $\mathcal{L}^1(X_2, \mathcal{F}_2, \mu_2)$. Finally,*

$$(\mu_1 \times \mu_2)(f) = \mu_1((\mu_2(f \mid 1)) = \mu_2(\mu_1(f \mid 2)). \tag{14.30}$$

In this statement of the theorem $\mu_2(f \mid 1)$ is the $\mu_2$ partial integral with the first variable fixed, regarded after integration as a function on $X_1$. Similarly, $\mu_1(f \mid 2)$ is the $\mu_1$ partial integral with the second variable fixed, regarded after integration as a function on $X_2$.

Fubini's theorem may also be stated with bound variables:

$$\int f(x,y)\, d(\mu_1 \times \mu_2)(x,y) = \int_{\Lambda_1} \left[\int f(x,y)\, d\mu_2(x)\right] d\mu_1(x) = \int_{\Lambda_2} \left[\int f(x,y)\, d\mu_1(x)\right] d\mu_2(y). \tag{14.31}$$

Here as before $\Lambda_1$ and $\Lambda_2$ are sets where the inner integral converges absolutely. The complement of each of these sets has measure zero.

Proof: By Tonelli's theorem we have that $\mu_2 \circ |f^{|1}|$ is in $\mathcal{L}^1(X_1, \mathcal{F}_1, \mu_1)$ and that $\mu_1 \circ |f^{|2}|$ is in $\mathcal{L}^2(X_2, \mathcal{F}_2, \mu_2)$. This is enough to show that $\mu_2(\Lambda_1^c) = 0$ and $\mu_1(\Lambda_2^c) = 0$. Similarly, by Tonelli's theorem we have

$$(\mu_1 \times \mu_2)(f) = (\mu_1 \times \mu_2)(f_+) - (\mu_1 \times \mu_2)(f_-) = \mu_1(\mu_2 \circ f_+^{|1}) - \mu_1(\mu_2 \circ f_-^{|1}). \tag{14.32}$$

Since $\Lambda_1$ and $\Lambda_2$ are sets whose complements have measure zero, we can also write this as

$$(\mu_1 \times \mu_2)(f) = \mu_1(1_{\Lambda_1}(\mu_2 \circ f_+^{|1})) - \mu_1(1_{\Lambda_1}(\mu_2 \circ f_-^{|1})). \tag{14.33}$$

Now for each fixed $x$ in $\Lambda_1$ we have

$$\mu_2(f^{|1}(x)) = \mu_2(f_+^{|1}(x)) - \mu_2(f_-^{|1}(x)). \tag{14.34}$$

This says that

$$\mu_2(f \mid 1) = 1_{\Lambda_1}(\mu_2 \circ f_+^{|1}) - 1_{\Lambda_1}(\mu_2 \circ f_-^{|1}). \tag{14.35}$$

Each function on the right hand side is a real function in $\mathcal{L}^1(X_1, \mathcal{F}_1, \mu_1)$. So

$$(\mu_1 \times \mu_2)(f) = \mu_1(\mu_2(f \mid 1)). \tag{14.36}$$

□

Tonelli's theorem and Fubini's theorem are often used together to justify an interchange of order of integration. Here is a typical pattern. Say that one can show that the iterated integral with the absolute value converges:

$$\int \left[\int |h(x,y)|\, d\nu(y)\right] d\mu(x) < \infty. \tag{14.37}$$

By Tonelli's theorem the product integral also converges:

$$\int |h(x,y)|\, d(\mu \times \nu)(x,y) < \infty. \tag{14.38}$$

Then from Fubini's theorem the integrated integrals are equal:

$$\int \left[ \int h(x, y)\, d\nu(y) \right] d\mu(x) = \int \left[ \int h(x, y)\, d\mu(x) \right] d\nu(y). \tag{14.39}$$

The outer integrals are each taken over a set for which the inner integral converges absolutely; the complement of this set has measure zero.

## 14.6 Supplement: Semirings and rings of sets

This section supplies the proof that finite linear combinations of indicator functions of rectangles form a vector lattice. It may be omitted on a first reading. The first and last results in this section are combinatorial lemmas that are proved in books on measure theory. See Chapter 3 of the text by Dudley [4].

Let $X$ be a set. A *ring* $\mathcal{R}$ of subsets of $X$ is a collection such that $\emptyset$ is in $\mathcal{R}$ and such that $A$ and $B$ in $\mathcal{R}$ imply $A \cap B$ is in $\mathcal{R}$ and such that $A$ and $B$ in $\mathcal{R}$ imply that $A \setminus B$ is in $\mathcal{R}$.

A *semiring* $\mathcal{D}$ of subsets of $X$ is a collection such that $\emptyset$ is in $\mathcal{D}$ and such that $A$ and $B$ in $\mathcal{D}$ imply $A \cap B$ is in $\mathcal{D}$ and such that $A$ and $B$ in $\mathcal{D}$ imply that $A \setminus B$ is a finite union of disjoint members of $\mathcal{D}$.

**Proposition 14.10** *Let $\mathcal{D}$ be a semiring of subsets of $X$. Let $\mathcal{R}$ be the ring generated by $\mathcal{D}$. Then $\mathcal{R}$ consists of all finite unions of members of $\mathcal{D}$.*

**Proposition 14.11** *Let $\mathcal{D}$ be a semiring of subsets of a set $X$. Let $\Gamma$ be a finite collection of subsets in $\mathcal{D}$. Then there exists a finite collection $\Delta$ of disjoint subsets in $\mathcal{D}$ such that each set in $\Gamma$ is a finite union of some subcollection of $\Delta$.*

Proof: For each non-empty subcollection $\Gamma'$ of $\Gamma$ consider the set $A_{\Gamma'}$ that is the intersection of the sets in $\Gamma'$ with the intersection of the complements of the sets in $\Gamma \setminus \Gamma'$. The sets $A_{\Gamma'}$ are in $\mathcal{R}$ and are disjoint. Furthermore, each set $C$ in $\Gamma$ is the finite disjoint union of the sets $A_{\Gamma'}$ such that $C \in \Gamma'$. The proof is completed by noting that by the previous proposition each of these sets $A_{\Gamma'}$ is itself a finite disjoint union of sets in $\mathcal{D}$. $\Box$

**Theorem 14.12** *Let $\mathcal{D}$ be a semiring of subsets of $X$. Let $L$ be the set of all finite linear combinations of indicator functions of sets in $\mathcal{D}$. Then $L$ is a vector lattice.*

Proof: The problem is to prove that $L$ is closed under the lattice operations. Let $f$ and $g$ be in $L$. Then $f$ is a finite linear combination of indicator functions of sets in $\mathcal{D}$. Similarly, $g$ is a finite linear combination of indicator functions of sets in $\mathcal{D}$. Take the union $\Gamma$ of these two collections of sets. These sets may not be disjoint, but there is a collection $\Delta$ of disjoint sets in $\mathcal{D}$ such that each set in the union is a disjoint union of sets in $\Delta$. Then $f$ and $g$ are each linear combinations of indicator functions of disjoint sets in $\Delta$. It follows that $f \wedge g$ and $f \vee g$ also have such a representation. $\Box$

**Theorem 14.13** *Let $X_1$ and $X_2$ be non-empty sets, and let $\mathcal{D}_1$ and $\mathcal{D}_2$ be semirings of subsets. Then the set of all $A \times B$ with $A \in \mathcal{D}_1$ and $B \in \mathcal{D}_2$ is a semiring of subsets of $X_1 \times X_2$.*

In the application to product measures the sets $\mathcal{D}_1$ and $\mathcal{D}_2$ consist of sets of finite measure. Thus each of $\mathcal{D}_1$ and $\mathcal{D}_2$ is a ring of subsets. It follows from the last theorem that the product sets form a semiring of subsets of the product space. The previous theorem then shows that the finite linear combinations form a vector lattice.

# Problems

1. Let $g$ be a real Borel function on the line that is in $\mathcal{L}^1$. Thus

$$\|g\|_1 = \int_{-\infty}^{\infty} |g(x)|\, dx < +\infty. \tag{14.40}$$

   Let $f$ be another such function. Show that there is a subset $\Lambda$ of the real line such that for each $x$ in $\Lambda$ the function $y \mapsto g(x-y)f(y)$ is in $\mathcal{L}^1$, and such that the complement of $\Lambda$ has measure zero.

   Define

$$h(x) = \int_{-\infty}^{\infty} g(x-y)f(y)\, dy \tag{14.41}$$

   for $x$ in $\Lambda$, and define $h(x) = 0$ for $x$ in the complement of $\Lambda$. Prove that $h$ is in $\mathcal{L}^1$ and that $\|h\|_1 \leq \|g\|_1 \|f\|_1$.

2. In the previous problem, show by example that it is possible that $\Lambda$ is a proper subset of the real line.

3. Let $\mu$ and $\nu$ be $\sigma$-finite measures, and let $k$ be a measurable function on the product space. Suppose that

$$\sup_y \int_{-\infty}^{\infty} |k(x,y)|\, d\mu(x) = M < +\infty. \tag{14.42}$$

   Let $f$ be absolutely integrable with respect to $\nu$. (Thus $f$ is measurable and its absolute value has finite $\nu$ integral.) For each $f$ define

$$h(x) = \int k(x,y) f(y)\, d\nu(y). \tag{14.43}$$

   Show using the same reasoning as in the first problem that

$$\|h\|_1 \leq M \|f\|_1. \tag{14.44}$$

4. Let $\mu$ and $\nu$ be $\sigma$-finite measures, and let $k \geq 0$ be a measurable function on the product space. Suppose that for each $y$

$$\int k(x, y)\, d\mu(x) = 1. \tag{14.45}$$

Let $f \geq 0$ be absolutely integrable, and define

$$h(x) = \int k(x, y) f(y)\, d\nu(y). \tag{14.46}$$

Show that

$$\|h\|_1 = \|f\|_1. \tag{14.47}$$

The interpretation is that $f$ is an initial probability density, and $k(x, y)$ is the transition probability density from $y$ to $x$. Then $h$ is the final probability density.

5. Show that if $\mu$ is not required to be $\sigma$-finite, then it is possible to have $\|f\|_1 = 1$ and $\|h\|_1 = 0$. Hint: Take the transition to go from each point to the same point.

# Chapter 15

# Probability

## 15.1 Coin-tossing

A basic probability model is that for coin-tossing. The set of outcomes of the experiment is $\Omega = 2^{\mathbf{N}_+}$. Let $b_j$ be the $j$th coordinate function. Let $f_{nk}$ be the indicator function of the set of outcomes that have the $k$ pattern in the first $n$ coordinates. Here $0 \leq k < 2^n$, and the pattern is given by the binary representation of $k$. If $S$ is the subset of $\{1, \ldots, n\}$ where the 1s occur, and $S^c$ is the subset where the 0s occur, then

$$f_{nk} = \prod_{j \in S} b_j \prod_{j \in S^c} (1 - b_j). \tag{15.1}$$

The expectation $\mu$ is determined by

$$\mu(f_{nk}) = p^j q^{n-j}, \tag{15.2}$$

where $j$ is the number of 1s in the binary expansion of $k$, or the number of points in $S$. It follows that if $S$ and $T$ are disjoint subsets of $\{1, \ldots, n\}$, then

$$\mu(\prod_{j \in S} b_j \prod_{j \in T} (1 - b_j)) = p^j q^\ell, \tag{15.3}$$

where $j$ is the number of elements in $S$, and $\ell$ is the number of elements in $T$.

It follows from these formulas that the probability of success on one trial is $\mu(b_j) = p$ and the probability of failure on one trial is $\mu(1-b_j) = q$. Similarly, for two trials $i < j$ the probabilities of two successes is $\mu(b_i b_j) = p^2$, the probability of success followed by failure is $\mu(b_i)(1 - b_j)) = pq$, the probability of failure followed by success is $\mu((1 - b_i)b_j = qp$, and the probability of two failures is $\mu((1 - b_i)(1 - b_j)) = q^2$.

## 15.2 Weak law of large numbers

**Theorem 15.1 (Weak law of large numbers)** *Let*

$$s_n = b_1 + \cdots + b_n \tag{15.4}$$

*be the number of successes in the first* $n$ *trials. Then*

$$\mu(s_n) = np \tag{15.5}$$

*and*

$$\mu((s_n - np)^2) = npq. \tag{15.6}$$

Proof: Expand $(s_n - np)^2 = \sum_{i=1}^n \sum_{j=1}^n (b_i - p)(b_j - p)$. The expectation of each of the cross terms vanishes. The expectation of each of the diagonal terms is $(1-p)^2 p + (0-p)^2 q = q^2 p + p^2 q = pq$. □

**Corollary 15.2 (Weak law of large numbers)** *Let*

$$f_n = \frac{b_1 + \cdots + b_n}{n} \tag{15.7}$$

*be the proportion of successes in the first* $n$ *trials. Then*

$$\mu(f_n) = p \tag{15.8}$$

*and*

$$\mu((f_n - p)^2) = \frac{pq}{n} \leq \frac{1}{4n}. \tag{15.9}$$

The quantity that is usually used to evaluate the error is the standard deviation, which is the square root of this quantity. The version that should be memorized is thus

$$\sqrt{\mu((f_n - p)^2)} = \frac{\sqrt{pq}}{\sqrt{n}} \leq \frac{1}{2\sqrt{n}}. \tag{15.10}$$

This $1/\sqrt{n}$ factor is what makes probability theory work (in the sense that it is internally self-consistent).

**Corollary 15.3** *Let*

$$f_n = \frac{b_1 + \cdots + b_n}{n} \tag{15.11}$$

*be the proportion of successes in the first* $n$ *trials. Then*

$$\mu(|f_n - p| \geq \epsilon) = \frac{pq}{n\epsilon^2} \leq \frac{1}{4n\epsilon^2}. \tag{15.12}$$

This corollary follows immediately from Chebyshev's inequality. It gives a perhaps more intuitive picture of the meaning of the weak law of large numbers. Consider a tiny $\epsilon > 0$. Then it says that if $n$ is sufficiently large, then, with probability very close to one, the experimental proportion $f_n$ differs from $p$ by less than $\epsilon$.

## 15.3 Strong law of large numbers

**Theorem 15.4** *Let*

$$s_n = b_1 + \cdots + b_n \tag{15.13}$$

*be the number of successes in the first $n$ trials. Then*

$$\mu(s_n) = np \tag{15.14}$$

*and*

$$\mu((s_n - np)^4) = n(pq^4 + qp^4) + 3n(n-1)(pq)^2. \tag{15.15}$$

*This is bounded by $(1/4)n^2$ for $n \geq 4$.*

Proof: Expand $(s_n - np)^4 = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n (b_i - p)(b_j - p)(b_k - p)(b_l - p)$. The expectation of each of the terms vanishes unless all four indices coincide or there are two pairs of coinciding indices. The expectation for the case when all four indices coincide is $(1-p)^4 p + (0-p)^4 q = q^4 p + p^4 q = pq(q^3 + p^3)$. There are $n$ such terms. The expectation when there are two pairs of coinciding indices works out to be $(pq)^2$. There are $3n(n-1)$ such terms.

The inequality then follows from $npq(q^3+p^3)+3n^2(pq)^2 \leq n/4+3/(16)n^2 \leq (1/4)n^2$ for $n \geq 4$. □

**Corollary 15.5** *Let*

$$f_n = \frac{b_1 + \cdots + b_n}{n} \tag{15.16}$$

*be the proportion of successes in the first $n$ trials. Then*

$$\mu(f_n) = p \tag{15.17}$$

*and*

$$\mu((f_n - p)^4) \leq \frac{1}{4n^2} \tag{15.18}$$

*for $n \geq 4$.*

**Corollary 15.6 (Strong law of large numbers)** *Let*

$$f_n = \frac{b_1 + \cdots + b_n}{n} \tag{15.19}$$

*be the proportion of successes in the first $n$ trials. Then*

$$\mu(\sum_{n=k}^{\infty} (f_n - p)^4) \leq \frac{1}{4(k-1)} \tag{15.20}$$

*for $k \geq 4$.*

This corollary has a remarkable consequence. Fix $k$. The fact that the expectation is finite implies that the sum converges almost everywhere. In particular, the terms of the sum approach zero almost everywhere. This means that $f_n \to p$ as $n \to \infty$ almost everywhere. This is the traditional formulation of the strong law of large numbers.

**Corollary 15.7 (Strong law of large numbers)** *Let*

$$f_n = \frac{b_1 + \cdots + b_n}{n} \tag{15.21}$$

*be the proportion of successes in the first $n$ trials. Then for $k \geq 4$*

$$\mu(\sup_{n\geq k} |f_n - p| \geq \epsilon) \leq \frac{1}{4(k-1)\epsilon^4}. \tag{15.22}$$

Proof: This corollary follows from the trivial fact that $\sup_{n\geq k} |f_n - p|^4 \leq \sum_{n=k}^{\infty}(f_n - p)^4$ and Chebyshev's inequality. □

This corollary give a perhaps more intuitive picture of the meaning of the strong law of large numbers. Consider a tiny $\epsilon > 0$. Then it says that if $k$ is sufficiently large, then, with probability very close to one, for the entire future history of $n \geq k$ the experimental proportions $f_n$ differ from $p$ by less than $\epsilon$.

## 15.4 Random walk

Let $w_j = 1 - 2b_j$, so that $b_j = 0$ gives $w_j = 1$ and $b_j = 1$ gives $w_j = -1$. Then the sequence $x_n = w_1 + \cdots w_n$ is called *random walk* starting at zero. In the case when $p = q = 1/2$ this is called symmetric random walk.

**Theorem 15.8** *Let $\rho_{01}$ be the probability that the random walk starting at zero ever reaches 1. Then this is a solution of the equation*

$$q\rho^2 - \rho + p = (q\rho - p)(\rho - 1) = 0. \tag{15.23}$$

*In particular, if $p = q = 1/2$, then $\rho_{01} = 1$.*

Proof: Let $\rho = \rho_{01}$. The idea of the proof is to break up the computation of $\rho$ into the case when the first step is positive and the case when the first step is negative. Then the equation

$$\rho = p + q\rho^2 \tag{15.24}$$

is intuitive. The probability of succeeding at once is $p$. Otherwise there must be a failure followed by getting from $-1$ to 0 and then from 0 to 1. However getting from $-1$ to 0 is of the same difficulty as getting from 0 to 1.

To make this intuition precise, lett $\tau_1$ be the first time that the walk reaches one. Then

$$\rho = \mu(\tau_1 < +\infty) = \mu(w_1 = 1, \tau_1 < +\infty) + \mu(w_1 = -1, \tau_1 < +\infty). \tag{15.25}$$

The value of the first term is $p$.

The real problem is with the second term. Write it as

$$\mu(w_1 = -1, \tau_1 < +\infty) = \sum_{k=2}^{\infty} \mu(w_1 = -1, \tau_0 = k, \tau_1 < +\infty) = \sum_{k=2}^{\infty} q\mu(\tau_1 = k-1)\rho = q\rho^2. \tag{15.26}$$

This gives the conclusion. It may be shown that when $p < q$ the correct solution is $\rho = p/q$. □

Notice the dramatic fact that when $p = q = 1/2$ the probability that the random walk gets to the next higher point is one. It is not hard to extend this to show that the probability that the random walk gets to any other point is also one. So the symmetric random walk must do a lot of wandering.

**Theorem 15.9** *Let $m_{01}$ be the expected time until the random walk starting at zero reaches 1. Then $m_{01}$ is a solution of*

$$m = 1 + 2qm. \tag{15.27}$$

*In particular, when $p = q = 1/2$ the solution is $m = +\infty$.*

Proof: Let $m = m_{01}$. The idea of the proof is to break up the computation of $\rho$ into the case when the first step is positive and the case when the first step is negative. Then the equation

$$m = p + q(1 + 2m). \tag{15.28}$$

is intuitive. The probability of succeeding at once is $p$, and this takes time 1. Otherwise $\tau_1 = 1 + (\tau_0 - 1) + (\tau_1 - \tau_0)$. However the average of the time $\tau_0 - 1$ to get from $-1$ to 0 is the same as the average of the time $\tau_1 - \tau_0$ to get from 0 to 1.

A more detailed proof is to write

$$m = \mu(\tau_1) = \mu(\tau_1 1_{w_1=1}) + \mu(\tau_1 1_{w_1=-1}). \tag{15.29}$$

The value of first term is $p$.

The second term is

$$\mu(\tau_1 1_{w_1=-1}) = \mu((1+(\tau_0-1)+(\tau_1-\tau_0))1_{w_1=-1}) = q+q\mu(\tau_1)+q\mu(\tau_1) = q(1+2m). \tag{15.30}$$

It may be shown that when $p > q$ the correct solution is $m = 1/(p - q)$. □

When $p = q = 1/2$ the expected time for the random walk to get to the next higher point is infinite. This is because there is some chance that the symmetric random walk wanders for a very long time on the negative axis before getting to the points above zero.

# Problems

1. Consider a random sample of size $n$ from a very large population. The experimental question is to find what proportion $p$ of people in the population have a certain opinion. The proportion in the random sample who have the opinion is $f_n$. How large must $n$ be so that the standard deviation of $f_n$ in this type of experiment is guaranteed to be no larger than one percent?

2. Recall that $f_n(x) \to f(x)$ as $n \to \infty$ means $\forall \epsilon > 0\, \exists N\, \forall n \geq N\, |f_n(x) - f(x)| < \epsilon$. Show that $f_n \to f$ almost everywhere is equivalent to

$$\mu(\{x \mid \exists \epsilon > 0\, \forall N\, \exists n \geq N\, |f_n(x) - f(x)| \geq \epsilon\}) = 0. \tag{15.31}$$

3. Show that $f_n \to f$ almost everywhere is equivalent to for all $\epsilon > 0$

$$\mu(\{x \mid \forall N\, \exists n \geq N\, |f_n(x) - f(x)| \geq \epsilon\}) = 0. \tag{15.32}$$

4. Suppose that the measure of the space is finite. Show that $f_n \to f$ almost everywhere is equivalent to for all $\epsilon > 0$

$$\lim_{N\to\infty} \mu(\{x \mid \exists n \geq N\, |f_n(x) - f(x)| \geq \epsilon\}) = 0. \tag{15.33}$$

Show that this is not equivalent in the case when the measure of the space may be infinite. Note: Convergence almost everywhere occurs in the strong law of large numbers.

5. Say that $f_n \to f$ *in measure* if for all $\epsilon > 0$

$$\lim_{N\to\infty} \mu(\{x \mid |f_N(x) - f(x)| \geq \epsilon\}) = 0. \tag{15.34}$$

Show that if the measure of the space is finite, then $f_n \to f$ almost everywhere implies $f_n \to f$ in measure. Note: Convergence in measure occurs in the weak law of large numbers.

# Part IV

# Metric Spaces

# Chapter 16

# Metric spaces

## 16.1 Metric space notions

A *metric space* $M, d$ is a set $M$ together with a *distance*function $d : M \times M \to [0, +\infty]$ such that for all $x, y, z$

1. $d(x, x) = 0$.
2. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality)
3. $d(x, y) = d(y, x)$ (symmetry).
4. $d(x, y) = 0$ implies $x = y$ (separatedness).
5. $d(x, y) < +\infty$ (finiteness).

The crucial properties are the first two, particularly the *triangle inequality.* There are generalizations where some of these properties are allowed to fail[13]. If the finiteness condition is allowed to fail, then we have an *extended metric.* If the separateness condition is allowed to fail, then we have a *pseudo-metric.* If the symmetry condition is allowed to fail, then we have a *quasi-metric.* The most general situation, where only properties 1 and 2 hold, is that of an extended pseudo-quasimetric, or what I shall call a *Lawvere metric.* Such structures are natural and important, as we shall see at the end of the chapter.

When the metric is understood from context, it is common to refer to a metric space $M, d$ by the underlying set $M$. Every subset of a metric space is itself a metric space with the *relative metric* obtained by restriction. In that case, the subset with its metric is called a *subspace.*

**Proposition 16.1** *Let $M$ be a metric space. Then For all $x, y, z$ in $M$ we have $|d(x, z) - d(y, z)| \leq d(x, y)$.*

Proof: From the triangle inequality $d(x, z) \leq d(x, y) + d(y, z)$ we obtain $d(x, z) - d(y, z) \leq d(x, y)$. On the other hand, from the triangle inequality we also have $d(y,z) \leq d(y,x) + d(x,z)$ which implies $d(y,z) - d(x,z) \leq d(y,x) = d(x,y)$. $\square$

In a metric space $M$ the *open ball* centered at $x$ of radius $\epsilon > 0$ is defined to be $B(x,\epsilon) = \{y \mid d(x,y) < \epsilon\}$. The *closed ball* centered at $x$ of radius $\epsilon > 0$ is defined to be $\overline{B(x,\epsilon)} = \{y \mid d(x,y) \leq \epsilon\}$. The *sphere* centered at $x$ of radius $\epsilon > 0$ is defined to be $S(x,\epsilon) = \{y \mid d(x,y) = \epsilon\}$.

Sometimes one wants to speak of the distance of a point from a non-empty set. This is defined to be $d(x,A) = \inf_{y \in A} d(x,y)$.

## 16.2 Normed vector spaces

One common way to get a metric is to have a norm on a vector space. A *norm* on a real vector space $V$ is a function from $V$ to $[0,+\infty)$ with the following three properties:

1. For all $x$ we have $\|x\| = 0$ if and only if $x = 0$.
2. For all $x,y$ we have $\|x+y\| \leq \|x\| + \|y\|$ (triangle inequality).
3. For all $x$ and real $t$ we have $\|tx\| = |t|\|x\|$.

The corresponding metric is then $d(x,y) = \|x-y\|$. Again the crucial property in the definition is the *triangle inequality*. The classic example, of course, is Euclidean space $\mathbb{R}^n$ with the usual square root of sum of squares norm. In the following we shall see that this $\ell^2_n$ norm is just one possibility among many.

## 16.3 Spaces of finite sequences

Here are some possible metrics on $\mathbb{R}^n$. The most geometrical metric is the $\ell^2_n$ metric given by the $\ell^2_n$ norm. This is $d_2(x,y) = \|x-y\|_2 = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$. It is the metric with the nicest geometric properties. A sphere in this metric is a nice round sphere.

Sometimes in subjects like probability one wants to look at the sum of absolute values instead of the sum of squares. The $\ell^1_n$ metric is $d_1(x,y) = \|x-y\|_1 = \sum_{k=1}^n |x_k - y_k|$. A sphere in this metric is actually a box with corners on the coordinate axes.

In other areas of mathematics it is common to look at the biggest or worst case. The $\ell^\infty_n$ metric is $d_\infty(x,y) = \|x-y\|_\infty = \max_{1 \leq k \leq n} |x_k - y_k|$. A sphere in this metric is a box with the flat sides on the coordinate axes.

Comparisons between these metrics are provided by

$$d_\infty(x,y) \leq d_2(x,y) \leq d_1(x,y) \leq n d_\infty(x,y). \tag{16.1}$$

The only one of these comparisons that is not immediate is $d_2(x,y) \leq d_1(x,y)$. But this follows from $d_2(x,y) \leq \sqrt{d_1(x,y) d_\infty(x,y)} \leq d_1(x,y)$.

## 16.4 Spaces of infinite sequences

A *sequence* is often taken to be a function defined on $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, but it is sometimes also convenient to regard a sequence as defined on $\mathbb{N}_+ = \{1, 2, 3, \ldots\}$. Often all that matters is that there is an ordered set $\mathbf{N}$ that is order isomorphic to either of these. In fact, in some cases even the order is not important. For instance, there are cases when the natural index set is $\mathbb{Z}$.

The $\ell^2$ metric is defined on the set of all infinite sequences such that $\|x\|_2^2 = \sum_{k=1}^\infty |x_k|^2 < \infty$. The metric is $d_2(x, y) = \|x - y\|_2 = \sqrt{\sum_{k=1}^\infty (x_k - y_k)^2}$. This is again a case with wonderful geometric properties. It is a vector space with a norm called real Hilbert space. The fact that the norm satisfies the triangle inequality is the subject of the following digression.

**Lemma 16.2 (Schwarz inequality)** *Suppose the inner product of two real sequences is to be defined by*

$$\langle x, y\rangle = \sum_{k=1}^\infty x_k y_k. \tag{16.2}$$

*If the two sequences $x, y$ are in $\ell^2$, then this inner product is absolutely convergent and hence well-defined, and it satisfies*

$$|\langle x, y\rangle| \leq \|x\|_2 \|y\|_2. \tag{16.3}$$

This well-known *Schwarz inequality* says that if we define the cosine of the angle between two non-zero vectors by $\langle x, y\rangle = \|x\|_2\|y\|_2 \cos(\theta)$, then $-1 \leq \cos(\theta) \leq 1$, and so the cosine has a reasonable geometrical interpretation. If we require the angle to satisfy $0 \leq \theta \leq \pi$, then the angle is also well-defined and makes geometrical sense.

The Schwarz inequality is just what is needed to prove the triangle inequality. The calculation is

$$\|x+y\|_2^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2. \tag{16.4}$$

The $\ell^1$ metric is defined on the set of all infinite sequences $x$ with $\|x\|_1 = \sum_{k=1}^\infty |x_k| < \infty$. The metric is $d_1(x, y) = \sum_{k=1}^\infty |x_k - y_k|$. This is the natural distance for absolutely convergence sequences. It is again a vector space with a norm. In this case it is not hard to prove the triangle inequality for the norm using elementary inequalities.

The $\ell^\infty$ metric is defined on the set of all bounded sequences. The metric is $d_\infty(x, y) = \sup_{1 \leq k < \infty} |x_k - y_k|$. That is, $d_\infty(x, y)$ is the least upper bound (supremum) of the $|x_k - y_k|$ for $1 \leq k < \infty$. This is yet one more vector space with a norm. The fact that this is a norm requires a little thought. The point is that for each $k$ we have $|x_k + y_k| \leq |x_k| + |y_k| \leq \|x\|_\infty + \|y\|_\infty$, so that $\|x\|_\infty + \|y\|_\infty$ is an upper bound for the set of numbers $|x_k + y_k|$. Since $\|x+y\|_\infty$ is the least upper bound for these numbers, we have $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$.

Comparisons between two of these these metrics are provided by

$$d_\infty(x, y) \le d_2(x, y) \le d_1(x, y). \tag{16.5}$$

As sets $\ell^1 \subset \ell^2 \subset \ell^\infty$. This is consistent with the fact that every absolutely convergent sequence is bounded. The proof of these inequalities is almost the same as for the finite-dimensional case. However the $\ell^\infty$ metric is defined by a supremum rather than by a maximum. So to bound it, one finds an upper bound for the set of all $|x_k - y_k$ and then argues that $\|x - y\|_\infty$ is the least such upper bound.

Yet another possibility is to try to define a metric on the product space $\mathbb{R}^{\mathbf{N}}$ of all sequences of real numbers. We shall often refer to this space as $\mathbb{R}^\infty$. This is the biggest possible metric space of sequences. In order to do this, it is helpful to first define a somewhat unusual metric on $\mathbb{R}$ by $d_b(s, t) = |s-t|/(1+|s-t|)$. We shall see below that the space $\mathbb{R}$ with this new metric is uniformly equivalent to the space $\mathbb{R}$ with its usual metric $d(s, t) = |s - t|$. However $d_b$ has the advantage that it is a metric that is bounded by one.

The metric on $\mathbb{R}^\infty$ is $d_p(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} d_b(x_k, y_k)$. This is called the product metric. The comparison between these metrics is given by the inequality

$$d_p(x, y) \le d_\infty(x, y). \tag{16.6}$$

The $d_p$ metric is an example of a metric on a vector space that is not given by a norm.

Two important subspaces of $\mathbb{R}^\infty$ are the *Hilbert cube* $[0, 1]^\infty$ and the *Cantor space* $\{0, 1\}^\infty$. For these two it is sometimes more convenient to use the norm $d_p(x, y) = \sum_{k=1}^{\infty} |x_k - y_k|/2^k$. The function $g : \{0, 1\}^\infty \to \mathbb{R}$ defined by $g(x) = \sum_{k=1}^{\infty} 2x_k/3^k$ is a bijection from the Cantor space to the *Cantor middle third set*.

There are other kinds of product space that come up from time to time. Let $T$ be the usual circle of circumference $2\pi$. Then $T^2$ is the product of two circles, so it is an ordinary *torus*. The space $T^\infty$ is thus a kind of generalized torus.

## 16.5 Spaces of bounded continuous functions

Here is another example. Let $X$ be a set. Let $B(X)$ be the set of all bounded functions on $X$. If $f$ and $g$ are two such functions, then $|f - g|$ is bounded, and so we can define the uniform metric by $d(f, g) = \|f - g\|_{\sup}$, where

$$\|f\|_{\sup} = \sup_s |f(s)| \tag{16.7}$$

is the *supremum norm*. This again is a normed vector space.

Suppose that $X$ is a metric space. A *continuous real function* is a map from $X$ to $\mathbb{R}$ that is continuous. Let $C(X)$ be the set of real continuous functions on $X$. Let $BC(X)$ be the set of bounded continuous real functions on $X$. This is the appropriate metric space for formulating the concept of *uniform* *convergence* of a sequence of continuous functions to a continuous function. Thus, the uniform convergence of $f_n$ to $g$ as $n \to \infty$ is equivalent to the condition $\lim_{n\to\infty} d_{\sup}(f_n, g) = 0$.

It should be remarked that all these examples have complex versions, where the only difference is that sequences of real numbers are replaced by sequences of complex numbers. So there is a complex Hilbert space, a space of bounded continuous complex functions, and so on. The version that is intended should be specified. Our default will be the real version, except in the context of Hilbert space and Fourier analysis.

## 16.6 Open and closed sets

A subset $U$ of a metric space $M$ is *open* if $\forall x\, (x \in U \Rightarrow \exists \epsilon\, B(x, \epsilon) \subset U)$. The following results are well-known facts about open sets.

**Theorem 16.3** *Let $\Gamma$ be a set of open sets. Then $\bigcup \Gamma$ is open.*

Proof: Let $x$ be a point in $\bigcup \Gamma$. Then there exists some $S$ in $\Gamma$ such that $x \in S$. Since $S$ is open there exists $\epsilon > 0$ with $B(x, \epsilon) \subset S$. However $S \subset \bigcup \Gamma$. So $B(x, \epsilon) \subset \bigcup \Gamma$. Hence $\bigcup \Gamma$ is open. □

Notice that $\bigcup \emptyset = \emptyset$, so the empty set is open.

**Theorem 16.4** *Let $\Gamma$ be a finite set of open sets. Then $\bigcap \Gamma$ is open.*

Proof: Let $x$ be a point in $\bigcap \Gamma$. Then $x$ is in each of the sets $S_k$ in $\Gamma$. Since each set $S_k$ is open, for each $S_k$ there is an $\epsilon_k > 0$ such that $B(x, \epsilon_k) \subset S_k$. Let $\epsilon$ be the minimum of the $\epsilon_k$. Since $\Gamma$ is finite, this number $\epsilon > 0$. Furthermore, $B(x, \epsilon) \subset S_k$ for each $k$. It follows that $B(x, \epsilon) \subset \bigcap \Gamma$. Hence $\bigcap \Gamma$ is open. □

Notice that under our conventions $\bigcap \emptyset = M$, so the entire space $M$ is open.

A subset $F$ of a metric space is *closed* if $\forall x\, (\forall \epsilon B(x, \epsilon) \cap F \neq \emptyset \Rightarrow x \in F)$. Here are some basic facts about closed sets.

**Theorem 16.5** *The closed subsets are precisely the complements of the open subsets.*

Proof: Let $U$ be a set and $F = M \setminus U$ be its complement. Then $x \in U \Rightarrow \exists \epsilon\, B(x, \epsilon) \subset U$ is logically equivalent to $\forall \epsilon\, \neg B(x, \epsilon) \subset U \Rightarrow x \notin U$. But this says $\forall \epsilon\, B(x, \epsilon) \cap F \neq \emptyset \Rightarrow x \in F$. From this it is evident that $F$ is closed precisely when $U$ is open. □

**Theorem 16.6** *A set $F$ in a metric space is an closed subset if and only if every convergent sequence $s : \mathbf{N} \to M$ with values $s_n \in F$ has limit $s_\infty \in F$.*

Proof: Suppose that $F$ is closed. Let $s$ be a convergent sequence with $s_n \in F$ for each $n$. Let $\epsilon > 0$. Then for $n$ sufficiently large $d(s_n, s_\infty) < \epsilon$, that is, $s_n \in B(s_\infty, \epsilon)$. This shows that $B(s_\infty, \epsilon) \cap F \neq \emptyset$. Since $\epsilon > 0$ is arbitrary, it follows that $s_\infty \in F$.

For the other direction, suppose that $F$ is not closed. Then there is a point $x \notin F$ such that $\forall \epsilon B(x, \epsilon) \cap F \neq \emptyset$. Then for each $n$ we have $B(x, 1/n) \cap F \neq \emptyset$. By the axiom of choice, we can choose $s_n \in B(x, 1/n) \cap F$. Clearly $s_n$ converges to $s_\infty = x$ as $n \to \infty$. Yet $s_\infty$ is not in $F$. □

Given an arbitrary subset $A$ of $M$, the *interior* $A^o$ of $A$ is the largest open subset of $A$. Similarly, the *closure* $\bar{A}$ of $A$ is the smallest closed superset of $A$. The set $A$ is *dense* in $M$ if $\bar{A} = M$.

## 16.7 Topological spaces

A *topological space* is a set $X$ together with a collection of subsets that is closed under unions and under finite intersections. These are the open subsets of $X$. More officially, a topology is a collection $\mathcal{T} \subset P(X)$ with the properties that $\Gamma \subset \mathcal{T}$ implies $\bigcup \Gamma \in \mathcal{T}$ and that $\Gamma \subset \mathcal{T}$, $\Gamma$ finite, implies $\bigcap \Gamma \in \mathcal{T}$.

Notice that it follows from the definition that $\bigcup \emptyset = \emptyset \in \mathcal{T}$ and that $\bigcap \emptyset = X \in \mathcal{T}$. (This last uses the convention that $X$ is the universe to which the intersection applies.) That is, the empty set and the whole space $X$ are always open subsets of $X$.

Every metric space defines a topological space. A property of a metric space that can be defined only in terms of the topology (the collection of open subsets) is called a topological property.

If $Y$ is a subset of a toplogical space $X$, then the *relative topology* of $Y$ consists of all the sets $U \cap Y$, where $U$ is an open subset of $X$. Thus every subset of a topological space is a topological space.

If $X$ is a metric space, and $Y$ is a subset of $X$, then $Y$ is also a metric space. The relative topology of $Y$ is then the same as the metric topology of $Y$.

A topological space is said to be *metrizable* if there is a metric that defines its topology.

**Theorem 16.7 (Urysohn's lemma (metric case))** *If $X$ is a metrizable topological space, then then for every pair $A, B$ of disjoint closed subsets there is a function $g : X \to [0, 1]$ that is zero on $A$ and one on $B$.*

Proof: let $A, B$ be disjoint closed sets. Then $x \mapsto d(x, A)$ and $x \mapsto d(x, B)$ are continuous functions that vanish precisely on $A$ and on $B$. Furthermore, $d(x, A) + d(x, B) > 0$ for every $x$. Let

$$g(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}. \tag{16.8}$$

Then $g$ is a continuous function from $X$ to $[0, 1]$ that is zero on $A$ and 1 on $B$. □

A topological space is *Hausdorff* if every pair of points is separated by a pair of disjoint open sets. It is clear that every metrizable topological space is Hausdorff.

A Hausdorff topological space is *regular* if every pair consisting of a closed set and a point not in the set is separated by a pair of disjoint open sets. A Hausdorff topological space is *normal* if every pair of disjoint closed sets is separated by a pair of disjoint open sets. It follows from the theorem that every metrizable topological space is normal. (Take the sets where $g < 1/3$ and where $g > 2/3$.)

A *base* for a topological space $X$ is a collection $\Gamma$ of open set such that every open set is a union of sets in $\Gamma$. In a metric space the balls $B(x, \epsilon)$ for $x$ in $X$ and $\epsilon > 0$ form a base.

A topological space $X$ is *second countable* if it has a countable base.

If $X$ is a topological space and $S$ is a subset, then $S$ is *dense* in $X$ if its closure is $X$.

A topological space $X$ is *separable* provided that there is a countable subset $S$ with closure $\bar{S} = X$. In other words, $X$ is separable if it has a countable dense subset.

**Theorem 16.8** *If $X$ is a second countable topological space, then $X$ is separable.*

Proof: Let $\Gamma$ be a countable base for the open subsets of $X$. Let $\Gamma' = \Gamma \setminus \{\emptyset\}$. Then $\Gamma'$ consists of non-empty sets. For each $U$ in $\Gamma'$ choose $x$ in $U$. Let $S$ be the set of all such $x$. Let $V = X \setminus \bar{S}$. Since $V$ is open, it is the union of those of its subsets that belong to $\Gamma$. Either there are no such subsets, or there is only the empty set. In either case, it must be that $V = \emptyset$. This proves that $\bar{S} = X$. □

**Theorem 16.9** *If $X$ is a separable metrizable topological space, then $X$ is second countable.*

Thus for metrizable spaces being separable is the same as being second countable. For a general topological space the most useful notion is that of being second countable. It is not true in general that a separable topological space is second countable.

The reason for introducing these concepts is that for second countable topological spaces it is relatively easy to characterize which spaces are metrizable. The *Urysohn metrization theorem* states that every second countable regular topological space is metrizable. A proof of this theorem may be found in Kelley [10]. Metrizable topological spaces are relatively common, and so it is reasonable to focus initially on them.

There are two reasons that a topological space might not be metrizable: it might not be second countable, or it might not regular. In later chapters it will become apparent that there are important examples of topological spaces that are not metrizable because they are very big, that is, not second countable. (A typical example is an uncountable product space.) However there are also simple and useful examples where where the space is not metrizable because it is not regular, or even Hausdorff.

Example: Here is an example of a topology that us useful in the theory of lower semicontinuous functions and hence in optimization. The underlying space for this topology is the set $(-\infty, +\infty]$. An non-trivial open set in this topology is defined to be an interval $(a, +\infty]$, where $a \in \mathbb{R}$. This topology is not metrizable. It is not even Hausdorff. Yet is is useful when describing a situation when $y$ is to be regarded as close to $y$ when either $y > x$ or $y$ is close to $x$ in the usual sense.

Example: Another simple example of how non-Hausdorff topological spaces is classification. Say that a topological space $X$ is classified into categories by some equivalence relation $E$. The *quotient space* $X/E$ consists of the equivalence classes. There is a classifying function $q : X \to X/E$ that sends each point of $X$ to its corresponding equivalence class. The topology on the quotient space $X/E$ is such that $U$ is an open subset of $X/E$ precisely when $q^{-1}[U]$ is an open subset of $X$.

As an example, say that one wants to classify the real numbers $\mathbb{R}$ into three categories: strictly negative, zero, strictly positive. The quotient space may be identified with the three points $-, 0, +$. There are $2^3 = 9$ subsets of this space, of which 6 are open. This space is not Hausdorff. The point is that although 0 is separated from each other real number, it is not separated from the set of strictly positive real numbers (or from the set of strictly negative real numbers).

This example looks less silly if one thinks of the problem of classifying the asymptotic behavior of the differential equation $dy/dt = -ky$ with $k > 0$. There are three kinds of asymptotic behaviors, depending on whether the initial value is strictly negative, zero, or strictly positive.

## 16.8 Continuity

Let $f$ be a map from a metric space $A$ to another metric space $B$. Then $f$ is said to be *continuous* at $a$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x$ we have that $d(x, a) < \delta$ implies $d(f(x), f(a)) < \epsilon$. O

Let $f$ be a map from a metric space $A$ to another metric space $B$. Then there are various notions of how it can respect the metric.

1. $f$ is a *contraction* if for all $x, y$ we have $d(f(x), f(y)) \leq d(x, y)$.
2. $f$ is *Lipschitz* (bounded slope) if there exists $M < \infty$ such that for all $x, y$ we have $d(f(x), f(y)) \leq Md(x, y)$.
3. $f$ is *uniformly continuous* if for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x, y$ we have that $d(x, y) < \delta$ implies $d(f(x), f(y)) < \epsilon$.
4. $f$ is *continuous* if for every $y$ and every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x$ we have that $d(x, y) < \delta$ implies $d(f(x), f(y)) < \epsilon$.

Clearly contraction implies Lipschitz implies uniformly continous implies continuous. The converse implications are false.

Let $A$ and $B$ be metric spaces. Suppose that there is a function $f : A \to B$ with inverse function $f^{-1} : B \to A$. There are various notions of equivalence of metric spaces.

1. The metric spaces are *isometric* if $f$ and $f^{-1}$ are both contractions.
2. The metric spaces are *Lipschitz equivalent* if $f$ and $f^{-1}$ are both Lipschitz.
3. The metric spaces are *uniformly equivalent* if $f$ and $f^{-1}$ are both uniformly continuous.
4. The metric spaces are *topologically equivalent* or topologically isomorphic or *homeomorphic*) if $f$ and $f^{-1}$ are both continuous.

Again there is a chain of implications for the various kinds of equivalence: isometric implies Lipschitz implies uniform implies topological.

The following theorem shows that continuity at a point is a topological property.

**Theorem 16.10** *Let $A$ and $B$ be metric spaces. Then $f : A \to B$ is continuous at $a$ if and only if for each open set $V \subset B$ with $f(a) \in V$ there is an open subset $U \subset A$ with $a \in U$ such that $f[U] \subset V$, or, what is the same, $U \subset f^{-1}[V]$.*

Proof: Suppose $f$ continuous at $a$. Consider an open set $V$ with $f(a) \in V$. Since $V$ is open, there is a ball $B(f(a), \epsilon) \subset V$. Since $f$ is continuous, there is a ball $U = B(a, \delta)$ such that $B(a, \delta) \subset f^{-1}[B(f(a), \epsilon)] \subset f^{-1}[V]$.

Suppose that the relation $f^{-1}$ satisfies the condition of the theorem at $a$. $\epsilon > 0$. The set $V = B(f(a), \epsilon)$ is open, so there is an open subset $U$ with $a \in U \subset f^{-1}[B(f(a), \epsilon)$. Since $U$ is open there is a $\delta > 0$ such that $B(a, \delta) \subset U \subset f^{-1}[B(f(a), \epsilon)]$. This shows that $f$ is continuous at $a$. □

The following theorem gives a particularly elegant description of continuity that shows that it is a topological property. It follows that the property of topological equivalence is also a topological property.

**Theorem 16.11** *Let $A$ and $B$ be metric spaces. Then $f : A \to B$ is continuous if and only if for each open set $V \subset B$, the set $f^{-1}[V] = \{x \in A \mid f(x) \in V\}$ is open.*

Proof: Suppose $f$ continuous. Then it is continuous at each point. Consider an open set $V$ whose inverse image under $f$ is not empty. Let $a$ be in $f^{-1}[V]$. Since $f$ is continuous at $a$ and $V$ is open, there is an open subset $U$ with $a \in U \subset f^{-1}[V]$. The union of the subsets $U$ for all $a$ in $f^{-1}[U]$ is $f^{-1}[V]$. So $f^{-1}[U]$ is open.

Suppose that the relation $f^{-1}$ maps open sets to open sets. Consider an $a$ and an open set $V$ with $f(a) \in V$. The set $U = f^{-1}[V]$ is open with $a \in U$. This shows that $f$ is continuous at $a$. Since this works for each $a$, it follows that $f$ is continuous. □

Example: As an example, let $f : X \to (-\infty, +\infty]$ be a real function, but take the topology for $\mathbb{R}$ to be the unusual one where the only non-trivial open sets are intervals $(a, +\infty)$, with $a$ real. The condition that $f$ is continuous in this sense is equivalent to the usual definition that $f$ is a lower semicontinuous function. Such functions are important in optimization problems. For instance, a lower semicontinuous function on a non-empty compact space always assumes its minimum (but not necessarily its maximum).

## 16.9 Uniformly equivalent metrics

Consider two metrics on the same set $A$. Then the identity function from $A$ with the first metric to $A$ with the second metric may be a contraction, Lipschitz, uniformly continuous, or continuous. There are corresponding notions of equivalence of metrics: the metrics may be the same, they may be Lipschitz equivalent, they may be uniformly equivalent, or they may be topologically equivalent.

For metric spaces the notion of uniform equivalence is particularly important. The following result shows that given a metric, there is a bounded metric that is uniformly equivalent to it. In fact, such a metric is

$$d_b(x, y) = \frac{d(x, y)}{1 + d(x, y)}. \tag{16.9}$$

The following theorem puts this in a wider context.

**Theorem 16.12** *Let $\phi : [0, +\infty) \to [0, +\infty)$ be a continuous function that satisfies the following three properties:*

1. *$\phi$ is increasing: $s \leq t$ implies $\phi(s) \leq \phi(t)$*
2. *$\phi$ is subadditive: $\phi(s + t) \leq \phi(s) + \phi(t)$*
3. *$\phi(t) = 0$ if and only if $t = 0$.*

*Then if $d$ is a metric, the metric $d'$ defined by $d'(x, y) = \phi(d(x, y))$ is also a metric. The identity map from the set with metric $d$ to the set with metric $d'$ is uniformly continuous with uniformly continuous inverse.*

Proof: The subadditivity is what is needed to prove the triangle inequality. The main thing to check is that the identity map is uniformly continuous in each direction.

Consider $\epsilon > 0$. Since $\phi$ is continuous at 0, it follows that there is a $\delta > 0$ such that $t < \delta$ implies $\phi(t) < \epsilon$. Hence if $d(x, y) < \delta$ it follows that $d'(x, y) < \epsilon$. This proves the uniform continuity in one direction.

The other part is also simple. Let $\epsilon > 0$. Let $\delta = \phi(\epsilon) > 0$. Since $\phi$ is increasing, $t \geq \epsilon \Rightarrow \phi(t) \geq \delta$, so $\phi(t) < \delta \Rightarrow t < \epsilon$. It follows that if $d'(x, y) < \delta$, then $d(x, y) < \epsilon$. This proves the uniform continuity in the other direction. □

In order to verify the subadditivity, it is sufficient to check that $\phi'(t)$ is decreasing. For in this case $\phi'(s + u) \leq \phi'(s)$ for each $u \geq 0$, so

$$\phi(s + t) - \phi(s) = \int_0^t \phi'(s + u)\, du \leq \int_0^t \phi'(u)\, du = \phi(t). \tag{16.10}$$

This works for the example $\phi(t) = t/(1+t)$. The derivative is $\phi'(t) = 1/(1+t)^2$, which is positive and decreasing.

## 16.10 Sequences

Consider a sequence $s : \mathbf{N} \to B$, where $B$ is a metric space. Then the *limit* of $s_n$ as $n \to \infty$ is $s_\infty$ provided that $\forall \epsilon > 0 \exists N \forall n \, (n \geq N \Rightarrow d(s_n, s_\infty) < \epsilon)$.

**Theorem 16.13** *If $A$ and $B$ are metric spaces, then $f : A \to B$ is continuous if and only if whenever $s$ is a sequence in $A$ converging to $s_\infty$, it follows that $f(s)$ is a sequence in $B$ converging to $f(s_\infty)$.*

Proof: Suppose that $f : A \to B$ is continuous. Suppose that $s$ is a sequence in $A$ converging to $s_\infty$. Consider arbitrary $\epsilon > 0$. Then there is a $\delta > 0$ such that $d(x, s_\infty) < \delta$ implies $d(f(x), f(s_\infty)) < \epsilon$. Then there is an $N$ such that $n \geq N$ implies $d(s_n, s_\infty) < \delta$. It follows that $d(f(s_n), f(s_\infty)) < \epsilon$. This is enough to show that $f(s)$ converges to $f(s_\infty)$.

The converse is not quite so automatic. Suppose that for every sequence $s$ converging to some $s_\infty$ the corresponding sequence $f(s)$ converges to $f(s_\infty)$. Suppose that $f$ is not continuous at some point $a$. Then there exists $\epsilon > 0$ such that for every $\delta > 0$ there is an $x$ with $d(x, a) < \delta$ and $d(f(x), f(a)) \geq \epsilon$. In particular, the set of $x$ with $d(x, a) < 1/n$ and $d(f(x), f(a)) \geq \epsilon$ is non-empty. By the axiom of choice, for each $n$ there is an $s_n$ in this set. Let $s_\infty = a$. Then $d(s_n, s_\infty) < 1/n$ and $d(f(s_n), f(s_\infty)) \geq \epsilon$. This contradicts the hypothesis that $f$ maps convergent sequences to convergent sequences. Thus $f$ is continuous at every point. □

One way to make this definition look like the earlier definitions is to define a metric on $\mathbb{N}_+$. Set

$$d^*(m, n) = \left| \frac{1}{m} - \frac{1}{n} \right|. \tag{16.11}$$

We may extend this to a metric on $\mathbb{N}_+ \cup \{\infty\}$ if we set $1/\infty = 0$.

**Theorem 16.14** *With the metric $d^*$ on $\mathbb{N}_+ \cup \{\infty\}$ defined above, the limit of $s_n$ as $n \to \infty$ is $s_\infty$ if and only if the function $s$ is continuous from the metric space $\mathbb{N}_+ \cup \{\infty\}$ to $B$.*

Proof: The result is obvious if we note that $n > N$ is equivalent to $d^*(n, \infty) = 1/n < \delta$, where $\delta = 1/N$. □

Another important notion is that of *Cauchy sequence*. A sequence $s : \mathbf{N} \to B$ is a Cauchy sequence if $\forall \epsilon \exists N \forall m \forall n \, ((m \geq N \wedge n \geq N) \Rightarrow d(s_m, s_n) < \epsilon)$.

**Theorem 16.15** *If we use the $d^*$ metric on $\mathbb{N}_+$ defined above, then for every sequence $s : \mathbb{N}_+ \to B$, $s$ is a Cauchy sequence if and only if $s$ is uniformly continuous.*

Proof: Suppose that $s$ is uniformly continuous. Then $\forall\epsilon > 0 \exists\delta > 0 (|1/m - 1/n| < \delta \Rightarrow d(s_m, s_n) < \epsilon)$. Temporarily suppose that $\delta'$ is such that $|1/m - 1/n| < \delta \Rightarrow d(s_m, s_n) < \epsilon)$. Take $N$ with $2/\delta' < N$. Suppose $m \geq N$ and $n \geq N$. Then $|1/m - 1/n| \leq 2/N < \delta'$. Hence $d(s_m, s_n) < \epsilon$. Thus $(m \geq N \wedge n \geq N) \Rightarrow d(s_m, s_n) < \epsilon$. From this it is easy to conclude that $s$ is a Cauchy sequence.

Suppose on the other hand that $s$ is a Cauchy sequence. This means that $\forall\epsilon > 0 \exists N \forall m \forall n \, ((m \geq N \wedge n \geq N) \Rightarrow d(s_m, s_n) < \epsilon)$. Temporarily suppose that $N'$ is such that $\forall m \forall n \, ((m \geq N' \wedge n \geq N') \Rightarrow d(s_m, s_n) < \epsilon)$. Take $\delta = 1/(N'(N'+1))$. Suppose that $|1/m - 1/n| < \delta$. Either $m < n$ or $n < m$ or $m = n$. In the first case, $1/(m(m+1)) = 1/m - 1/(m+1) < 1/m - 1/n < 1/(N'(N'+1))$, so $m > N'$, and hence also $n > N'$. So $d(s_m, s_n) < \epsilon$ Similarly, in the second case both $m > N'$ and $n > N'$, and again $d(s_m, s_n) < \epsilon$. Finally, in the third case $m = n$ we have $d(s_m, s_n) = 0 < \epsilon$. So we have shown that $|1/m - 1/n| < \delta \Rightarrow d(s_m, s_n) < \epsilon$. □

## 16.11 Supplement: Lawvere metrics and semicontinuity

There is a generalization of the notion of metric space that includes both pseudometric spaces and ordered sets. The fundamental idea is that of a *Lawvere metric*. A Lawvere metric space is a set $M$ together with a function $d : M \times M \to [0, +\infty]$ with the following two properties:

1. For all $x$ we have $d(x, x) = 0$
2. For all $x, y, z$ we have $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

This is a considerable generalization of the notion of pseudometric space. In fact, what we have is an extended pseudo-quasimetric. Not only are the values 0 and $+\infty$ allowed as distances, but also the symmetry property is dropped.

Lawvere metric is not the usual name for this concept. In various places it is called a generalized metric or a quasi-pseudometric or an extended pseudo-quasimetric. Unfortunately the terms "generalized" and "quasi" are used in other ways in various branches of mathematics. The term Lawvere metric seems appropriate, since Lawvere [11] recognized the significance of this notion in the context of category theory.

Each Lawvere metric defines a notion of open ball. The ball $B(x, \epsilon)$ consists of all $y$ with $d(x, y) < \epsilon$. There is also a corresponding topology, where the open sets are unions of open balls.

The notion of Lawvere metric includes the notion of ordered set. Let $d(x, y) = 0$ if $x \leq y$. Otherwise let $d(x, y) = +\infty$. Then the Lawvere metric axioms are satisfied. This shows a Lawvere metric space is a quantitative versions of an ordered set. From this point of view an open ball in an ordered set is an interval of the form $\{y \mid x \leq y\}$.

There are yet other important examples of Lawvere metrics. Consider the interval of extended real numbers $[-\infty, +\infty)$. Define the *upper Lawvere metric* on this interval by $d_U(x, y) = y - x$ if $x \leq y$, otherwise $d_U(x, y) = 0$ if $y \leq x$. Thus it costs the usual amount to go upward, but going downward is free. The epsilon ball about $x$ is the set of all $y$ with $d_U(x, y) < \epsilon$. This is just the set of all $y$ with $y < x + \epsilon$.

Similarly, consider the interval of extended real numbers $(-\infty, +\infty]$, and define the *lower Lawvere metric* by $d_L(x, y) = x - y$ if $y \leq x$ and $d_L(x, y) = 0$ for $x \leq y$. Now one pays a price for going downward.

A function from a metric space $M$ to $[-\infty, +\infty)$ is said to be *upper semi-continuous* if for every $u$ and every $\epsilon > 0$ there is a $\delta > 0$ such that all $v$ with $d(u, v) < \delta$ satisfy $d_U(f(u), f(v)) < \epsilon$, that is, $f(v) < f(u)+\epsilon$. An example of an upper semicontinuous function is one that is continuous except where it jumps up at a single point. It is easy to fall from this peak. The indicator function of a closed set is upper semicontinuous. The infimum of a non-empty collection of upper semicontinuous functions is upper semicontinuous. This generalizes the statement that the intersection of a collection of closed sets is closed.

There is a corresponding notion of *lower semicontinuous function* from $M$ to $(-\infty, +\infty]$. An example of a lower semicontinuous function is one that is continuous except where jumps down at a single point. The indicator function of an open set is lower semicontinuous. The supremum of a non-empty collection of lower semicontinuous functions is lower semicontinuous. This generalizes the fact that the union of a collection of open sets is open.

# Problems

1. In this problem $\ell^p$ denotes the space of real functions $x$ defined on the strictly positive natural numbers with $\|x\|_p^p = \sum_{n=1}^{\infty} |x_n|^p < +\infty$. Suppose that $n \mapsto c_n = na_n$ is in $\ell^2$. Show that $n \mapsto a_n$ is in $\ell^1$.

2. Here $[0, 1]$ is the closed interval of real numbers. Consider $[0, 1]^{\mathbf{N}_+}$ with uniform metric $d_\infty$ (a subset of the metric space $\ell^\infty$). Also, consider the Hilbert cube $[0, 1]^{\mathbf{N}_+}$ with product metric $d_p$ given by $d_p(x, y) = \sum_{n=1}^{\infty} |x_n - y_n|/2^n$. Is the identity function $\iota_1$ from $[0, 1]^{\mathbf{N}_+}, d_\infty$ to $[0, 1]^{\mathbf{N}_+}, d_p$ continuous? Is the identity function $\iota_2$ from $[0, 1]^{\mathbf{N}_+}, d_p$ to $[0, 1]^{\mathbf{N}_+}, d_\infty$ continuous? In each case answer yes or no, and provide a proof or a counterexample.

3. Regard $\ell^2$ as a subset of $\mathbb{R}^\infty$. Find a sequence of points in the unit sphere of $\ell^2$ that converges in the $\mathbb{R}^\infty$ sense to zero.

4. Let $X$ be a metric space. Give a careful proof using precise definitions that $BC(X)$ is a closed subset of $B(X)$.

5. Give four examples of bijective functions from $\mathbb{R}$ to $\mathbb{R}$: an isometric equivalence, a Lipschitz but not isometric equivalence, a uniform but not Lipschitz equivalence, and a topological but not uniform equivalence.

6. Show that for $F$ a linear transformation of a normed vector space to itself, $F$ continuous at zero implies $F$ Lipschitz (bounded slope).

7. Let $K$ be an infinite matrix with $\|K\|_{1,\infty} = \sup_n \sum_m |K_{mn}| < \infty$. Show that $F(x)_m = \sum_n K_{mn} x_n$ defines a Lipschitz function from $\ell^1$ to itself.

8. Let $K$ be an infinite matrix with $\|K\|_{\infty,1} = \sup_m \sum_n |K_{mn}| < \infty$. Show that $F(x)_m = \sum_n K_{mn} x_n$ defines a Lipschitz function from $\ell^\infty$ to itself.

9. Let $K$ be an infinite matrix with $\|K\|_{2,2}^2 = \sum_m \sum_n |K_{mn}|^2 < \infty$. Show that $F(x)_m = \sum_n K_{mn} x_n$ defines a Lipschitz function from $\ell^2$ to itself.

10. Let $K$ be an infinite matrix with $\|K\|_{1,\infty} < \infty$ and $\|K\|_{\infty,1} < \infty$. Show that $F(x)_m = \sum_n K_{mn} x_n$ defines a Lipschitz function from $\ell^2$ to itself.

11. Let $X$ be a topological space. Say that $X$ is *connected* if there is no partition of $X$ into two open subsets. A subset $A$ of $X$ is connected if $A$ is connected with respect to the relative topology (where a subset of $A$ is open if it is the intersection of an open set of $X$ with $A$). Show that if $A$ is connected, then $\bar{A}$ is connected.

12. Let $X$ be a topological space. Say $p \sim q$, or $p$ is connected to $q$, if there exists a connected subset $A$ of $X$ with $p \in A$ and $q \in A$. This is an equivalence relation, and the equivalence classes are called the connected components of $X$. If $p$ is in $X$, let $C_p$ be the connected component with $p \in C_p$. Show that $C_p$ is connected. Show that $C_p$ is the largest connected set with $p \in C$. Show that $C_p$ is closed.

13. Let $X$ be a topological space. Say $p \leftrightarrow q$ if there there is no open partition $U, V$ of $X$ with $p \in U$ and $q \in X$. (We might say in this case that $p$ is "allied" with $q$.) This is also an equivalence relation. Show that $p \sim q$ implies $p \leftrightarrow q$.

14. Let $g : \{0,1\}_+^{\mathbb{N}} \to \mathbf{R}$ be given by $g(x) = \sum_{k=1}^\infty 2x_k/3^k$. This is a bijection from the Cantor space to the middle third Cantor set. Let $d_p(x,y) = \sum_{k=1}^\infty |x_k - y_k|/2^k$ be the metric on the Cantor space. Show that the Cantor space and the Cantor set are uniformly equivalent. Hint: Show that $|g(x) - g(y)| \leq 2d_p(x,y)$. Show that if $|g(x) - g(y)| < 1/3^m$, then $d_p(x,y) \leq 1/2^{m-1}$.

15. Consider the Cantor space or the Cantor set. Show that each pair of distinct points is disconnected. Hint: Show that the points are not allied.

16. A $G_\delta$ subset of a topological space is a countable intersection of open subsets. What is the cardinality of the collection of $G_\delta$ subsets of $\mathbb{R}^n$? Establish your result by giving upper and lower estimates.

# Chapter 17

# Metric spaces and metric completeness

## 17.1 Completeness

Let $A$ be a metric space. Then $A$ is *complete* means that every Cauchy sequence with values in $A$ converges. In this section we give an alternative perspective on completeness that makes this concept seem particularly natural.

If $z$ is a point in a metric space $A$, then $z$ defines a function $f_z : A \to [0, +\infty)$ by

$$f_z(x) = d(z, x). \tag{17.1}$$

This function has the following three properties:

1. $f_z(y) \leq f_z(x) + d(x, y)$
2. $d(x, y) \leq f_z(x) + f_z(y)$
3. $\inf f_z = 0$.

Say that a function $f : A \to [0, +\infty)$ is a virtual point if it has the three properties:

1. $f(y) \leq f(x) + d(x, y)$
2. $d(x, y) \leq f(x) + f(y)$
3. $\inf f = 0$.

We shall see that a metric space is complete if and only if every virtual point is a point. That is, it is complete iff whenever $f$ is a virtual point, there is a point $z$ in the space such that $f = f_z$.

It will be helpful later on to notice that the first two conditions are equivalent to $|f(y) - d(x, y)| \leq f(x)$. Also, it follows from the first condition and symmetry that $|f(x) - f(y)| \leq d(x, y)$. Thus virtual points are contractions, and in particular they are continuous.

**Theorem 17.1** *A metric space is complete if and only if every virtual point is given by a point.*

Proof: Suppose that every virtual point is a point. Let $s$ be a Cauchy sequence of points in $A$. Then for each $x$ in $A$, $d(s_n, x)$ is a Cauchy sequence in $\mathbb{R}$. This is because $|d(s_m, x) - d(s_n, x)| \leq d(s_m, s_n)$. However every Cauchy sequence in $\mathbb{R}$ converges. Define $f(x) = \lim_{n\to\infty} d(s_n, x)$. It is easy to verify that $f$ is a virtual point. By assumption it is given by a point $z$, so $f(x) = f_z(x) = d(z, x)$. But $d(s_n, z)$ converges to $f(z) = d(z, z) = 0$, so this shows that $s_n \to z$ as $n \to \infty$.

Suppose on the other hand that every Cauchy sequence converges. Let $f$ be a virtual point. Let $s_n$ be a sequence of points such that $f(s_n) \to 0$ as $n \to \infty$. Then $d(s_m, s_n) \leq f(s_m) + f(s_n) \to 0$ as $m, n \to \infty$, so $s_n$ is a Cauchy sequence. Thus it must converges to a limit $z$. Since $f$ is continuous, $f(z) = 0$. Furthermore, $|f(y) - d(z, y)| \leq f(z) = 0$, so $f = f_z$. □

**Theorem 17.2** *Let $A$ be a dense subset of the metric space $\bar{A}$. Let $M$ be a complete metric space. Let $f : A \to M$ be uniformly continuous. Then there exists a unique uniformly continuous function $\bar{f} : \bar{A} \to M$ that extends $f$.*

Proof: Regard the function $f$ as a subset of $\bar{A} \times M$. Define the relation $\bar{f}$ to be the closure of $f$. If $x$ is in $\bar{A}$, let $s_n \in A$ be such that $s_n \to x$ as $n \to \infty$. Then $s_n$ is a Cauchy sequence in $A$. Since $f$ is uniformly continuous, it follows that $f(s_n)$ is a Cauchy sequence in $M$. Therefore $f(s_n)$ converges to some $y$ in $M$. This shows that $(x, y)$ is the relation $\bar{f}$. So the domain of $\bar{f}$ is $\bar{A}$.

Let $\epsilon > 0$. By uniform continuity there is a $\delta > 0$ such that for all $x, u'$ in $A$ we have that $d(x', u') < \delta$ implies $d(f(x'), f(u')) < \epsilon/3$.

Now let $(x, y) \in \bar{f}$ and $(u, v) \in \bar{f}$ with $d(x, u) < \delta/3$. There exists $x'$ in $A$ such that $f(x') = y'$ and $d(x', x) < \delta/3$ and $d(y', y) < \epsilon/3$. Similarly, there exists $u'$ in $A$ such that $f(u') = v'$ and $d(u', u) < \delta/3$ and $d(v', v) < \epsilon/3$. It follows that $d(x', u') \leq d(x', x) + d(x, u) + d(u, u') < \delta$. Hence $d(y, v) \leq d(y, y') + d(y', v') + d(v', v) < \epsilon$. Thus $d(x, u) < \delta/3$ implies $d(y, v) < \epsilon$. This is enough to show that $\bar{f}$ is a function and is uniformly continuous. □

A *normed vector space* is a vector space with a norm. A *Banach space* is a vector space with a norm that is a complete metric space. Here are examples of complete metric spaces. All of them except for $\mathbb{R}^\infty$ are Banach spaces. Notice that $\ell^\infty$ is the special case of $B(X)$ when $X$ is countable. For $BC(X)$ we take $X$ to be a metric space, so that the notion of continuity is defined.

Examples:

1. $\mathbb{R}^n$ with either the $\ell^1_n$, $\ell^2_n$, or $\ell^\infty_n$ metric.
2. $\ell^1$.
3. $\ell^2$.
4. $\ell^\infty$.

5. $\mathbb{R}^\infty$ with the product metric.
6. $B(X)$ with the uniform metric.
7. $BC(X)$ with the uniform metric.

In these examples the points of the spaces are real functions. There are obvious modifications where one instead uses complex functions. Often the same notation is used for the two cases, so one must be alert to the distinction.

## 17.2 Uniform equivalence of metric spaces

**Theorem 17.3** *Let $A$ be a metric space, and let $M$ be a complete metric space. Suppose that there is a uniformly continuous bijection $f : A \to M$ such that $f^{-1}$ is continuous. Then $A$ is complete.*

Proof: Suppose that $n \mapsto s_n$ is a Cauchy sequence with values in $A$. Since $f$ is uniformly continuous, the composition $n \mapsto f(s_n)$ is a Cauchy sequence in $M$. Since $M$ is complete, there is a $y$ in $M$ such that $f(s_n) \to y$ as $n \to \infty$. Let $x = f^{-1}(y)$. Since $f^{-1}$ is continuous, it follows that $s_n \to x$ as $n \to \infty$. □

**Corollary 17.4** *The completeness property is preserved under uniform equivalence.*

It is important to understand that completeness is not a topological invariant. For instance, take the function $g : \mathbb{R} \to (-1, 1)$ defined by $g(x) = \sinh(x)$. This is a topological equivalence. Yet $\mathbb{R}$ is complete, while $(-1, 1)$ is not complete.

It is customary to define a metric on $[-\infty, +\infty]$ that makes it a complete metric space. One way to do this is to define the map $h : [-\infty, +\infty] \to [-1, 1]$ by $h(x) = \sinh(x)$ for $x$ in $\mathbb{R}$, while $h(_\infty) = -1$ and $h(+\infty) = 1$. Then the distance between two points in $[-\infty, +\infty]$ is the usual distance between their images under $h$. However, one must be careful. With this metric on $[-\infty, +\infty]$ the subset $(-\infty, +\infty)$ has its usual topology, but does not inherit its usual metric. In fact, the subset $(-\infty, +\infty)$ is not complete with respect to the inherited metric.

## 17.3 Completion

**Theorem 17.5** *Every metric space is densely embedded in a complete metric space.*

This theorem says that if $A$ is a metric space, then there is a complete metric space $F$ and an isometry from $A$ to $F$ with dense range.

Proof: Let $F$ consist of all the virtual points of $A$. These are continuous functions on $A$. The distance $\bar{d}$ between two such functions is the usual sup norm $\bar{d}(f, g) = \sup_{x \in A} d(f(x), g(x))$. It is not hard to check that the virtual points form a complete metric space of continuous functions. The embedding sends each point $z$ in $a$ into the corresponding $f_z$. Again it is easy to verify that this embedding preserves the metric, that is, that $\bar{d}(f_z, f_w) = d(z, w)$. Furthermore, the range of this embedding is dense. The reason for this is that for each virtual point $f$ and each $\epsilon > 0$ there is an $x$ such that $f(x) < \epsilon$. Then $|f(y) - f_x(y)| = |f(y) - d(x, y)| \leq f(x) < \epsilon$. This shows that $\bar{d}(f, f_x) \leq \epsilon$. □

The classic example is the completion of the rational number system **Q**. A virtual point of **Q** is a function whose graph is in the general shape of a letter V. When the bottom tip of the V is at a rational number, then the virtual point is already a point. However most of these V functions have tips that point to a gap in the rational number system. Each such gap in the rational number system corresponds to the position of an irrational real number in the completion.

## 17.4 The Banach fixed point theorem

If $f$ is a Lipschitz function from a metric space to another metric space, then there is a constant $C < +\infty$ such that for all $x$ and $y$ we have $d(f(x), f(y)) \leq Cd(x, y)$. The set of all $C$ is a set of upper bounds for the quotients, and so there is a least such upper bound. This is called the least Lipschitz constant of the function.

A Lipschitz function is a contraction if its least Lipschitz constant is less than or equal to one. It is a *strict contraction* if its least Lipschitz constant is less than one.

**Theorem 17.6 (Banach fixed point theorem)** *Let $A$ be a complete metric space. Let $f : A \to A$ be a strict contraction. Then $f$ has a unique fixed point. For each point in $A$, its orbit converges to the fixed point.*

Proof: Let $a$ be a point in $A$, and let $s_k = f^{(k)}(a)$. Then by induction $d(s_k, s_{k+1}) \leq M^k d(s_0, s_1)$. Then again by induction $d(s_m, s_{m+p})) \leq \sum_{k=m}^{p-1} M^k d(s_0, s_1) \leq K^m/(1-K)d(s_0, s_1)$. This is enough to show that $s$ is a Cauchy sequence. By completeness it converges to some $s_\infty$. Since $f$ is continuous, this is a fixed point. □

Recall that a Banach space is a complete normed vector space. The Banach fixed point theorem applies in particular to a linear transformations of a Banach space to itself that is a strict contraction.

For instance, consider one of the Banach spaces of sequences. Let $f(x) = Kx + u$, where $K$ is a matrix, and where $u$ belongs to the Banach space. The function $f$ is Lipschitz if and only if multiplication by $K$ is Lipschitz. If the Lipschitz constant is strictly less than one, then the Banach theorem gives the solution of the linear system $x - Kx = u$.

To apply this, first look at the Banach space $\ell^\infty$. Define $\|K\|_{\infty \to \infty}$ to be the least Lipschitz constant. Define

$$\|K\|_{\infty,1} = \sup_m \sum_{n=1}^{\infty} |K_{mn}|. \tag{17.2}$$

Then it is not difficult to see that $\|K\|_{\infty\to\infty} = \|K\|_{\infty,1}$.

For another example, consider the Banach space $\ell^1$. Define $\|K\|_{1\to 1}$ to be the least Lipschitz constant. Define

$$\|K\|_{1,\infty} = \sup_n \sum_{m=1}^{\infty} |K_{mn}|. \tag{17.3}$$

Then it is not difficult to see that $\|K\|_{1\to 1} = \|K\|_{1,\infty}$.

The interesting case is the Hilbert space $\ell^2$. Define $\|K\|_{2\to 2}$ to be the least Lipschitz constant. Define

$$\|K\|_{2,2} = \sqrt{\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} K_{mn}^2}. \tag{17.4}$$

Then an easy application of the Schwarz inequality will show that $\|K\|_{2\to 2} \leq \|K\|_{2,2}$. However this is usually not an equality!. A somewhat more clever application of the Schwarz inequality will show that $\|K\|_{2\to 2} \leq \sqrt{\|K\|_{1,\infty}\|K\|_{\infty,1}}$. Again this is not in general an equality. Finding the least Lipschitz constant is a non-trivial task. However one or the other of these two results will often give useful information.

## 17.5 Coerciveness

A continuous function defined on a compact space assumes its minimum (and its maximum). This result is both simple and useful. However in general the point where the minimum is assumed is not unique. Furthermore, the condition that the space is compact is too strong for many applications. A result that only uses completeness could be helpful, and the following is one of the most useful results of this type.

**Theorem 17.7** *Let $M$ be a complete metric space. Let $f$ be a continuous real function on $M$ that is bounded below. Let $a = \inf\{f(x) \mid x \in M\}$. Suppose that there is an increasing function $\phi$ from $[0, +\infty)$ to itself such that $\phi(t) = 0$ only for $t = 0$ with the coercive estimate*

$$a + \phi(d(x, y)) \leq \frac{f(x) + f(y)}{2}. \tag{17.5}$$

*Then there is a unique point $p$ where $f(p) = a$. That is, there exists a unique point $p$ where $F$ assumes its minimum value.*

Proof: Let $s_n$ be a sequence of points such that $f(s_n) \to a$ as $n \to \infty$. Consider $\epsilon > 0$. Let $\delta = \phi(\epsilon) > 0$. Since $\phi$ is increasing, $\phi(t) < \delta$ implies $t < \epsilon$. For large enough $m, n$ we can arrange that $\phi(d(s_m, s_n)) < \delta$. Hence $d(s_m, s_n) < \epsilon$. Thus $s_n$ is a Cauchy sequence. Since $M$ is complete, the sequence converges to some $p$ in $M$. By continuity, $f(p) = a$. Suppose also that $f(q) = a$. Then from the inequality $d(p, q) = 0$, so $p = q$. □

This theorem looks impossible to use in practice, because it seems to requires a knowledge of the infimum of the function. However the following result shows that there is a definite possibility of a useful application.

**Corollary 17.8** *Let $M$ be a closed convex subset of a Banach space. Let $f$ be a continuous real function on $M$. Say that $a = \inf_{x \in M} f(x)$ is finite and that there is a $c > 0$ such that the strict convexity condition*

$$c\|x - y\|^2 \leq \frac{f(x) + f(y)}{2} - f(\frac{x + y}{2}) \tag{17.6}$$

*is satisfied. Then there is a unique point $p$ in $M$ with $f(p) = a$.*

Proof: Since $M$ is convex, $(x + y)/2$ is in $M$, and so $a \leq f((x + y)/2)$. □

## 17.6 Supplement: The regulated integral

The traditional integral used in rigorous treatments of calculus is the Riemann integral. This may be developed using ideas involving order. By contrast, the regulated integral is based on metric space ideas. It is even simpler, but it is sufficient for many purposes. The functions that are integrable in this sense are known as regulated functions. Each continuous function is regulated, so this notion of integral is good for many calculus application. Furthermore, it works equally well for integrals with values in a Banach space.

Let $[a, b] \subset \mathbf{R}$ be a closed interval. Consider a partition $a \leq a_0 < a_1 < \ldots < a_n = b$ of the interval. A *general step function* is a function $f$ from $[a, b]$ to $\mathbf{R}$ that is constant on each open interval $(a_i, b_{i+1})$ of such a partition. For each general step function $f$ there is an integral $\lambda(f)$ that is the sum

$$\lambda(f) = \int_a^b f(x)\, dx = \sum_{i=0}^{n-1} f(c_i)(a_{i+1} - a_i), \tag{17.7}$$

where $a_i < c_i < a_{i+1}$.

Let $R([a, b])$ be the closure of the space $S$ of general step functions in the complete metric space $B([a, b])$ consisting of all bounded real functions. This called the space of *regulated functions*. Since every continuous function is a regulated function, we have $C([a, b]) \subset R([a, b])$.

The function $\lambda$ defined on the space $S$ of general step functions is a Lipschitz function with Lipschitz constant $b - a$. In particular it is uniformly continuous, and so it extends by uniform continuity to a function on the closure $R([a, b])$.

This extended function is also denoted by $\lambda$ and is the regulated integral. In particular, the regulated integral is defined on $C([a,b])$ and agrees with the integral for continuous functions that is used in elementary calculus.

# Problems

1. The usual definition of Cauchy sequence $\forall\epsilon > 0\exists N\forall m \geq N\forall n \geq N\, d(s_m, s_n) < \epsilon$ involves four quantifiers. It is proposed to replace this with a new definition involving three quantifiers of the form $\forall\epsilon > 0\exists N\forall n \geq N\, d(s_N, s_n) < \epsilon$. Is this equivalent? For each direction of the implication, give a proof or give a counterexample.

2. Consider the space $\ell^2$ of real sequences $j \mapsto s_j$ for which the norm $\|s\|_2 = \sqrt{\sum_{j=1}^{\infty} s_j^2} < +\infty$. Let $\delta_n$ be the unit basis vector determined by $\delta_{nj} = 1$ if $j = n$, zero otherwise. Let $n \mapsto a_n$ be a sequence in $R^\infty$. Let $S$ be the subset consisting of the points $a_n\delta_n$. Give necessary and sufficient conditions for $S$ to be totally bounded, with proof.

3. Let $c_0$ be the subset of $\ell^\infty$ consisting of all sequences that converge to zero. Show that $c_0$ is a complete metric space.

4. A sequence $s$ with values on a metric space $M$ is said to be *fast Cauchy* if $\sum_{n=1}^{\infty} d(s_n, s_{n+1}) < +\infty$. Prove that every Cauchy sequence has a fast Cauchy subsequence. Prove that $M$ is complete if and only if every fast Cauchy sequence is convergent. Note: A sequence is exponentially fast Cauchy if $d(s_n, s_{n+1} \leq 1/2^{n+1}$. This is more concrete, and the same results hold.

5. A map $f : M \to N$ of metric spaces is said to be *open map* if for every $x$ in $M$ and every $\epsilon > 0$ there exists a $\delta > 0$ such that we have $B(f(x), \delta) \subset f[B(x, \epsilon)]$. Prove that if $f$ is open if and only if it sends open subsets of $M$ to open subsets of $N$.

6. A map $f : M \to N$ of metric spaces is said to be *uniformly open* if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for all $x \in N$ we have $B(f(x), \delta) \subset f[B(x, \epsilon)]$. Prove that if $M$ is complete and $f : M \to N$ is continuous, uniformly open, and surjective, then $N$ is complete. Hint: Let $n \mapsto y_n$ be a Cauchy sequence in $N$. Constructive inductively an exponentially fast Cauchy sequence $k \mapsto x_k$ in $M$ such that $f(x_k) = y_{N_k}$.

7. Let $A$ be a dense subset of the metric space $\bar{A}$. Let $M$ be a complete metric space. Let $f : A \to M$ be continuous. It does not follow in general that there is a continuous function $\bar{f} : \bar{A} \to M$ that extends $f$. (a) Give an example of a case when the closure $\bar{f}$ of the graph is a function on $A$ but is not defined on $\bar{A}$. (b) Give an example when the closure $\bar{f}$ of the graph is a relation defined on $\bar{A}$ but is not a function.

8. Let $C([0,1])$ be the space of continuous real functions on the closed unit interval. Give it the metric $d_1(f,g) = \int_0^1 |f(x) - g(x)|\,dx$. Let $h$ be a discontinuous step function equal to 0 on half the interval and to 1 on the other half. Show that the map $f \mapsto \int_0^1 |f(x) - h(x)|\,dx$ is a virtual point of $C([0,1])$ (with the $d_1$ metric) that does not come from a point of $C([0,1])$.

9. Let $E$ be a complete metric space. Let $f : E \to E$ be a strict contraction with constant $C < 1$. Consider $z$ in $E$ and $r$ with $r \geq d(f(z), z)/(1 - C)$. Then $f$ has a fixed point in the ball consisting of all $x$ with $d(x, z) \leq r$. Hint: First show that this ball is a complete metric space.

10. Prove: A metric space $M$ is complete if and only if for every sequence $B_i$ of epsilon balls in $M$ the conditions 1) for each $i$ the ball $B_{i+1}$ is a subset of $B_i$ and 2) the sequence $\epsilon_i$ of the radii of the epsilon balls converges to zero together imply that there is a unique point $z$ of $M$ in each $B_i$. (The importance of this idea is seen in the very similar construction in the proof of the Baire category theorem of the next chapter.)

# Chapter 18

# Metric spaces and compactness

## 18.1 Total boundedness

The notion of compactness is meaningful and important in general topological spaces. However it takes a quantitative form in metric spaces, and so it is worth making a special study in this particular setting. A metric space is complete when it has no nearby missing points (that is, when every virtual point is a point). It is compact when, in addition, it is well-approximated by finite sets. The precise formulation of this approximation property is in terms of the following concept.

A metric space $M$ is *totally bounded* if for every $\epsilon > 0$ there exists a finite subset $F$ of $M$ such that the open $\epsilon$-balls centered at the points of $F$ cover $M$.

We could also define $M$ to be totally bounded if for every $\epsilon > 0$ the space $M$ is the union of finitely many sets each of diameter at most $2\epsilon$. For some purposes this definition is more convenient, since it does not require the sets to be balls.

The notion of total boundedness is quantitative. If $M$ is a metric space, then there is a function that assigns to each $\epsilon > 0$ the smallest number $N$ such that $M$ is the union of $N$ sets each of diameter at most $2\epsilon$. The slower the growth of this function, the better the space is approximated by finitely many points.

For instance, consider a box of side $2L$ in a Euclidean space of dimension $k$. Then the $N$ is roughly $(L/\epsilon)^k$. This shows that the covering becomes more difficult as the size $L$ increases, but also as the dimension $k$ increases.

**Theorem 18.1** *Let $f : K \to M$ be a uniformly continuous surjection. If $K$ is totally bounded, then $M$ is totally bounded.*

**Corollary 18.2** *Total boundedness is invariant under uniform equivalence of metric spaces.*

## 18.2 Compactness

For metric spaces we can say that a metric space is *compact* if it is both complete and totally bounded.

**Lemma 18.3** *Let $K$ be a metric space. Let $F$ be a subset of $K$. If $F$ is complete, then $F$ is a closed subset of $K$. Suppose in addition that $K$ is complete. If $F$ is a closed subset of $K$, then $F$ is complete.*

Proof: Suppose $F$ is complete. Say that $s$ is a sequence of points in $F$ that converges to a limit $a$ in $K$. Then $s$ is a Cauchy sequence in $F$, so it converges to a limit in $F$. This limit must be $a$, so $a$ is in $F$. This proves that $F$ is a closed subset of $K$. Suppose for the converse that $K$ is complete and $F$ is closed in $K$. Let $s$ be a Cauchy sequence in $F$. Then it converges to a limit $a$ in $K$. Since $F$ is closed, the point $a$ must be in $F$. This proves that $F$ is complete. □

**Lemma 18.4** *Let $K$ be a totally bounded metric space. Let $F$ be a subset of $K$. Then $F$ is totally bounded.*

Proof: Let $\epsilon > 0$. Then $K$ is the union of finitely many sets, each of diameter bounded by $2\epsilon$. Then $F$ is the union of the intersections of these sets with $F$, and each of these intersections has diameter bounded by $2\epsilon$. □

**Theorem 18.5** *Let $K$ be a compact metric space. Let $F$ be a subset of $K$. Then $F$ is compact if and only if it is a closed subset of $K$.*

Proof: Since $K$ is compact, it is complete and totally bounded. Suppose $F$ is compact. Then it is complete, so it is a closed subset of $K$. For the converse, suppose $F$ is a closed subset of $K$. It follows that $F$ is complete. Furthermore, from the last lemma $F$ is totally bounded. It follows that $F$ is compact. □

Examples:

1. The unit sphere (cube) in $\ell^\infty$ is not compact. In fact, the unit basis vectors $\delta_n$ are spaced by 1.
2. The unit sphere in $\ell^2$ is not compact. The unit basis vectors $\delta_n$ are spaced by $\sqrt{2}$.
3. The unit sphere in $\ell^1$ is not compact. The unit basis vectors $\delta_n$ are spaced by 2.

Examples:

1. Let $c_k \geq 1$ be a sequence that increases to infinity. The squashed solid rectangle of all $x$ with $c_k|x_k| \leq 1$ for all $k$ is compact in $\ell^\infty$.
2. Let $c_k \geq 1$ be a sequence that increases to infinity. The squashed solid ellipsoid of all $x$ with $\sum_{k=1}^\infty c_k x_k^2 \leq 1$ is compact in $\ell^2$.
3. Let $c_k \geq 1$ be a sequence that increases to infinity. The squashed region of all $x$ with $\sum_{k=1}^\infty c_k|x_k| \leq 1$ is compact in $\ell^1$.

## 18.3 Countable product spaces

Let $M_j$ for $j \in \mathbf{N}$ be a sequence of metric spaces. Let $\prod_j M_j$ be the product space consisting of all functions $f$ such that $f(j) \in M_j$. Let $\phi(t) = t/(1+t)$. Define the *product metric* by

$$d(f,g) = \sum_{j=1}^{\infty} \frac{1}{2^j} \phi(d(f(j), g(j))). \quad (18.1)$$

The following results are elementary.

**Lemma 18.6** *If each $M_j$ is complete, then $\prod_j M_j$ is complete.*

**Lemma 18.7** *If each $M_j$ is totally bounded, then $\prod_j M_j$ is totally bounded.*

**Theorem 18.8** *If each $M_j$ is compact, then $\prod_j M_j$ is compact.*

Examples:

1. The product space $\mathbb{R}^\infty$ is complete but not compact.

2. The closed unit ball (solid cube) in $\ell^\infty$ is a compact subset of $\mathbb{R}^\infty$ with respect to the $\mathbb{R}^\infty$ metric. In fact, it is a countably infinite product $[0,1]^\infty$ of compact spaces $[0,1]$. What makes this work is that the $\mathbb{R}^\infty$ metric measures the distances for various coordinates in increasingly less stringent ways. This example is called the *Hilbert cube*.

3. In the last example the Hilbert cube was defined as the countable infinite product $[0,1]^\infty$ of the unit interval $[0,1]$ with itself, with some metric uniformly equivalent to the $\mathbb{R}^\infty$ metric on this cube. An example of such a metric is $d(x,y) = \sqrt{\sum_n |x_k - y_k|^2 a_n^2}$, with a fixed sequence of $a_n > 0$ such that $\sum_n a_n^2 < +\infty$. A more geometric way of thinking of the Hilbert cube is as the countable product of the spaces $[0, a_n]$, regarded as a subspace of $\ell^2$. The metric on $\ell^2$ is natural geometrically. In this picture the Hilbert cube is compact because it is more and more compressed as the dimension gets bigger.

4. The unit sphere (cube faces) in $\ell^\infty$ is not compact with respect to the $\mathbb{R}^\infty$ metric, in fact, it is not even closed. The sequence $\delta_n$ converges to zero. The zero sequence is in the closed ball (solid cube), but not in the sphere.

## 18.4 The Bolzano-Weierstrass property

The notion of *subsequence* depends on the concept of increasing injective function from $\mathbf{N}$ to $\mathbf{N}$. Let $r : \mathbf{N} \to Nb$ be such a function. Then $r$ is characterized by the property that $m < n$ implies $r_m < r_n$. If $s : \mathbf{N} \to M$ is a sequence, then $s \circ r$ is a subsequence. That is, a subsequence of $n \mapsto s_n$ is $j \mapsto s_{r_j}$, where $r$ is increasing and injective.

**Theorem 18.9 (Bolzano-Weierstrass property)** *A metric space $M$ is compact if and only if every sequence with values in $M$ has a subsequence that converges to a point of $M$.*

Proof: Suppose that $M$ is compact. Thus it is totally bounded and complete. Let $s$ be a sequence with values in $M$. Since $M$ is bounded, it is contained in a ball of radius $C$.

By induction construct a sequence of balls $B_j$ of radius $C/2^j$ and a decreasing sequence of infinite subsets $N_j$ of the natural numbers such that for each $k$ in $N_j$ we have $s_k$ in $B_j$. For $j = 0$ this is no problem. If it has been accomplished for $j$, cover $B_j$ by finitely many balls of radius $C/2^{j+1}$. Since $N_j$ is infinite, there must be one of these balls such that $s_k$ is in it for infinitely many of the $k$ in $N_j$. This defines $B_{j+1}$ and $N_{j+1}$.

Let $r$ be a strictly increasing sequence of numbers such that $r_j$ is in $N_j$. Then $j \mapsto s_{r_j}$ is a subsequence that is a Cauchy sequence. By completeness it converges.

The converse proof is easy. The idea is to show that if the space is either not complete or not totally bounded, then there is a sequence without a convergent subsequence. In the case when the space is not complete, the idea is to have the sequence converge to a point in the completion. In the case when the space is not totally bounded, the idea is to have the terms in the sequence separated by a fixed distance. □

The theorem shows that for metric spaces the concept of compactness is invariant under topological equivalence. In fact, it will turn out that compactness is a purely topological property.

## 18.5 Compactness and continuous functions

**Theorem 18.10** *Let $K$ be a compact metric space. Let $L$ be another metric space. Let $f : K \to L$ be a continuous function. Then $f$ is uniformly continuous.*

Proof: Suppose $f$ were not uniformly continuous. Then there exists $\epsilon > 0$ such that for each $\delta > 0$ the set of pairs $(x, y)$ with $d(x, y) < \delta$ and $d(f(x), f(y)) \geq \epsilon$ is not empty. Consider the set of pairs $(x, y)$ with $d(x, y) < 1/n$ and $d(f(x), f(y)) \geq \epsilon$. Choose $s_n$ and $t_n$ with $d(s_n, t_n) < 1/n$ and $d(f(s_n), f(t_n)) \geq \epsilon$. Since $K$ is compact, there is a subsequence $u_k = s_{r_k}$ that converges to some limit $a$. Then also $v_k = t_{r_k}$ converges to $a$. But then $f(u_k) \to f(a)$ and $f(v_k) \to f(a)$ as $k \to \infty$. In particular, $d(f(u_k), f(v_k)) \to d(a, a) = 0$ as $k \to \infty$. This contradicts the fact that $d(f(u_k), f(v_k)) \geq \epsilon$. □

A corollary of this result is that for compact metric spaces the concepts of uniform equivalence and topological equivalence are the same.

**Theorem 18.11** *Let $K$ be a compact metric space. Let $L$ be another metric space. Let $f : K \to L$ be continuous. Then $f[K]$ is compact.*

Proof: Let $t$ be a sequence with values in $f[K]$. Choose $s_k$ with $f(s_k) = t_k$. Then there is a subsequence $u_j = s_{r_j}$ with $u_j \to a$ as $j \to \infty$. It follows that

$t_{r_j} = f(s_{r_j}) = f(u_j) \to f(a)$ as $j \to \infty$. This shows that $t$ has a convergence subsequence. □

The classic application of this theorem is to the case when $f : K \to \mathbb{R}$, where $K$ is a non-empty metric space. Then $f[K]$ is a non-empty compact subset of $\mathbb{R}$. However, a non-empty compact set of real numbers has a least element and a greatest element. Therefore there is a $p$ in $K$ where $f$ assumes its minimum value, and there is a $q$ in $K$ where $f$ assumes its maximum value.

## 18.6 The Heine-Borel property

It is striking that while completeness is a metric property, and total boundedness is a metric property, according to the theorem of this section compactness is a purely topological property. In fact, it can be formulated entirely in terms of open subsets, with no mention of the metric.

An *open cover* of a topological space $K$ is a collection $\Gamma$ of open sets with $K \subset \bigcup \Gamma$. The *Heine-Borel property* says that if $\Gamma$ is an open cover of $K$, then there is a finite subcollection $\Gamma_0$ that is an open cover of $K$. This is a purely topological property.

An equivalent statement of the Heine-Borel property is the *finite intersection property* following. Let $\Delta$ be a collection of closed subsets of $K$. Suppose that for each finite subcollection $\Delta_0$ of $\Delta$ the intersection $\bigcap \Delta_0 \neq \emptyset$. Then $\bigcap \Delta \neq \emptyset$.

**Theorem 18.12** *The metric space $K$ is compact if and only if $K$ has the Heine-Borel property.*

Proof: Suppose that the metric space $K$ is compact. Then it has the Bolzano-Weierstrass property, and in addition it is totally bounded. Let $\Gamma$ be an open cover of $K$. The main point of the following proof is to show that there is an $\epsilon > 0$ such that the sets in $\Gamma$ all overlap by at least $\epsilon$. More precisely, the claim is that there is an $\epsilon > 0$ such that for every $x$ there is an open set $U$ in $\Gamma$ such that $B(x, \epsilon) \subset U$.

Otherwise, there would be a sequence $\epsilon_n \to 0$ and a sequence $x_n$ such that $B(x_n, \epsilon_n)$ is not a subset of an open set in $\Gamma$. By the Bolzano-Weierstrass property there is a subsequence $x_{n_k}$ that converges to some $x$. Since $\Gamma$ is a cover, there is a $U$ in $\Gamma$ such that $x \in U$. Since $U$ is open, there is a $\delta > 0$ such that $B(x, \delta) \subset U$. Take $k$ so large that $d(x_{n_k}, x) < \delta/2$ and $\epsilon_{n_k} < \delta/2$. Then $d(y, x) \leq d(y, x_{n_k}) + d(x_{n_k}, x)$, so $B(x_{n_k}, \epsilon_{n_k}) \subset B(x, \delta) \subset U$. This is a contradiction. So there must be an overlap by some $\epsilon > 0$.

By total boundedness, there are points $x_1, \ldots, x_r$ such that the balls $B(x_j, \epsilon)$ cover $K$. For each $j$ let $U_j$ be an open set in $\Gamma$ such that $B(x_j, \epsilon) \subset U_j$. Then the $U_j$ form an finite cover of $K$. This completes the proof of the Heine-Borel property.

The converse is much easier. Suppose that $x_n$ form an infinite sequence of distinct points with no convergent subsequence. Then for each $x$ in $K$ there is a ball $B(x, \epsilon_x)$ with only finitely many elements of the sequence in it. The balls $B(x, \epsilon_x)$ cover $K$, but cannot have a finite subcover. □

## 18.7 Semicontinuity

A function from a metric space $M$ to $[-\infty, +\infty)$ is said to be *upper semicontinuous* if for every $u$ and every $r > f(u)$ there is a $\delta > 0$ such that all $v$ with $d(u, v) < \delta$ satisfy $f(v) < r$. An example of an upper semicontinuous function is one that is continuous except where it jumps up at a single point. It is easy to fall from this peak. The indicator function of a closed set is upper semicontinuous. The infimum of a non-empty collection of upper semicontinuous functions is upper semicontinuous. This generalizes the statement that the intersection of a collection of closed sets is closed.

There is a corresponding notion of lower semicontinuous function. A function from a metric space $M$ to $(-\infty, +\infty]$ is said to be *lower semicontinuous* if for every $u$ and every $r < f(u)$ there is a $\delta > 0$ such that all $v$ with $d(u, v) < \delta$ satisfy $f(v) > r$. An example of a lower semicontinuous function is one that is continuous except where jumps down at a single point. The indicator function of an open set is lower semicontinuous. The supremum of a non-empty collection of lower semicontinuous functions is lower semicontinuous. This generalizes the fact that the union of a collection of open sets is open.

**Theorem 18.13** *Let $K$ be compact and not empty. Let $f : K \to (-\infty, +\infty]$ be lower semicontinuous. Then there is a point $p$ in $K$ where $f$ assumes its minimum value.*

Proof: Let $a$ be the infimum of the range of $f$. Suppose that $s$ is a sequence of points in $K$ such that $f(s_n) \to a$. By compactness there is a strictly increasing sequence $g$ of natural numbers such that the subsequence $j \mapsto s_{g_j}$ converges to some $p$ in $K$. Consider $r < f(p)$. The lower semicontinuity implies that for sufficiently large $j$ the values $f(s_{g_j}) > r$. Hence $a \geq r$. Since $r < f(p)$ is arbitrary, we conclude that $a \geq f(p)$. □

There is a corresponding theorem for the maximum of an upper semicontinuous function on a compact space that is not empty.

## 18.8 Compact sets of continuous functions

Let $A$ be a family of functions on a metric space $M$ to another metric space. Then $A$ is *equicontinuous* if for every $x$ and every $\epsilon > 0$ there is a $\delta > 0$ such that for all $f$ in $A$ the condition $d(x, y) < \delta$ implies $d(f(x), f(y)) < \epsilon$. Thus the $\delta$ does not depend on the $f$ in $A$.

Similarly, $A$ is *uniformly equicontinuous* if for every $\epsilon > 0$ there is a $\delta > 0$ such that for all $f$ in $A$ the condition $d(x, y) < \delta$ implies $d(f(x), f(y)) < \epsilon$. Thus the $\delta$ does not depend on the $f$ in $A$ or on the point in the domain.

Finally, $A$ is *equiLipschitz* if there is a constant $C$ such that for all $f$ in $A$ the condition $d(x, y) < \delta$ implies $d(f(x), f(y)) < Cd(x, y)$ is satisfied.

It is clear that equiLipschitz implies uniformly equicontinuous implies equicontinuous.

**Lemma 18.14** *Let $K$ be a compact metric space. If $A$ is an equicontinuous set of functions on $K$, then $A$ is a uniformly equicontinuous set of functions on $K$.*

Let $K, M$ be metric spaces, and let $BC(K \to M)$ be the metric space of all bounded continuous functions from $K$ to $M$. The distance between two functions is given by the supremum over $K$ of the distance of their values in the $M$ metric. When $M$ is complete, this is a complete metric space. When $K$ is compact or $M$ is bounded, this is the same as the space $C(K \to M)$ of all continuous functions from $K$ to $M$. A common case is when $M = [-m, m] \subset \mathbb{R}$, a closed bounded interval of real numbers.

**Theorem 18.15 (Arzelà-Ascoli)** *Let $K$ and $M$ be totally bounded metric spaces. Let $A$ be a subset of $C(K \to M)$. If $A$ is uniformly equicontinuous, then $A$ is totally bounded.*

Proof: Let $\epsilon > 0$. By uniform equicontinuity there exists a $\delta > 0$ such that for all $f$ in $A$ and all $x, y$ the condition $d(x, y) < \delta$ implies that $|f(x) - f(y)| < \epsilon/4$. Furthermore, there is a finite set $F \subset K$ such that every point in $K$ is within $\delta$ of a point of $F$. Finally, there is a finite set $G$ of points in $M$ that are within $\epsilon/4$ of every point in $M$. The set $G^F$ is finite.

For each $h$ in $G^F$ let $D_h$ be the set of all $g$ in $A$ such that $g$ is within $\epsilon/4$ of $h$ on $F$. Every $g$ is in some $D_h$. Each $x$ in $K$ is within $\delta$ of some $a$ in $F$. Then for $g$ in $D_h$ we have

$$|g(x) - h(a)| \leq |g(x) - g(a)| + |g(a) - h(a)| < \epsilon/4 + \epsilon/4 = \epsilon/2. \qquad (18.2)$$

We conclude that each pair of functions in $D_h$ is within $\epsilon$ of each other. Thus $A$ is covered by finitely many sets of diameter $\epsilon$. □

In practice the way to prove that $A$ is uniformly equicontinuous is to prove that $A$ is equiLipschitz with constant $C$. Then the theorem shows in a rather explicit way that $A$ is totally bounded. In fact, the functions are parameterized to within a tolerance $\epsilon$ by functions from the finite set $F$ of points spaced by $\delta = \epsilon/(4C)$ to the finite set $G$ of points spaced by $\epsilon/4$.

**Corollary 18.16 (Arzelà-Ascoli)** *Let $K, M$ be compact metric spaces. Let $A$ be a subset of $C(K \to M)$. If $A$ is equicontinuous, then its closure $\bar{A}$ is compact.*

Proof: Since $K$ is compact, the condition that $A$ is equicontinuous implies that $A$ is uniformly equicontinuous. By the theorem, $A$ is totally bounded. It follows easily that the closure $\bar{A}$ is totally bounded. Since $M$ is compact and hence complete, $C(K \to M)$ is complete. Since $\bar{A}$ is a closed set of a complete space, it is also complete. The conclusion is that $\bar{A}$ is compact. □

The theorem has consequences for existence results. Thus every sequence of functions in $A$ has a subsequence that converges in the metric of $C(K \to M)$ to a function in the space.

## 18.9 Summary

This summary is a brief comparison of compactness and completeness.

A topological space $K$ is compact iff whenever Gamma is a collection of closed subsets with the property that every finite subcollection has non-empty intersection, then Gamma has non-empty intersection. Every closed subset $A$ of a compact space $K$ is compact. Every compact subset $K$ of a Hausdorff space $X$ is closed.

Compactness is preserved under topological equivalence. In fact the image of a compact space under a continuous map is compact.

A metric space $M$ is complete if every Cauchy sequence in $M$ converges to a point in $M$. A closed subset $A$ of a complete metric space $M$ is complete. A complete subset $A$ of a metric space $M$ is closed.

Completeness is preserved under uniform equivalence. In fact, the image of a complete space under a continuous uniformly open map is complete.

A compact metric space $K$ is complete. For compact metric spaces topological equivalence is the same as uniform equivalence. In fact, every continuous map from a compact metric space to another metric space is uniformly continuous.

# Problems

1. Let $c_k \geq 1$ be a sequence that increases to infinity. Show that the squashed solid ellipsoid of all $x$ with $\sum_{k=1}^{\infty} c_k x_k^2 \leq 1$ is compact in $\ell^2$.

2. Prove that the squashed solid ellipsoid in $\ell^2$ is not homeomorphic to the closed unit ball in $\ell^2$.

3. Let $c_k \geq 1$ be a sequence that increases to infinity. Is the squashed ellipsoid of all $x$ with $\sum_{k=1}^{\infty} c_k x_k^2 = 1$ compact in $\ell^2$?

4. Is the squashed ellipsoid in $\ell^2$ completely metrizable? Hint: Show that the set $\sum_k c_k x_k^2 \leq 1$ is a $G_\delta$ in $\ell^2$. Show that the set $\sum_k c_k x_k^2 \geq 1$ is a $G_\delta$ in $\ell^2$.

5. Consider a metric space $A$ with metric $d$. Say that there is another metric space $B$ with metric $d_1$. Suppose that $A \subset B$, and that $d_1 \leq d$ on $A \times A$. Finally, assume that there is a sequence $f_n$ in $A$ that approaches $h$ in $B \setminus A$ with respect to the $d_1$ metric. Show that $A$ is not compact with respect to the $d$ metric. (Example: Let $A$ be the unit sphere in $\ell^2$ with the $\ell^2$ metric, and let $B$ be the closed unit ball in $\ell^2$, but with the $\mathbb{R}^\infty$ metric.)

6. Is the metric space of continuous functions on $[0, 1]$ to $[-1, 1]$ with the sup norm compact? Prove or disprove. (Hint: Consider the completion with respect to the metric $d_1(f, g) = \int_0^1 |f(x) - g(x)|\, dx$. Construct a sequence as in the previous problem. Also, check directly that the sequence is not totally bounded.)

7. Consider the situation of the Arzelà-Ascoli theorem applied to a set $A \subset C(K)$ with bound $m$ and Lipschitz constant $C$. Suppose that the number of $\delta$ sets needed to cover $K$ grows like $(L/\delta)^k$, a finite dimensional behavior (polynomial in $1/\delta$). What is the growth of the number of $\epsilon$ sets needed to cover $A \subset C(K)$? It this a finite dimensional rate?

8. Let $M$ be a metric space. Show that a real function $f$ on $M$ is continuous if and only if its restriction to each compact subset $K$ is continuous. Hint: Use sequences.

9. Let $M$ be a metric space. Let $f_n$ be a sequence of continuous real functions on $M$ such that for each compact subset $K$ of $M$ we have $f_n \to f$ uniformly on $K$. Prove that $f$ is continuous.

# Part V

# Polish Spaces

# Chapter 19

# Completely metrizable topological spaces

## 19.1 Completely metrizable spaces

The central focus of this part is Polish spaces: topological spaces that are metrizable with a complete metric and that are separable. These include most of the spaces on which analysis is done. This leads to the subject of standard spaces: measurable spaces isomorphic to the space of Borel subsets of a Polish space. The concluding result of this chapter will center around a remarkable uniqueness result: up to isomorphism there is only one uncountable standard measurable space.

However, initially we concentrate on a more general class of topological spaces, those that are metrizable with a complete metric. Such a space is said to be *completely metrizable*. Sometimes such a space is called *topologically complete*. However this term is used more general topological contexts, so a reader consulting other references must remain alert. If necessary use a term such as "metrically topologically complete".

If the hypothesis of a theorem says that a certain space is a complete metric space, and the conclusion of the theorem is purely topological property of this space, then it is clear that the conclusion follows for an arbitrary completely metrizable topological spaces. Often this observation is taken for granted. On the other hand, if the conclusion of a theorem says that a certain metric space is a completely metrizable topological space, then it does not follow that it is a complete metric space.

Consider a topological space. A subset is a $G_\delta$ if it is a countable intersection of open sets. A subset is a $F_\sigma$ if it is a countable union of closed sets. The German origin of $G_\delta$ is Gebiet-Durchschnitt, which means open intersection. The French origin of $F_\sigma$ is fermé-somme, which means closed union.

**Theorem 19.1** *Let $T$ be a completely metrizable topological space. Suppose* *that $S$ is dense in $T$. Then $S$ is completely metrizable if and only if it is a $G_\delta$ in $T$.*

To motivate the if part of the proof, look at the special case when $S$ is open and dense in $T$. Let $B = T \setminus S$. Suppose that $T$ has the metric $d$. Since $S$ is open it follows that $d(x, B) > 0$ for each $x$ in $S$. Consider the metric $e$ defined on $S$ by

$$e(x, y) = d(x, y) + \left| \frac{1}{d(x, B)} - \frac{1}{d(y, B)} \right|. \tag{19.1}$$

It is easy to see that $d$ and $e$ define the same topology on $S$. However with the metric $e$ the space $S$ is complete. To see this, consider a Cauchy sequence $x_m$ with respect to the metric $e$. It is also a Cauchy sequence with respect to the metric $d$, so it converges to a point $x$ in $T$. On the other hand, the numbers $1/d(x_m, B)$ also form a Cauchy sequence, so they converge to a number $a \neq 0$. Thus $1/d(x, B) = a$, and so $x$ is in $S$.

Proof: Suppose that $S$ is a $G_\delta$ in $T$. Then $S = \bigcap_n U_n$, where $U_n$ is open in $T$. Let $B_n = T \setminus U_n$. Suppose that $T$ has the metric $d$. Since $S \subset U_n$ and $U_n$ is open it follows that $d(x, B_n) > 0$ for each $x$ in $S$. Let $b(t) = t/(1 + t)$. This is the transformation that creates bounded metrics. Consider the metric $e$ defined on $S$ by

$$e(x, y) = d(x, y) + \sum_n \frac{1}{2^n} \left| \frac{1}{d(x, B_n)} - \frac{1}{d(y, B_n)} \right|. \tag{19.2}$$

It is easy to see that $d$ and $e$ define the same topology on $S$. However with the metric $e$ the space $S$ is complete. To see this, consider a Cauchy sequence $x_m$ with respect to the metric $e$. It is also a Cauchy sequence with respect to the metric $d$, so it converges to a point $x$ in $T$. On the other hand, for each $n$ the numbers $1/d(x_m, B_n)$ also form a Cauchy sequence, so they converge to a number $a_n \neq 0$. Thus $1/d(x, B_n) = a_n$, and so $x$ is in the complement of $B_n$. Since this works for each $n$, it follows that $x$ is in $S$. Thus $S$ with this metric is complete.

For the converse, suppose that $S$ has a metric $e$ that defines its topology and that with this metric $S$ is complete. For each $n$ let

$$U_n = \{x \in T \mid \exists \delta > 0 \, \forall y \in S \forall z \in S \, (d(y, x) < \delta, d(z, x) < \delta \Rightarrow e(y, z) < 1/n)\}. \tag{19.3}$$

The first thing to note is that for each $n$ we have $S \subset U_n$. This is because on $S$ the metrics $d$ and $e$ define the same topology. So for each $n$ there exists $\delta > 0$ so that if $y$ and $z$ are each within $\delta$ of $x$ with respect to $d$, then $y$ and $z$ are each within $1/(2n)$ of $x$ with respect to $e$. It follows that $y$ and $z$ are within $1/n$ of each other with respect to $e$.

Next, each set $U_n$ is open. Suppose that $x$ is in $U_n$. Then there is a corresponding $\delta$. Suppose that $x'$ is within $\delta/2$ of $x$ with respect to $d$. Then if $y$ and $z$ are within $\delta/2$ of $x'$ with respect to $d$, then they are within $\delta$ of $x$ with respect to each $d$. This is enough to show that $x'$ is in $U_n$.

Finally, the intersection of the $U_n$ is $S$. To see this, suppose that $x$ is a point such that for each $n$ we have $x$ in $U_n$. Since $S$ is dense in $T$, there is a sequence $m \mapsto x_m$ of points in $S$ such that $x_m \to x$ as $m \to \infty$. Consider a particular value of $n$. Since $x$ is in $U_n$, there is a value of $\delta$ so that $d(x_m, x) < \delta$ and $d(x_k, x) < \delta$ implies $e(x_m, x_k) < 1/n$. However for $m, k$ large enough we can guarantee that $d(x_m, x) < \delta$ and $d(x_k, x) < \delta$. So for these $m, k$ we have $e(x_m, x_k) < 1/n$. Thus the $x_m$ form a Cauchy sequence with respect to $e$. Therefore the $x_m$ converge to some limit in $S$. This limit must be $x$, so $x$ is in $S$. The conclusion of the discussion is that $S$ is the intersection of the open sets $U_n$, and hence $S$ is a $G_\delta$.
□

## 19.2 Locally compact metrizable spaces

A topological space is $M$ is *compact* if it has the Heine-Borel property, that is, every open cover has a finite subcover. A topological space is $M$ is *locally compact* if for each $p$ in $M$ there is an open subset $U$ and a compact subset $K$ with $p \in U \subset K$.

**Theorem 19.2** *Let $X$ be a metrizable topological space. Let $M$ be a subset of $X$ such that $M$ is locally compact and $M$ is dense in $X$. Then $M$ is an open subset of $X$.*

Proof: Suppose that $p$ is in $M$. Since $M$ is locally compact, there is an open subset $U$ of $M$ and a compact subset $K$ of $M$ such that $p \in U \subset K$. Furthermore, there is an open subset $W$ of $X$ such that $W \cap M = U$. Since $\bar{M} = X$ it follows that $W = W \cap \bar{M} \subset \overline{W \cap M} \subset \bar{U}$. Since $K$ is compact in $M$, it follows that $K$ is compact in $X$, and consequently $K$ is closed in $X$. It follows that $\bar{U} \subset K$. This shows that $W$ is open in $X$ with $p \in W \subset K \subset M$. This suffices to prove that $M$ is an open subset of $K$. □

**Corollary 19.3** *Let $M$ be a locally compact metrizable space. Then $M$ is a completely metrizable topological space.*

Proof: Let $X$ be the completion of $M$. Then $M$ is dense in $M$. It follows from the theorem that $M$ is an open subset of $X$. In particular, $M$ is a $G_\delta$ in $X$. It follows from an earlier theorem that $M$ is topologically complete. □

## 19.3 Closure and interior

For each subset $A$ of a topological space, its *closure* is the smallest closed set of which $A$ is a subset. The closure of $A$ is denoted $\bar{A}$. The closure operation satisfies the *Kuratowski closure axioms*:

1. $\bar{\emptyset} = \emptyset$.

2. $A \subset \bar{A}$.

3. $\overline{\bar{A}} = A$.

4. $\overline{(A \cup B)} = \bar{A} \cup \bar{B}$.

For each subset $A$ of a topological space $X$, its *interior* is the largest open subset of $A$. The interior of $A$ is denoted $A^\circ$. The relation between closure and interior is that $(\bar{A})^c = A^{c\circ}$. In other words, just as complementation interchanges closed sets and open sets, it also interchanges closure and interior operations.

A dense set $B$ with $\bar{B} = X$ is good at approximation. On the other hand, an empty interior set $A$ with $A^\circ = \emptyset$, that is, $\overline{A^c} = X$, is a set whose complement is good at approximation. Every point is near a point that is not in $A$.

The operation $A^{c\circ}$ sends $A$ into the interior of its complement, that is, into the set of all points that are isolated from $A$. This operation may be iterated. The operation $A^{c\circ c\circ} = (\bar{A})^\circ$ sends $A$ into the set of all points that are isolated from the points that are isolated from $A$. In other word, every point near a point in $(\bar{A})^\circ$ is approximated by points in $A$.

A set $B$ has *dense interior* if $\overline{B^\circ} = X$. For a set $B$ with dense interior every point $x$ in $X$ may be approximated by points in the interior of $B$. Thus such a set is extremely good at approximation.

A set $A$ such that $(\bar{A})^\circ = \emptyset$ is said to be *nowhere dense*. This is the dual notion to dense interior. A nowhere dense set $A$ has a complement that is extremely good in approximation, since $\overline{A^{c\circ}} = X$. In other words, every point is near a point not approximated by $A$.

The collection of nowhere dense subsets is an *ideal* of subsets in the collection of all subsets. This means that the nowhere dense subsets are closed under finite intersections and countable unions, and furthermore, if $A$ is nowhere dense and $B$ is an arbitrary subset, then $A \cap B$ is nowhere dense.

Notice for future use that when a set $A$ fails to be nowhere dense, that means that there exists a point $x$ that is in the interior of $\bar{A}$.

If $A$ is a subset of a topological space $X$, and $A^c = X \setminus A$ is its complement, then the *boundary* of $A$ is the closed subset $\partial A = \bar{A} \setminus A^\circ$. If $A$ is open, or if $A$ is closed, then $\partial A$ has empty interior, and so in particular $\partial A$ is nowhere dense.

## 19.4 The Baire category theorem

The Baire category theorem is a theory of subsets of a completely metrizable topological space. Some subsets are good at approximation (residual sets) and other subsets have complements that are good at approximation (meager sets). Another name for meager set is set of first category; hence the terminology.

Let $S$ be a topological space. A subset $M$ is *meager* if $M$ is a countable union of nowhere dense subsets. A subset $R$ of $S$ is *residual* if $R$ is a countable intersection of dense interior subsets. The property of being meager or residual is a purely topological property.

For an example, take $S = [0, 1]$. Every finite subset is nowhere dense. Every countable subset is meager. The indicator function of the rationals fails to be nowhere dense, but it is meager. On the other hand, an uncountable subset of $[0, 1]$ can be nowhere dense. The Cantor set is an example.

The collection of meager subsets is a *$\sigma$-ideal* of subsets in the collection of all subsets. This means that the meager subsets are closed under finite intersections and countable unions, and furthermore, if $A$ is meager and $B$ is an arbitrary subset, then $A \cap B$ is meager. A *non-meager subset* is, of course, a subset that is not meager. The Baire theorem proved below will establish that every subset with non-empty interior of a completely metrizable topological space is non-meager.

Sometimes another terminology is used. A subset $M$ is *first category* if it is meager. A non-meager subset is *second category.* We shall not use this terminology.

**Theorem 19.4 (Baire category theorem)** *Let $S$ be a non-empty completely metrizable topological space. If $R$ is residual in $S$, then $R$ is dense in $S$. Equivalently, if $M$ is meager in $S$, then $M$ has empty interior.*

Proof: It is sufficient to prove that if $G_n$ is a sequence of sets each with dense interior, then the intersection $R$ is dense.

Let $x_1$ be an arbitrary point in $S$. Let $\epsilon_1$ be an arbitrary number with $0 < \epsilon_1 < 1$. The task is to prove that there is an $x$ in $R$ such that $d(x, x_1) < \epsilon_1$.

Construct inductively $x_n$ and $0 < \epsilon_n$ with $\overline{B(x_{n+1}, \epsilon_{n+1})} \subset B(x_n, \epsilon_n/2) \cap G_n^\circ$. In particular $d(x_{n+1}, x_n) < \epsilon_n/2$ and $\epsilon_{n+1} \leq \epsilon_n/2$. The reason this can be done is that the interior $G_n^\circ$ is dense, and so it must have non-empty intersection with the open ball $B(x_n, \epsilon_n/2)$. Then $d(x_m, x_n) < \epsilon_n \sum_{k=1}^{\infty} 1/2^k = \epsilon_n$ for $m \geq n$. This proves that the $x_n$ form a Cauchy sequence. Since $S$ is a complete metric space, there is an $x$ such that $x_n \to x$ as $n \to \infty$. Then $d(x, x_n) \leq \epsilon_n$. Furthermore, $d(x_{n+1}, x) \leq \epsilon_{n+1}$ implies $x \in G_n^\circ$. Thus there is an element $x$ in $R$ with $d(x, x_1) < \epsilon_1$. □

Meager sets are sets whose complements are very good at approximation, yet they have a stability property under countable unions. The basic properties of meager subsets of completely metrizable topological spaces are summarized here:

- $A$ nowhere dense $(\bar{A})^\circ = \emptyset \Rightarrow A$ meager $\Rightarrow A$ has empty interior ($A^\circ = \emptyset$).
- The class of meager subsets is closed under countable unions.
- Consequence: A countable union of closed sets with empty interiors has empty interior.

Residual sets are very good at approximation, yet they have a stability property under countable intersections. The basis properties of residual subsets of completely metrizable topological spaces are summarized here:

1. $B$ dense interior ($\overline{B^\circ} = X$) $\Rightarrow B$ residual $\Rightarrow B$ dense ($\bar{B} = X$).

2. The class of residual subsets is closed under countable intersection.

3. A countable intersection of dense open sets is dense.

**Theorem 19.5** *A subset $R$ of a completely metrizable topological space $S$ is residual if and only if there is a subset $W \subset R$ that is a dense $G_\delta$ in $S$.*

Proof: If there is a subset $W$ that is a dense $G_\delta$, then $W$ is a countable intersection of open sets $U_n$. Furthermore, if $W$ is dense, then each $U_n$ is dense. It follows that $W$ is residual, and so $R$ is residual.

If $R$ is residual, then $R$ it is the intersection of sets $B_n$ with $\overline{B_n^\circ} = S$. Let $W$ be the intersection of the sets $B_n^\circ$. Then $W$ is a $G_\delta$. Since $W$ is also residual, it follows from the Baire theorem that $W$ is dense. □

Example: A trivial but illuminating illustration of Baire category ideas is the fact that the plane cannot be written as a countable union of lines. Each line is nowhere dense, but the plane is not meager. Notice however that the countable union of lines can be dense in the plane.

There is also a completely elementary proof of the same fact. Consider a fixed circle. Then each line intersects the circle in at most two points, so the union of lines cannot even include the circle.

There are many less trivial examples of the use of the Baire category theorem. What follows is a series of arguments that proves the remarkable fact that a lower semicontinuous real function cannot be discontinuous at every point.

**Theorem 19.6** *Suppose that $f$ is LSC on a complete metric space with real values. Then there exists a non-empty open subset on which $f$ is bounded above.*

Proof: Let $S_n$ be the subset where $f \leq n$. Since $f$ is LSC, the set $S_n$ is closed. Since $f$ has real values, the union of the $S_n$ is a complete metric space. Such a space cannot be meager. Therefore one of the $S_n$ must fail to be nowhere dense. Since $S_n$ is closed, it follows that $S_n$ has non-empty interior. This gives a non-empty open set on which $f \leq n$. □

**Lemma 19.7** *Suppose that $f$ is defined on a non-empty complete metric space and has real values. Suppose that $f$ is LSC but fails to be USC at every point. Then $f$ is unbounded above.*

Proof: To say that $f$ is USC at $x$ is to say that for every $k$ there exists an open set with $x$ in it such that for all $y$ in this open set the values $f(y) < f(x) + 1/k$. To say that $f$ is not USC at $x$ is to say that for some $k$ there are points $y$ arbitrarily close to $x$ such that $f(y) \geq f(x) + 1/k$. Let $A_k$ be the set of all $x$ such that there are $y$ arbitrarily close to $x$ with $f(x) \geq f(x) + 1/k$. If $f$ fails to be USC at every point, then the union of the $A_k$ is a complete metric space. It follows that one of the $A_k$ must fail to be nowhere dense. So there is a non-empty open set $U \subset \bar{A}_k$.

There exists $x_0$ in $U$ and in $A_k$. This will be the starting point for an inductive construction. Suppose that we have $x_i$ in $U$ and in $A_k$. Then there exists $y$ in $U$ with $f(y) \geq f(x_i)+1/k$. There exists a sequence of points that are in $U$ and in $A_k$ and converge to $y$. Since $f$ is LSC at $y$, there must be a point $x_{i+1}$ that is in the range of this sequence such that $f(x_{i+1}) > f(y) - 1/(2k)$. Thus $f(x_{i+1}) > f(x_i) + 1/(2k)$. It is clear that $f(x_i) > f(x_0) + i/(2k)$. This is enough to show that $f$ is unbounded above. □

**Theorem 19.8** *There is no real function on a non-empty complete metric space that is LSC and nowhere continuous.*

Proof: Suppose that $f$ were such a function. Since $f$ is LSC, there is an non-empty open set on which $f$ is bounded above. This open set is a topologically complete metric space. Since $f$ is nowhere USC on this space, it is unbounded above on it. This is a contradiction. Thanks to Leonid Friedlander for this argument. □

# Problems

1. Let $M$ be a metric space. Let $f$ be a real function on $M$. Show that the points of $M$ where $f$ is continuous form a $G_\delta$ subset. Hint: First prove that $f$ is continuous at $a$ if and only if for every $n$ there exists real $z$ and $\delta > 0$ such that for all $x$ $(d(x,a) < \delta \Rightarrow d(f(x), z) < 1/n)$. Let $U_n$ be the set of all $a$ for which there exists real $z$ and $\delta > 0$ such that for all $x$ $(d(x,a) < \delta \Rightarrow d(f(x), z) < 1/n)$. Show that $U_n$ is open.

2. (a) Consider a complete metric space. Show that every dense $G_\delta$ subset is residual.
   (b) Show that $\mathbb{Q}$ is not a residual subset of $\mathbb{R}$.
   (c) Show that is impossible for a real function on $\mathbb{R}$ to be continuous precisely on $\mathbb{Q}$.

3. Show that $\mathbb{R}^2$ is not a countable union of circles.

4. Let $M$ be a complete metric space. Let $f_n$ be a sequence of continuous real functions such that $f_n$ to $f$ pointwise. Show that there is a $k$ in $\mathbb{N}$ and a non-empty open subset $U$ such that $|f|$ is bounded by $k$ on $U$. Hint: Let $F_k$ be the set of all $x$ such that for each $n$ $|f_n(x)| \leq k$. Use the fact that $M$ is not meager.

5. Let $L$ be the subset of $C([0,1])$ consisting of Lipschitz functions. Show that $L$ is a meager subset. Hint: Let $F_k$ be the set of all $f$ such that for all $x, y$ we have $|f(x) - f(y)| \leq k|x - y|$.

6. Show that the middle third Cantor subset of $\mathbb{R}$ is equal to its boundary. Show that it is an uncountable nowhere dense subset of $\mathbb{R}$.

7. Recall that $f : X \to \mathbb{R}$ is lower semicontinuous (LSC) if and only if the inverse image of each interval $(a, +\infty)$, where $-\infty \leq a \leq +\infty$, is open

in $X$. Show that if $f_n \uparrow f$ pointwise and each $f_n$ is LSC, then $f$ is LSC. (This holds in particular if each $f_n$ is continuous.)

8. Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ that is LSC but is discontinuous at almost every point (with respect to Lebesgue measure).

# Chapter 20

# Polish topological spaces

## 20.1 The role of Polish spaces

Recall that a completely metrizable topological space is a space whose topology is given by some complete metric. A separable completely metrizable topological space is called a *Polish space*. Every compact metrizable space is a Polish space.

It may be shown [10] that every compact Hausdorff space is normal, and every locally compact Hausdorff space is regular. The Urysohn metrization theorem says that every second countable regular space is metrizable. It follows that every second countable locally compact Hausdorff space is metrizable. We have already seen that a locally compact metrizable space is completely metrizable. This proves the following theorem.

**Theorem 20.1** *Every second countable locally compact Hausdorff space is a Polish space.*

Among topological spaces the Polish spaces are particularly nice for analysis. The separable locally compact metrizable spaces are a more restrictive class; they are the same as the locally compact Polish spaces. The nicest of all are the compact metrizable spaces, which are the compact Polish spaces.

The Euclidean space $\mathbb{R}^n$ is a locally compact Polish space. So one could think of the concept of locally compact Polish space as capturing the idea of the topology of Euclidean space. However, some analysts prefer to work with the more general concept of locally compact Hausdorff space.

An important example of a Polish space that is not locally compact is the real Hilbert space $\ell^2$ of square-summable real sequences. This is the infinite dimensional analog of Euclidean space $\mathbb{R}^n$. Similarly, the spaces $\ell^1$ and $\ell^\infty$ are not locally compact. Perhaps surprisingly, the space $\mathbf{R}^\infty$ is also not locally compact. Certainly, each parallelepiped $\prod_n [-L_n, L_n]$ bounded in each coordinate direction is compact, but non-empty open subsets are only bounded in finitely many coordinate directions.

## 20.2 Embedding a Cantor space

A standard example of an compact metric space is the *Cantor product space* $2^{\mathbf{N}}$. This is space is uncountable. However in a certain sense it is the smallest such space, as is shown by the following result.

**Proposition 20.2** *Let $B$ be a Polish space. Suppose that $B$ is uncountable. Then there exists a compact subset $D \subset B$ that is homeomorphic to $2^{\mathbf{N}}$.*

Proof: Let $C$ be the set of all $y$ in $B$ such that there exists a countable open subset $U$ with $y \in U$. Since $B$ is separable, we can take the open sets to belong to a countable base. This implies that $C$ is itself a countable set.

Let $D = B \setminus C$. Since $B$ is uncountable, it follows that $D$ is also uncountable. Each $y$ in $D$ has the property that each open set $U$ with $y \in U$ is uncountable.

The next task is to construct for each sequence $\omega_1, \ldots, \omega_m$ a corresponding closed subset $F_{\omega_1,\ldots,\omega_m}$ of $D$. Each closed set has a non-empty interior and has diameter at most $1/m$. For different sequences the corresponding closed subsets are disjoint. The closed sets decrease as the sequence is extended. This is done inductively. Start with $D$. Since it is uncountable, it has at least two points. Construct the first two sets $F_0$ and $F_1$ to satisfy the desired properties. Say that the closed subset $F_{\omega_1,\ldots,\omega_m}$ has been defined. Since it non-empty interior, there are uncountably many points in it. Take two points. About each of these points construct subsets $F_{\omega_1,\ldots,\omega_m,0}$ and $F_{\omega_1,\ldots,\omega_m,1}$ with the desired properties.

These closed subsets of $D$ are also non-empty subsets of the complete metric space $B$ with decreasing diameter. By the completeness of $B$, for each infinite sequence $\omega$ in $2^{\mathbf{N}}$ the intersection of the corresponding sequence of closed sets has a single element $g(\omega) \in B$. It is not hard to see that $g : 2^{\mathbf{N}} \to B$ is a continuous injection. Since $2^{\mathbf{N}}$ is compact, it is a homeomorphism onto a compact subset $D$ of $B$. □

## 20.3 Embedding in the Hilbert cube

Another standard example of an compact metric space is the Hilbert cube, the product space $[0,1]^{\mathbf{N}}$. In a certain sense it is the largest such space, as is shown by the following result.

**Theorem 20.3** *Let $B$ be a separable metrizable space. Then there exists a compact subset $T \subset [0,1]^{\mathbf{N}}$ and a dense subset $S$ of $T$ such that $B$ is homeomorphic to $S$.*

Proof: Let $B$ be a complete separable metric space. There exists a metric on $B$ with values bounded by one such that the map between the two metric spaces is uniformly continuous. So we may as well assume that that $B$ is a complete separable metric space with metric $d$ bounded by one.

Since $B$ is separable, there is a sequence $s : \mathbf{N} \to B$ that is an injection with dense range. Let $I = [0,1]$ be the unit interval, and consider the space $I^{\mathbf{N}}$ with the product metric $d_p$. Define a map $f : B \to I^{\mathbf{N}}$ by $f(x)_n = d(x, s_n)$. This is a homeomorphism $f$ from $B$ to a subset $S \subset I^{\mathbf{N}}$. □

If a topological space $B$ is homeomorphic to a dense subspace $S$ of a compact space $T$, then $T$ is said to be a *compactification* of $B$. The above result shows that every separable metrizable space $B$ has a Polish compactification $T$. In view of the construction, it seems reasonable to call this the *Hilbert cube compactification* of a separable metrizable space.

**Corollary 20.4** *Let $B$ be a Polish space. Then there exists a compact subset $T \subset [0,1]^{\mathbf{N}}$ and a dense subset $S$ of $T$ that is a $G_\delta$ in $T$ such that $B$ is homeomorphic to $S$.*

# Problems

1. Describe the Hilbert cube compactification of the open interval $(0, 1)$. How many extra points are adjoined?

# Chapter 21

# Standard measurable spaces

## 21.1 Measurable spaces

As we know, a measurable space is a set $X$ together with a given $\sigma$-algebra of subsets. Many concepts for topological spaces carry over to measurable spaces. If $Z$ is a subset of $X$, then there is a relative $\sigma$-algebra induced on $X$, so that in this way $Z$ becomes a measurable space. Also, if $\Gamma$ is a partition of $X$, then there is a quotient $\sigma$-algebra induced on $\Gamma$, so that again $\Gamma$ becomes a measurable space.

Of course if $X$ is a topological space, then with its Borel $\sigma$-algebra it also becomes a measurable space. Such a measurable space, where the $\sigma$-algebra is generated by a topology, is sometimes called a Borel space.

It is not hard to see that if $X$ is a topological space and $Z$ is a subset, then the Borel measurable structure on $Z$ coming from the relative topology of $Z$ is the same as the relative measurable structure on $Z$ coming from the Borel measurable structure on $X$.

The corresponding result for quotient spaces is false. If $X$ is a topological space, and $\Gamma$ is a partition of $X$, then the Borel measurable structure on $\Gamma$ coming from the quotient topology of $\Gamma$ may be coarser than the quotient measurable structure on $\Gamma$ coming from the Borel measurable structure on $X$. Example: Let $X = \mathbb{R}$ and let the partition $\Gamma$ of $X$ consist of the intervals $[n, n+1)$ for integer $n$ in $\mathbb{Z}$. Thus the partition looks like $\mathbb{Z}$. The quotient topology on $\mathbb{Z}$ is the trivial topology with just the empty set and the whole space as open sets. Thus topology does not seem very useful for classification. The Borel measurable structure generated by this topology is also trivial. However the other direction gives us what we need. The measurable structure on the quotient space that comes from the Borel measurable structure on $\mathbb{R}$ consists of all subsets.

The general picture is that the topology on a quotient space may be too coarse to be of interest, but the measurable structure on a quotient space may be exactly what is appropriate. So even if a measurable structure is relatively uninformative compared to a topological structure, it may be all that is available for classification.

## 21.2 Bernstein's theorem for measurable spaces

This section gives the proof of Bernstein's theorem in the case when sets and subsets are replaced by measure spaces and measurable subsets. It is the same dynamical systems argument that gave the theorem for sets.

**Lemma 21.1** *Suppose that $C$ is a measurable space space. Suppose that $A \subset B \subset C$ are measurable subsets. Suppose that the map $\phi : C \to A$ is an isomorphism of measurable spaces. Then there exists a map $\psi : C \to B$ that is an isomorphism of measurable spaces.*

Proof: Think of $\phi$ as a dynamical system on $C$. Let $D = C \setminus A$. This is the part of the space that consists of starting points for the action of $\phi$ as a shift. It is a measurable subset. Since $\phi : C \to A$ is an isomorphism, it maps measurable subsets to measurable subsets. It follows that each iterate $\phi^n$ maps measurable subsets to measurable subsets. The part of the space on which $\phi$ acts as a shift is the countable union of measurable subsets $\phi^n[D]$ for $n = 0, 1, 2, 3, \ldots$, and hence it is a measurable subset $O(D)$. Let $E$ be the complement of $O(D)$. The set $E$ consists of the part $C$ on which $\phi$ is a bijection. The points in the measurable set $E$ consist of the intersection of all the $\phi^n[A]$ for $n = 0, 1, 2, 3, \ldots$. That is, each point in $E$ comes from an arbitrarily remote past.

Decompose $D$ into the two measurable subsets $F = C \setminus B$ and $G = B \setminus A$. Then $O(D)$ is the union of $O(F)$ with $O(G)$, where $O(D)$ is the union of the $\phi^n[F]$ and $O(G)$ is the union of the $\phi^n[G]$. These are all measurable subsets. Let $\psi : C \to B$ agree with with $\phi$ on $O(F)$, and let $\psi$ be the identity on $0(G)$. On $E$ on can either make $\psi$ agree with $\phi$, or it can be set to be the identity. Then the starting points for $\phi$ as a shift are $F$. The range of $\psi$ is the union of $O(F) \setminus F$ with $0(G)$ and with $E$. This is just $C \setminus F = B$. Thus $\psi$ gives a measurable isomorphism of $C$ with $B$. □

**Theorem 21.2** *Let $X$ and $Y$ be measurable spaces. Let $f : X \to Y$ and $g : Y \to X$ be measurable functions with images that are measurable subsets. Suppose that $f$ and $g$ are each isomorphisms onto their images. Then there is an isomorphism of measure spaces $h : X \to Y$.*

Proof: The composition $\phi = g \circ f : X \to X$ maps $X$ to itself. Since $g$ is an isomorphism from $Y$ to the image of $g$ in $X$, the map $g$ sends the image of $f$ to a measurable subset of $X$. This measurable subset is the image of $\phi$. Also $\phi$ is an isomorphism onto its image. From the lemma, there is an isomorphism $\psi$ from $X$ to the image of $g$. Then $h = g^{-1} \circ \psi$ is the desired isomorphism from $X$ to $Y$. □

## 21.3 A unique measurable structure

In the next sections we argue that the setting where $X$ is a Polish the Borel measurable structure is very close to being unique. The only possibilities are those associated with a countable set or with the unit interval $[0, 1]$.

Two measurable spaces are isomorphic if there is a bijection between the underlying sets that preserves the measurable subsets. If the spaces are topological spaces, and the measurable subsets are the Borel subsets, then the spaces are sometimes said to be Borel isomorphic.

Examples:

1. The measurable spaces $(0, 1)$ and $[0, 1)$ are Borel isomorphic. These spaces are homeomorphic to $(0, \infty)$ and $[0, +\infty)$. So to prove this, it is sufficient to show that $(0, +\infty)$ and $[0, +\infty)$ are Borel isomorphic. An isomorphism is given by $f(x) = n + 1 - (x - n)$ on $n < x \leq n + 1$. This is a Borel isomorphism, but it is far from being continuous.

2. The measurable space $[0, 1)$ and $[0, 1]$ are Borel isomorphic. It is obvious that $[0, 1)$ is homeomorphic to $(0, 1]$, so it is sufficient to show that $(0, 1]$ and $[0, 1]$ are Borel isomorphic. However $(0, 1)$ is a Borel subset of $(0, 1]$, and $[0, 1)$ is a Borel subset of $[0, 1]$. So we can take the isomorphism we got in the previous example between these subsets, and send 1 to 1.

**Theorem 21.3** *Let $X$ and $Y$ be two uncountable Polish topological spaces. Then $X$ and $Y$ are isomorphic as measurable space.*

This theorem implies that for every uncountable Polish space the associated measurable space is isomorphic to the measurable space associated with the unit interval $[0, 1]$. In other words, for most practical purposes, there is just one measurable space of interest.

There is a more general form of theorem that applies to uncountable separable metrizable spaces that Borel subsets of Polish spaces. See the text by Dudley [4] for this stronger version. The slightly more elementary Polish space version is proved in the following two sections of this chapter.

Here are some consequences. Say that a measurable space is a *standard measurable space* if it isomorphic, as a measurable space, to a Polish space with the Borel $\sigma$-algebra. Often this is called a standard Borel space.

Examples:

1. Both $[0, 1]$ and $(-\infty, +\infty)$ are standard measurable space, since they are separable complete metric spaces and hence Polish spaces.

2. $(0, 1)$ and $[0, 1)$ are standard measurable spaces, and in fact are Borel isomorphic to $[0, 1]$. It is true that these are not complete spaces with their usual metrics, but they are are still Polish spaces.

**Corollary 21.4** *Every standard measurable space is isomorphic, as a measurable space, to a countable set with the discrete $\sigma$-algebra or to the unit interval $[0,1]$ with the usual Borel $\sigma$-algebra.*

Proof: If the space is uncountable, then it is isomorphic as a measurable space to $[0,1]$ with its Borel structure, since $[0,1]$ is a complete separable metric space. If the space is countable, then every subset is a Borel set. So it is isomorphic to a finite set or to $\mathbf{N}$ with the discrete topology. □

**Corollary 21.5** *Every standard measurable space is isomorphic, as a measurable space, to the Borel structure associated with a compact metric space.*

Proof: The unit interval $[0,1]$ is compact. Every finite set is compact. The remaining case is that of a countable infinite set. Take the space to be $\mathbf{N} \cup \{\infty\}$, the one point compactification of the natural numbers. This is a countable compact metric space. Every subset is a Borel set. □

## 21.4 Measurable equivalence of Cantor space and Hilbert cube

The strategy of the proof is simple. The first part is to show that for every uncountable Polish space has a Cantor set embedded in it. The second part is to show that every Polish space may be placed inside a Hilbert cube. The third part is to place the Hilbert cube inside the Cantor set. The final step is to argue that if the space has a Cantor set inside and is also inside a Cantor set, then it may be matched up with a Cantor set. This is done by a dynamical systems argument.

The first part is a known result. Suppose that $X$ is an uncountable Polish space. We have seen that there is a subset $K$ of $X$ that is homeomorphic to the Cantor set.

The second part depends on the known embedding theorem. Suppose $X$ is a Polish space. We have also seen that there subset $S$ of the Hilbert cube with closure $T$ such that $B$ is homeomorphic to $S$, and $S$ is a $G_\delta$ subset of $T$.

By the definition of the relative topology on $T$, Each open subset $U$ of $T$ is of the form $U = T \cap V$, where $V$ is open an open subset of the Hilbert cube. Since $S$ is a $G_\delta$, it is of the form $S = \bigcap_n U_n$, where each $U_n = T \cap V_n$ is open in $T$. Thus $S = T \cap \bigcap_n V_n$. Since $T$ is closed, it is a Borel subset of the Hilbert cube. Each $V_n$ is a open set, so $\bigcap_n V_n$ is also a Borel subset of the Hilbert cube. We conclude that $S$ is a Borel subset of the Hilbert cube.

The third part of the proof is an explicit construction.

**Lemma 21.6** *There is a continuous bijective Borel measurable function from a Borel subset $Y$ of $2^{\mathbf{N}_+}$ onto $[0,1]^{\mathbf{N}_+}$ with a Borel measurable inverse. In particular, $[0,1]^{\mathbf{N}_+}$ is isomorphic as a measure space with the Borel subset $Y$ of $2^{\mathbf{N}_+}$.*

Proof: There is a continuous function from $2^{\mathbf{N}_+}$ onto $[0,1]$ that is injective on a Borel subset $W$ of $2^{\mathbf{N}_+}$. The inverse of this continuous function is a measurable function from $[0,1]$ to the Borel subset $W$.

The continuous function from $2^{\mathbf{N}_+}$ to $[0,1]$ defines a continuous function from $2^{\mathbf{N}_+ \times \mathbf{N}_+}$ to $[0,1]^{\mathbf{N}_+}$. It maps $Y' = W^{\mathbf{N}_+}$ injectively onto $[0,1]^{\mathbf{N}_+}$ and has a measurable inverse. Since there is a bijection of $\mathbf{N}_+$ with $\mathbf{N}_+ \times \mathbf{N}_+$, there is a continuous bijection of $2^{\mathbf{N}_+}$ with $2^{\mathbf{N}_+ \times \mathbf{N}_+}$. This gives a continuous bijection of a subset $Y \subset 2^{\mathbf{N}_+}$ with $Y' \subset 2^{\mathbf{N}_+ \times \mathbf{N}_+}$, which in turn goes bijectively to $[0,1]^{\mathbf{N}_+}$. □

The assertion of the theorem is that if $X$ and $X'$ be uncountable Polish topological spaces, then the measurable space $X$ is isomorphic to the measurable space $X'$. It is enough to show that they are isomorphic to the cantor space $C = 2^{\mathbf{N}_+}$. Here is the remainder of the proof of the theorem.

Proof: Let $X$ be an uncountable Polish space. Then there exists a compact subset $K \subset X$ that is homeomorphic to $2^{\mathbf{N}_+}$. Furthermore, $X$ is isomorphic to a Borel subset $S$ of $[0,1]^{\mathbf{N}_+}$. Since there is a measurable isomorphism of $[0,1]^{\mathbf{N}_+}$ with a Borel subset $Y$ of $2^{\mathbf{N}_+}$, there is a measurable isomorphism of $S$ with a Borel subset $B$ of $C = 2^{\mathbf{N}_+}$.

These constructions give a measurable isomorphism of $X$ with $B$. Let $A \subset B$ be the image of $K$ under this isomorphism. Then $A \subset B \subset C$, where $A$ is measurable isomorphic to $C$. The lemma for the proof of Bernstein's theorem shows that $B$ must also be measurable isomorphic to $C$. This reasoning shows that every such $X$ is measurable isomorphic to $C$. □

## 21.5 A unique measure structure

There is another striking result that is an easy consequence of the theorem. This says that if $X$ is an uncountable Polish space, and $\mu$ is a finite non-zero Borel measure with no point masses, then the measure space $(X, \mathcal{B}, \mu)$ is isomorphic to Lebesgue measure on some closed and bounded interval $[0, M]$ of real numbers, with $0 < M < +\infty$. There is a corresponding result for a $\sigma$-finite measure, where the interval $[0, +\infty)$ is also allowed. In other words, for most practical purposes there is just one class of continuous $\sigma$-finite measure spaces of interest, classified by total mass $M$.

**Theorem 21.7** *Let $\mu$ be a finite Borel measure on an uncountable Polish space $X$. Let $M = \mu(X)$ be the total mass and suppose that $M > 0$. Suppose also there are no one point sets with non-zero measure. Then the measure space $(X, \mathcal{B}, \mu)$ is isomorphic to the measure space $([0, M], \mathcal{B}, \lambda)$.*

This is not a difficult result, since one can reduce the problem to the analysis of a finite measure on $[0,1]$. If the measure has no point masses, then it is given by an increasing continuous function from $[0,1]$ to $[0, M]$. This maps the measure to Lebesgue measure on $[0, M]$. The only problem that there may be intervals where the function is constant, so it is not bijective. This can be fixed. See the Chapter 15, Section 5 of the Royden text [17] for the detailed proof.

The conclusion is that for continuous finite measure spaces of this type the only invariant under isomorphism is the total mass. Otherwise all such measure spaces look the same. There is a corresponding theorem for $\sigma$-finite measures. In this situation there is the possibility of infinite total mass, corresponding to Lebesgue measure on the interval $[0, +\infty)$.

In the end, the measure spaces of practical interest are isomorphic to a countable set with point measures, an interval with Lebesgue measure, or a disjoint union of the two.

## Problems

1. Give an explicit construction to prove that the closed unit interval $[0, 1]$ is Borel isomorphic to the unit circle $T$.

# Chapter 22

# Measurable classification

## 22.1 Standard and substandard measurable spaces

A measurable space is standard if its $\sigma$-algebra is the Borel $\sigma$-algebra of some Polish space. It may be shown [4] that every measurable subset of a standard measurable space is a standard measurable space.

Let $Z$ be a measurable space. Then $Z$ is said to be *countably separated* if there is a countable family of measurable subsets that separate points.

**Theorem 22.1** *Let $Z$ be a measurable space. Then there is a measurable injection of $Z$ into a standard measurable space if and only if its $\sigma$-algebra $\mathcal{F}$ is countably separated.*

Proof: Let $A_1, A_2, A_3, \ldots$ be a sequence of measurable subsets that separate points. Define a function from $Z$ to the Cantor space by $f : Z \to \{0,1\}^{\mathbf{N}_+}$ by $f(x)_n = 1_{A_n}(x)$. Then $f$ is an injection.

The subsets $\{\omega \mid \omega_n = 1\}$ for $n = 1, 2, 3, \ldots$ generate the $\sigma$-algebra of the Cantor space. The inverse images of these sets are the $Y_n$. This shows that $f$ is a measurable function from $Z$ to $A$. $\square$

A measurable space is a *substandard* if it along with its $\sigma$-algebra is isomorphic to a subset of a standard measurable space with its $\sigma$-algebra. (The subset need not be measurable). This following is an easy characterization of substandard measurable spaces [9, 20]. Consider a measurable space with its $\sigma$-algebra. This is said to be *countably generated* if there is a countable sequence of measurable sets that generate the $\sigma$-algebra. The condition in the theorem below is that the $\sigma$-algebra is countably generated and separates points. These imply that the $\sigma$-algebra is countably separated, so both hypothesis and conclusion are stronger than in the previous result.

**Theorem 22.2** *Let $Z$ be a measurable space. Then there is a measurable injection of $Z$ into a standard measurable space that preserves measurable subsets if and only if its $\sigma$-algebra $\mathcal{F}$ is countably generated and separates points.*

Proof: To say that $\mathcal{F}$ separates points says that for every pair of of distinct points $x \neq y$ there is an element $B$ of $F$ such that $x \in B$ and $y \notin B$. Let $A_1, A_2, A_3, \ldots$ be a sequence of measurable subsets that generate $\mathcal{F}$. Then the $A_n$ also separate points. Define a function from $Z$ to the Cantor space by $f : Z \to \{0,1\}^{\mathbf{N}_+}$ by $f(x)_n = 1_{A_n}(x)$. Then $f$ is an injection.

The subsets $\{\omega \mid \omega_n = 1\}$ for $n = 1, 2, 3, \ldots$ generate the $\sigma$-algebra of the Cantor space. Let $Y$ be the subset of the Cantor space that is the image of $f$. Then $f$ maps $A_n$ onto the sets $f[A_n]$ that are the intersections of the subsets $\{\omega \mid \omega_n = 1\}$ with $A$. The $A_n$ generate the $\sigma$-algebra of $X$. The images $f[A_n]$ in $Y$ generate the $\sigma$-algebra of $Y$ induced from the $\sigma$-algebra of the Cantor space. This shows that $f$ is a measurable isomorphism of $Z$ with $A$. $\Box$

If $Z$ is standard, then $Z$ is substandard, and if $Z$ is substandard, then $Z$ is countably generated. For a countably generated $Z$ there is an measurable injection $f$ of $Z$ into a standard measurable space $Y$. This may be thought of as a classification of the points of $Z$ by a reasonable parameter space $Y$. The classification is nicer if $Z$ is substandard, since then $f$ can be a measurable isomorphism onto its range. It is particularly nice if $Z$ is standard, since in that case $f$ can be a measurable isomorphism onto a measurable subset.

## 22.2 Classification

Measurable spaces that are not standard measurable spaces arise in classification problems. This section is an introduction to this subject. Many of these ideas go back to work of Mackey [15].

Say that $X$ is a Polish space. Then by definition its Borel structure is a standard measurable structure. If $E$ is an equivalence relation on $X$, then the quotient space $X/E$ also has a measurable structure. If $q : X \to X/E$ is the natural map, then $W$ is a measurable subset of $X/E$ precisely when $q^{-1}[W]$ is a measurable measurable subset of $X$.

**Proposition 22.3** *Suppose that $X$ is a Polish space and $E$ is an equivalence relation. Then there is a measurable injection of $X/E$ into a standard measurable space $Y$ if and only if there is a measurable function $\theta : X \to Y$ such that*

$$xEy \Leftrightarrow \theta(x) = \theta(y). \tag{22.1}$$

Proof: Suppose that there is a measurable injection $f : X/E \to Y$, where $Y$ is a standard measurable space. Let $\theta$ be the composition of the natural map from $X/E$ with the map $f$ from $X/E$ to $Y$.

For the converse, suppose that there is such a map $\theta$ from $X$ to $Y$. For $t$ in $X/E$ take $x \in t$ and define $f(x) = \theta(t)$. This is well-defined and injective from $X/E$ to $Y$. It is easy to see from the definition of the quotient measurable structure that it is a measurable function. $\Box$

The above proposition describes when an equivalence relation $E$ on $X$ has a Polish parameter space $Y$ for the equivalence classes. (In this case some authors [9, 8] call $E$ a Borel smooth equivalence relation.) This seems almost a minimal requirement for effective classification. The parameterization need not be particularly nice, but at least it should be injective and measurable!

An equivalence relation on a Polish space $X$ is a *Borel equivalence relation* if it is a Borel subset of the product topological space. It follows that the equivalence classes are each Borel subsets of $X$. A Borel probability measure $\mu$ on $X$ is *ergodic* with respect to the equivalence relation $E$ if every invariant Borel subset (union of equivalence classes) has $\mu$ probability zero or one. The following result is from Becker and Kechris [1].

**Theorem 22.4** *Let $X$ be a Polish space, and let $E$ be a Borel equivalence relation. Suppose that there exists a Borel probability measure $\mu$ on $X$ that is ergodic for $E$. Suppose also that each equivalence class has $\mu$ measure zero. Then there is no measurable injection of the quotient measurable space $X/G$ into a standard measurable space.*

Proof: The measure $\mu$ on $X$ maps to a measure $\tilde{\mu}$ on $X/E$. The measure $\tilde{\mu}$ has the property that every measurable subset has measure zero or one. Furthermore, each one point subset has $\tilde{\mu}$ measure zero.

Suppose that $X/E$ were countably separated. Then there would be a measurable injection of $X/E$ into the unit interval $[0, 1]$. The probability measure $\tilde{\mu}$ would map to a Borel measure on $[0, 1]$. This would also have the property that every Borel subset has measure zero or one, and every one point set has measure zero. From the latter property, the distribution function would be continuous. But then by the intermediate value theorem there would be subsets with every possible measure between zero and one. □

Hjorth [8] presents another approach to results of this type. This approach makes no mention of measure. Instead it is required that $G$ is a topological group that is a Polish space, and $G$ acts continuously on the Polish space $X$. Suppose that (1) every orbit is dense, and (2) every orbit is meager. Then the quotient space $X/G$ cannot be injected measurably into a standard measurable space.

## 22.3 Orbits of dynamical systems

Consider the case where a group acts on $X$. The equivalence classes are the orbits of the group. In this case there are many classical cases where there is a measure that makes the action ergodic. See the book by Sinai [19] for an elementary introduction. Thus the conclusion is that the coset spaces cannot be parameterized, even measurably, by a Polish space.

Here is an example where the group is $\mathbb{Z}$, and it acts on the circle $T$ by rotation by an irrational angle.

**Theorem 22.5** *Let $T$ be the circle of circumference one with the rotationally invariant probability measure. Let $\alpha$ be an irrational number. Then rotation by $\alpha$ is ergodic.*

The group action is given by $n \cdot x = x + n\alpha$ modulo 1. Two points $x, x'$ are in the same orbit if there are $n, n'$ with $n \cdot x = n' \cdot x'$. This says that $x - x' = (n' - n)\alpha$ modulo 1. The proof that this is an ergodic action is given in the chapter on Fourier series.

The group does not have to be discrete. Consider the case is where the Polish space $X = T^2$ is the torus, and the group consists of the reals $\mathbb{R}$.

**Theorem 22.6** *Let $T^2$ be the that is the product of two circles, each of circumference one. The measure is Lebesgue measure. Let $\alpha$ and $\beta$ be numbers such that whenever $p$ and $q$ are integers with $p\alpha + q\beta = 0$, then $p = q = 0$. Then rotation by $\alpha, \beta$ is ergodic.*

Again the proof may be found in the chapter on Fourier series. These examples are quite concrete. In each case an equivalence relation on a Polish space gives rise to a quotient space. This quotient space appears to have no measurable parameterization by a Polish space. Thus some quite natural mathematical objects (quotient spaces) are apparently unclassifiable.

# Problems

1. Consider a measurable space $X$ with its $\sigma$-algebra $\mathcal{F}$ of subsets. It is said to be countably separated if there is a sequence $n \mapsto A_n$ of sets in $\mathcal{F}$ such that for every $x \neq y$ there is an $n$ with $x \in A_n$ and $y \notin A_n$.

   (a) Let $X$ be the unit interval with the Borel $\sigma$-algebra. Prove that $X$ is countably separated.

   (b) Let $X$ be the unit interval with the $\sigma$-algebra generated by the one point subsets. Prove that $X$ is not countably separated. Hint: Describe the $\sigma$-algebra explicitly. Which ordered pairs of points in the unit square can be separated by a countable sequence of elements of this $\sigma$-algebra?

2. Let $Z$ be an uncountable measurable space. Prove that $Z$ is standard if and only there is a sequence $A_1, A_2, A_3$ of measurable subsets that generate the $\sigma$-algebra of all measurable subsets, and such that the function $f : Z \to \{0,1\}^{\mathbf{N}_+}$ defined by $f(x)_n = 1_{A_n}(x)$ is a bijection.

3. Consider a rational rotation of the circle of circumference one. Describe the space of orbits. Describe its topology and its measurable structure.

# Part VI

# Function Spaces

# Chapter 23

# Function spaces

## 23.1 Spaces of continuous functions

This section records notations for spaces of real functions. In some contexts it is convenient to deal instead with complex functions; usually the changes that are necessary to deal with this case are minor. Our default is to take the functions as real functions, except in the context of Hilbert space and Fourier analysis.

Let $X$ be a topological space. The space $C(X)$ consists of all real continuous functions. The space $B(X)$ consists of all real bounded functions. It is a Banach space in a natural way. The space $BC(X)$ consists of all bounded continuous real functions. It is a somewhat smaller Banach space.

Consider now the special case when $X$ is a locally compact Hausdorff space. Thus each point has a compact neighborhood. For example $X$ could be $\mathbb{R}^n$. The space $C_c(X)$ consists of all continuous functions, each one of which has compact support. The space $C_0(X)$ is the closure of $C_c(X)$ in $BC(X)$. It is itself a Banach space. It is the space of continuous functions that vanish at infinity.

The relation between these spaces is that $C_c(X) \subset C_0(X) \subset BC(X)$. They are all equal when $X$ compact. When $X$ is locally compact, then $C_0(X)$ is the best behaved.

Recall that a *Banach space* is a normed vector space that is complete in the metric associated with the norm. In the following we shall need the concept of the *dual space* of a Banach space $E$. The dual space $E^*$ consists of all continuous linear functions from the Banach space to the real numbers. (If the Banach space has complex scalars, then we take continuous linear function from the Banach space to the complex numbers.) The dual space $E^*$ is itself a Banach space, where the norm is the Lipschitz norm.

For certain Banach spaces $E$ of functions the linear functionals in the dual space $E^*$ may be realized in a more concrete way. For example, suppose that $X$ is a Polish space (a separable completely metrizable space) that is locally compact. (This is equivalent to being a second countable locally compact Hausdorff space.) If $E = C_0(X)$, then its dual space $E^* = M(X)$ is a Banach space consisting of finite signed Borel measures. (A finite signed measure $\sigma$ is the difference $\sigma = \sigma_+ - \sigma$ of two finite positive measures $\sigma_\pm$.) If $\sigma$ is in $M(X)$, then it defines the linear functional $f \mapsto \int f(x)\, d\sigma(x)$, and all elements of the dual space $E^*$ arise in this way.

## 23.2 The Stone-Weierstrass theorem

Let $X$ be a compact Hausdorff space, and let $C(X)$ be the space of continuous real functions on $X$. It is fairly difficult to approximate in $C(X)$, since this is uniform approximation. Nevertheless, there is a powerful way of finding a dense subset of $C(X_)$, given by the *Stone-Weierstrass theorem* .

A collection of functions is an *algebra of functions* if it is a vector space of functions that is also closed under pointwise multiplication. A collection of functions separates points if for each $x \neq y$ there is a function $f$ in the collection with $f(x) \neq f(y)$.

**Theorem 23.1 (Stone-Weierstrass)** *Consider a algebra of real functions in $C(X)$ that includes the constant functions. Suppose that it separates points. Then it is dense in $C(X)$.*

There are proofs of this theorem in Folland [5] and Dudley [4]. The classic application is the original Weierstrass approximation theorem. In that case $X = [a, b]$ and the functions in the collection consist of all real polynomials. The conclusion is that every real continuous function may be uniformly approximated on $[a, b]$ by a real polynomial.

Here is another example. Take $X$ to be the circle of circumference $2\pi$. Consider the collection of functions that are finite real linear combinations of $\cos(nx)$ for $n = 0, 1, 2, 3, \ldots$ and $\sin(nx)$ for $n = 1, 2, 3, \ldots$. Trigonometric identities show that these form an algebra of functions. The Stone-Weierstrass theorem shows that every real continuous periodic function may be approximated uniformly by such trigonometric functions.

There is also a version of the Stone-Weierstrass theorem for the space of complex functions, but it also requires that the collection be invariant under taking the complex conjugation. This gives a more transparent way of treating the last example. Take $X$ to be the circle of circumference $2\pi$. Consider the collection of functions that are finite complex linear combinations of functions $e^{inx}$ for $n$ in $\mathbb{Z}$. It is completely elementary that this is an algebra of functions. Furthermore, the complex conjugate of such a function $e^{inx}$ is another function $e^{-inx}$ of the same kind. It follows that these functions are uniformly dense in the space $C(X)$ of continuous complex periodic functions.

## 23.3 Pseudometrics and seminorms

A pseudometric is a function $d : P \times P \to [0, +\infty)$ that satisfies $d(f, f) \geq 0$ and $d(f, g) \leq d(f, h) + d(h, g)$ and such that $d(f, f) = 0$. If in addition $d(f, g) = 0$

implies $f = g$, then $d$ is a metric.

The theory of pseudometric spaces is much the same as the theory of metric spaces. The main difference is that a sequence can converge to more than one limit. However each two limits of the sequence have distance zero from each other, so this does not matter too much.

Given a pseudometric space $P$, there is an associated metric space $M$. This is defined to be the set of equivalence classes of $P$ under the equivalence relation $fEg$ if and only if $d(f, g) = 0$. In other words, one simply defines two points $r, s$ in $P$ that are at zero distance from each other to define the same point $r' = s'$ in $M$. The distance $d_M(a, b)$ between two points $a, b$ in $M$ is defined by taking representative points $p, q$ in $P$ with $p' = a$ and $q' = b$. Then $d_M(a, b)$ is defined to be $d(p, q)$.

A *seminorm* is a function $f \mapsto \|f\| \geq 0$ on a vector space $E$ that satisfies $\|cf\| = |c|\|f\|$ and the triangle inequality $\|f + g\| \leq \|f\| + \|g\|$ and such that $\|0\| = 0$. If in addition $\|f\| = 0$ implies $f = 0$, then it is a norm. Each seminorm on $E$ defines a pseudo-metric for $E$ by $d(f, g) = \|f - g\|$. Similarly, a norm on $E$ defines a metric for $E$.

Suppose that we have a seminorm on $E$. Then the set of $h$ in $E$ with $\|h\| = 0$ is a vector subspace of $E$. The set of equivalence classes in the construction of the metric space is itself a vector space in a natural way. So for each vector space with a seminorm we can associate a new quotient vector space with a norm.

## 23.4 $\mathcal{L}^p$ spaces

In this and the next sections we introduce the spaces $\mathcal{L}^p(X, \mathcal{F}, \mu)$ and the corresponding quotient spaces $L^p(X, \mathcal{F}, \mu)$.

Fix a set $X$ and a $\sigma$-algebra $\mathcal{F}$ of measurable functions. Let $0 < p < \infty$. Define

$$\|f\|_p = \mu(|f|^p)^{\frac{1}{p}}. \tag{23.1}$$

Define $\mathcal{L}^p(X, \mathcal{F}, \mu)$ to be the set of all $f$ in $\mathcal{F}$ such that $\|f\|_p < \infty$.

**Theorem 23.2** *For $0 < p < \infty$, the space $\mathcal{L}^p$ is a vector space.*

Proof: It is obvious that $L^p$ is closed under scalar multiplication. The problem is to prove that it is closed under addition. However if $f$, $g$ are each in $L^p$, then

$$|f + g|^p \leq [2(|f| \vee |g|)]^p \leq 2^p(|f|^p + |g|^p). \tag{23.2}$$

Thus $f + g$ is also in $L^p$. $\square$

The function $x^p$ is increasing for every $p > 0$. In fact, if $\phi(p) = x^p$ for $x \geq 0$, then $\phi'(p) = px^{p-1} \geq 0$. However it is convex only for $p \geq 1$. This is because in that case $\phi''(x) = p(p-1)x^{p-2} \geq 0$.

Let $a \geq 0$ and $b \geq 0$ be weights with $a + b = 1$. For a convex function we have the inequality $\phi(au + bv) \leq a\phi(u) + b\phi(v)$. This is the key to the *Minkowski inequality*.

**Theorem 23.3** *(Minkowski inequality) If $1 \le p < \infty$, then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{23.3}$$

Proof: Let $c = \|f\|_p$ and $d = \|g\|_p$. Then by the fact that $x^p$ is increasing and convex

$$\left|\frac{f+g}{c+d}\right|^p \le \left(\frac{c}{c+d}\left|\frac{f}{c}\right| + \frac{d}{c+d}\left|\frac{g}{d}\right|\right)^p \le \frac{c}{c+d}\left|\frac{f}{c}\right|^p + \frac{d}{c+d}\left|\frac{g}{d}\right|^p. \tag{23.4}$$

Integrate. This gives

$$\mu\left(\left|\frac{f+g}{c+d}\right|^p\right) \le 1. \tag{23.5}$$

Thus $\|f+g\|_p \le c+d$. □

The preceding facts show that $\mathcal{L}^p$ is a vector space with a seminorm. It is a fact that $\mu(|f|^p) = 0$ if and only if $f = 0$ almost everywhere. Thus for $f$ in $\mathcal{L}^p$ we have $\|f\|_p = 0$ if and only if $f = 0$ almost everywhere.

**Theorem 23.4** *(dominated convergence for $\mathcal{L}^p$) Let $0 < p < \infty$. Let $f_n \to f$ pointwise. Suppose that there is a $g \ge 0$ in $\mathcal{L}^p$ such that each $|f_n| \le g$. Then $f_n \to f$ in $\mathcal{L}^p$, that is, $\|f_n - f\|_p \to 0$.*

Proof: If each $|f_n| \le g$ and $f_n \to f$ pointwise, then $|f| \le g$. Thus $|f_n - f| \le 2g$ and $|f_n - f|^p \le 2^p g^p$. Since $g^p$ has finite integral, the integral of $|f_n - f|^p$ approaches zero, by the usual dominated convergence theorem. □

It would be an error to think that just because $g_n \to g$ in the $\mathcal{L}^p$ sense it would follow that $g_n \to g$ almost everywhere. Being close on the average does not imply being close at a particular point. Consider the following example. For each $n = 1, 2, 3, \ldots$, write $n = 2^k + j$, where $k = 0, 1, 2, 3, \ldots$ and $0 \le j < 2^k$. Consider a sequence of functions defined on the unit interval $[0, 1]$ with the usual Lebesgue measure. Let $g_n = 1$ on the interval $[j/2^k, (j+1)/2^k]$ and $g_n = 0$ elsewhere in the unit interval. Then the $\mathcal{L}^1$ seminorm of $g_n$ is $1/2^k$, so the $g_n \to 0$ in the $\mathcal{L}^1$ sense. On the other hand, given $x$ in $[0, 1]$, there are infinitely many $n$ for which $g_n(x) = 0$ and there are infinitely many $n$ for which $g_n(x) = 1$. So pointwise convergence fails at each point.

Say that a seminormed vector space is sum complete if every absolutely convergent series is convergent to some limit. Recall that it is complete (as a pseudometric space) if every Cauchy sequence converges to some limit.

**Lemma 23.5** *Consider a seminormed vector space. If the space is sum complete, then it is complete.*

Proof: Suppose that $E$ is a seminormed vector space that is sum complete. Suppose that $g_n$ is a Cauchy sequence. This means that for every $\epsilon > 0$ there is an $N$ such that $m, n \ge N$ implies $\|g_m - g_n\| < \epsilon$. The idea is to show that $g_n$ has a subsequence that converges very rapidly. Let $\epsilon_k$ be a sequence such that $\sum_{k=1}^{\infty} \epsilon_k < \infty$. In particular, for each $k$ there is an $N_k$ such that $m, n \ge N_k$

implies $\|g_m - g_n\| < \epsilon_k$. The desired subsequence is the $g_{N_k}$. Define a sequence $f_1 = g_{N_1}$ and $f_j = g_{N_j} - g_{N_{j-1}}$ for $j \geq 2$. Then

$$g_{N_k} = \sum_{j=1}^{k} f_j. \tag{23.6}$$

Furthermore,

$$\sum_{j=1}^{\infty} \|f_j\| \leq \|f_1\| + \sum_{j=2}^{\infty} \epsilon_{j-1} < \infty. \tag{23.7}$$

This says that the series is absolutely convergence. Since $E$ is sum complete, the series converges to some limit, that is, there exists a $g$ such that the subsequence $g_{N_k}$ converges to $g$. Since the sequence $g_n$ is Cauchy, it also must converge to the same $g$. Thus $E$ is complete. Thus the theorem follows. □

**Theorem 23.6** *For $1 \leq p < \infty$ the space $\mathcal{L}^p$ is complete.*

Proof: Suppose that $\sum_{j=1}^{\infty} f_j$ is absolutely convergent in $\mathcal{L}^p$, that is,

$$\sum_{j=1}^{\infty} \|f_j\|_p = B < \infty. \tag{23.8}$$

Then by using Minkowski's inequality

$$\|\sum_{j=1}^{k} |f_j|\|_p \leq \sum_{j=1}^{k} \|f_j\|_p \leq B. \tag{23.9}$$

By the monotone convergence theorem $h = \sum_{j=1}^{\infty} |f_j|$ is in $\mathcal{L}^p$ with $\mathcal{L}^p$ seminorm bounded by $B$. In particular, it is convergent almost everywhere. It follows that the series $\sum_{j=1}^{\infty} f_j$ converges almost everywhere to some limit $g$. The sequence $\sum_{j=1}^{k} f_j$ is dominated by $h$ in $\mathcal{L}^p$ and converges pointwise to $\sum_{j=1}^{\infty} f_j$. Therefore, by the dominated convergence theorem, it converges to the same limit $g$ in the $\mathcal{L}^p$ seminorm. □

**Corollary 23.7** *If $1 \leq p < \infty$ and if $g_n \to g$ in the $\mathcal{L}^p$ seminorm sense, then there is a subsequence $g_{N_k}$ such that $g_{N_k}$ converges to $g$ almost everywhere.*

Proof: Let $g_n \to g$ as $n \to \infty$ in the $\mathcal{L}^p$ sense. Then $g_n$ is a Cauchy sequence in the $\mathcal{L}^p$ sense. Let $\epsilon_k$ be a sequence such that $\sum_{k=1}^{\infty} \epsilon_k < \infty$. Let $N_k$ be a subsequence such that $n \geq N_k$ implies $\|g_n - g_{N_k}\|_p < \epsilon_k$. Define a sequence $f_k$ such that

$$g_{N_k} = \sum_{j=1}^{k} f_j. \tag{23.10}$$

Then $\|f_j\|_p = \|g_{N_j} - g_{N_{j-1}}\|_p \leq \epsilon_{j-1}$ for $j \geq 2$. By the monotone convergence theorem

$$h = \sum_{j=1}^{\infty} |f_j| \tag{23.11}$$

converges in $\mathcal{L}^p$ and is finite almost everywhere. It follows that

$$g = \sum_{j=1}^{\infty} f_j \tag{23.12}$$

converges in $\mathcal{L}^p$ and also converges almost everywhere In particular, $g_{N_k} \to g$ as $k \to \infty$ almost everywhere. □

In order to complete the picture, define

$$\|f\|_\infty = \inf\{M \geq 0 \mid |f| \leq M \text{ almost everywhere }\}. \tag{23.13}$$

This says that $\|f\|_\infty \leq M$ if and only if $\mu(|f| > M) = 0$. In other words, $M < \|f\|_\infty$ if and only if $\mu(|f| > M) > 0$. The space $\mathcal{L}^\infty(X, \mathcal{F}, \mu)$ consists of all functions $f$ in $\mathcal{F}$ such that $\|f\|_\infty < \infty$. The number $\|f\|_\infty$ is called the *essential supremum* of $f$ with respect to $\mu$. The space $\mathcal{L}^\infty(X, \mathcal{F}, \mu)$ is a vector space with a seminorm. The following theorem is also simple.

**Theorem 23.8** *The space $\mathcal{L}^\infty(X, \mathcal{F}, \mu)$ is complete.*

Among the $\mathcal{L}^p$ spaces the most important are $\mathcal{L}^1$ and $\mathcal{L}^2$ and $\mathcal{L}^\infty$. Convergence in $\mathcal{L}^1$ is also called convergence in mean, more precisely, in *mean absolute value*. Convergence in $\mathcal{L}^2$ is convergence in *root mean square*. (Sometimes this is abbreviated to RMS.) Convergence in $\mathcal{L}^\infty$ is a measure theory version of uniform convergence.

## 23.5 Dense subspaces of $\mathcal{L}^p$

In this section we see that it is easy to approximate in $L^p$, at least for $1 \leq p < \infty$. The advantage is that the approximation does not have to be uniform.

For a function to be in $\mathcal{L}^p(X, \mathcal{F}, \mu)$ it is not only required that $\mu(|f|^p) < \infty$, but also that $f$ is measurable, that is, that $f$ is in $\mathcal{F}$. This requirement has important consequences for approximation.

**Theorem 23.9** *Let $X$ be a set, $\mathcal{F}$ a $\sigma$-algebra of real measurable functions on $X$, and $\mu$ an integral. Consider the space $\mathcal{L}^p(X, \mathcal{F}, \mu)$ for $1 \leq p < \infty$. Let $L$ be a vector lattice of functions with $L \subset \mathcal{L}^p(X, \mathcal{F}, \mu)$. Suppose that the smallest monotone class including $L$ is $\mathcal{F}$. Then $L$ is dense in $\mathcal{L}^p(X, \mathcal{F}, \mu)$. That if, if $f$ is in $\mathcal{L}^p(X, \mathcal{F}, \mu)$ and if $\epsilon > 0$, then there exists $h$ in $L$ with $\|h - f\|_p < \epsilon$.*

Proof: We know that the smallest monotone class including $L^+$ is $\mathcal{F}^+$. Let $g$ be in $L^+$. Let $S_g$ be the set of all $f \geq 0$ such that $f \wedge g$ is in the $\mathcal{L}^p$ closure of $L$. Clearly $L^+ \subset S_g$, since if $f$ is in $L^+$ then so is $f \wedge g$. Furthermore, $S_g$ is closed under increasing and decreasing limits. Here is the proof for increasing limits. Say that $f_n$ is in $S_g$ and $f_n \uparrow f$. Then $f_n \wedge g \uparrow f \wedge g \leq g$. By the $\mathcal{L}^p$ monotone convergence theorem, $\|f_n \wedge g - f \wedge g\|_p \to 0$. Since $f_n \wedge g$ is in the $\mathcal{L}^p$ closure of $L$, it follows that $f \wedge g$ is also in the $\mathcal{L}^p$ closure of $L$. However this says that $f$ is in $S_g$. It follows from this discussion that $\mathcal{F}^+ \subset S_g$. Since $g$ is arbitrary, this proves that $f$ in $\mathcal{F}^+$ and $g$ in $L^+$ implies $f \wedge g$ is in the $\mathcal{L}^p$ closure of $L$.

Let $f \geq 0$ be in $\mathcal{L}^p$. Let $S'_f$ be the set of all $h \geq 0$ such that $f \wedge h$ is in the $\mathcal{L}^p$ closure of $L$. From the preceding argument, we see that $L \subset S'_f$. Furthermore, $S'_f$ is closed under increasing and decreasing limits. Here is the proof for increasing limits. Say that $h_n$ is in $S'_f$ and $h_n \uparrow h$. Then $f \wedge h_n \uparrow f \wedge h \leq f$. By the $\mathcal{L}^p$ monotone convergence theorem, $\|f \wedge h_n - f \wedge h\|_p \to 0$. Since $f \wedge h_n$ is in the $\mathcal{L}^p$ closure of $L$, it follows that $f \wedge h$ is also in the $\mathcal{L}^p$ closure of $L$. However this says that $h$ is in $S_f$. It follows from this discussion that $\mathcal{F}^+ \subset S'_f$. Since $f$ is arbitrary, this proves that $f \geq 0$ in $\mathcal{L}^p$ and $h$ in $\mathcal{F}^+$ implies $f \wedge h$ is in the $\mathcal{L}^p$ closure of $L$. Take $h = f$. Thus $f \geq 0$ in $\mathcal{L}^p$ implies $f$ is in the $\mathcal{L}^p$ closure of $L$. □

**Corollary 23.10** *Take $1 \leq p < \infty$. Consider the space $\mathcal{L}^p(\mathbb{R}, \mathcal{B}, \mu)$, where $\mu$ is a measure that is finite on compact subsets. Let $L$ be the space of step functions, or let $L$ be the space of continuous functions with compact support. Then $L$ is dense in $\mathcal{L}^p(\mathbb{R}, \mathcal{B}, \mu)$.*

This result applies in particular to the case $\mu = \lambda$ of Lebesgue measure. Notice that nothing like this is true for $\mathcal{L}^\infty(\mathbb{R}, \mathcal{B}, \lambda)$. The uniform limit of a sequence of continuous functions is continuous, and so if we start with continuous functions and take uniform limits, we stay in the class of continuous functions. But functions in $\mathcal{L}^\infty(\mathbb{R}, \mathcal{B}, \lambda)$ can be discontinuous in such a way that cannot be fixed by changing the function on a set of measure zero. Even a step function has this property.

Remark. So people might argue that the so-called delta function $\delta(x)$ is in $\mathcal{L}^1$, since it has integral $\int_{-\infty}^{\infty} \delta(x) = 1$. Actually the delta function is a measure, not a function, so this is not correct. But there is a stronger sense in which this is not correct. Let $h$ be an arbitrary continuous function with compact support. Look at the distance from $\delta(x)$ to $h(x)$. This is the integral $\int_{-\infty}^{\infty} |\delta(x) - h(x)|\, dx$ which always has a value one or bigger. The delta function is thus not even close to being in $\mathcal{L}^1$.

## 23.6 The quotient space $L^p$

The space $\mathcal{L}^p(X, \mathcal{F}, \mu)$ is defined for $1 \leq p \leq \infty$. It is a vector space with a seminorm, and it is complete. One can associate with this the space $L^p(X, \mathcal{F}, \mu)$, where two elements $f, g$ of $\mathcal{L}^p(X, \mathcal{F}, \mu)$ define the same element of $L^p(X, \mathcal{F}, \mu)$ provided that $f = g$ almost everywhere with respect to $\mu$. Then this is a vector space with a norm, and it is complete. In other words, it is a Banach space.

This passage from a space of functions $\mathcal{L}$ to the corresponding quotient space $L^p$ is highly convenient, but also confusing. The elements of $L^p$ are not functions, and so they do not have values defined at particular points of $X$. Nevertheless they are come from functions.

It is convenient to work with the spaces $L^p$ abstractly, but to perform all calculations with the corresponding functions in $\mathcal{L}^p$. For this reason people often use the notation $L^p$ to refer to either space, and we shall follow this practice in most of the following, unless there is a special point to be made.

However be warned, these spaces can be very different. As an example, take the space $\mathcal{L}^1(\mathbb{R}, \mathcal{B}, \delta_a)$, where $\delta_a$ is the measure that assigns mass one to the point $a$. Thus the corresponding integral is $\delta_a(f) = f(a)$. This space consists of all Borel functions, so it is infinite dimensional. However two such functions are equal almost everywhere with respect to $\delta_a$ precisely when they have the same values at the point $a$. Thus the quotient space $L^1(\mathbb{R}, \mathcal{B}, \delta)$ is one dimensional. This is a much smaller space. But it captures the notion that from the perspective of the measure $\delta_a$ the points other than $a$ are more or less irrelevant.

## 23.7 Duality of $L^p$ spaces

In this section we describe the duality theory for the Banach spaces $L^p$. We begin with the *arithmetic-geometric mean inequality.* This will immediately give the famous *Hölder inequality.*

**Lemma 23.11** *(arithmetic-geometric mean inequality) Let $a \geq 0$ and $b \geq 0$ with $a + b = 1$. Let $z > 0$ and $w > 0$. Then $z^a w^b \leq az + bw$.*

Proof: Since the exponential function is convex, we have $e^{au+bv} \leq ae^u + be^v$. Set $z = e^u$ and $w = e^v$. □

**Lemma 23.12** *Let $p > 1$ and $q > 1$ with $1/p + 1/q = 1$. If $x > 0$ and $y > 0$, then*

$$xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q. \tag{23.14}$$

Proof: Take $a = 1/p$ and $b = 1/q$, and substitute $e^a = x$ and $e^b = y$. □

**Theorem 23.13** *(Hölder's inequality) Suppose that $1 < p < \infty$ and that $1/p + 1/q = 1$. Then*

$$|\mu(fg)| \leq \|f\|_p \|g\|_q. \tag{23.15}$$

Proof: It is sufficient to prove this when $\|f\|_p = 1$ and $\|g\|_p = 1$. However by the lemma

$$|f(x)||g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q. \tag{23.16}$$

Integrate. □

This lemma shows that if $g$ is in $L^q(\mu)$, with $1 < q < \infty$, then the linear functional defined by $f \mapsto \mu(fg)$ is continuous on $L^p(\mu)$, where $1 < p < \infty$ with $1/p + 1/q = 1$. This shows that each element of $L^q(\mu)$ defines an element of the dual space of $L^p(\mu)$. It may be shown that every element of the dual space arises in this way. Thus the dual space of $L^p(\mu)$ is $L^q(\mu)$, for $1 < p < \infty$.

Notice that we also have a Hölder inequality in the limiting case:

$$|\mu(fg)| \leq \|f\|_1 \|g\|_\infty. \tag{23.17}$$

This shows that every element $g$ of $L^\infty(\mu)$ defines an element of the dual space of $L^1(\mu)$. It may be shown that if $\mu$ is $\sigma$-finite, then $L^\infty(\mu)$ is the dual space of $L^1(\mu)$.

On the other hand, each element $f$ of $L^1(\mu)$ defines an element of the dual space of $L^\infty(\mu)$. However in general this does not give all elements of the dual space of $L^\infty(\mu)$.

The most important spaces are $L^1$, $L^2$, and $L^\infty$. The nicest by far is $L^2$, since it is a Hilbert space. The space $L^1$ is also common, since it measures the total amount of something. The space $L^\infty$ goes together rather naturally with $L^1$. Unfortunately, the theory of the spaces $L^1$ and $L^\infty$ is more delicate than the theory of the spaces $L^p$ with $1 < p < \infty$. Ultimately this is because the spaces $L^p$ with $1 < p < \infty$ have better convexity properties.

Here is a brief summary of the facts about duality. The dual space of a Banach space is the space of continuous linear scalar functions on the Banach space. The dual space of a Banach space is a Banach space. Let $1/p + 1/q = 1$, with $1 \leq p < \infty$ and $1 < q \leq \infty$. (Require that $\mu$ be $\sigma$-finite when $p = 1$.) Then the dual of the space $L^p(X, \mathcal{F}, \mu)$ is the space $L^q(X, \mathcal{F}, \mu)$. The dual of $L^\infty(X, \mathcal{F}, \mu)$ is not in general equal to $L^1(X, \mathcal{F}, \mu)$. Typically $L^1(X, \mathcal{F}, \mu)$ is not the dual space of anything. The fact that is often used instead is that $M(X)$ is the dual of $C_0(X)$.

There is an advantage to identifying a Banach space $E^*$ as the dual space of another Banach space $E$. This can be done for $E^* = M(X)$ and for $E^* = \mathrm{L}^q(X, \mathcal{F}, \mu)$ for $1 < q \leq \infty$ (with $\sigma$-finiteness in the case $q = \infty$). Then $E^*$ is the space of all continuous linear functionals on the original space $E$. There is a corresponding notion of pointwise convergence in $E^*$, called weak$^*$ convergence, and this turns out to have useful properties that make it a convenient technical tool.

Spaces of sequences provide a particularly illuminating example. Let $c_0$ be the space of all sequences that converge to zero. It may be thought of as a space of continuous functions that vanish at infinity. Its dual space is $\ell^1$, the space of absolutely summable sequences. In this context $\ell^1$ is analogous to a space of finite signed measures. On the other hand, we may think of $\ell^1$ as a space of integrable functions, so its dual space is $\ell^\infty$. This gives a concrete example where the double dual $\ell^\infty$ is larger than the original space $c_0$.

## 23.8 Supplement: Orlicz spaces

It is helpful to place the theory of $L^p$ spaces in a general context. Clearly, the theory depends heavily on the use of the functions $x^p$ for $p \geq 1$. This is a convex function. The generalization is to use a more or less arbitrary convex function.

Let $H(x)$ be a continuous function defined for all $x \geq 0$ such that $H(0) = 0$ and such that $H'(x) > 0$ for $x > 0$. Then $H$ is an increasing function. Suppose that $H(x)$ increases to infinity as $x$ increases to infinity. Finally, suppose that $H''(x) \geq 0$. This implies convexity.

Example: $H(x) = x^p$ for $p > 1$.

Example: $H(x) = e^x - 1$.

Example: $H(x) = (x + 1)\log(x + 1)$.

Define the size of $f$ by $\mu(H(|f|))$. This is a natural notion, but it does not have good scaling properties. So we replace $f$ by $f/c$ and see if we can make the size of this equal to one. The $c$ that accomplishes this will be the norm of $f$.

This leads to the official definition of the Orlicz norm

$$\|f\|_H = \inf\{c > 0 \mid \mu\left(H\left(|f/c|\right)\right) \leq 1. \tag{23.18}$$

When this norm is finite, then $f$ is said to belong to the *Orlicz space* corresponding to the function $H$.

It is not difficult to show that if this norm is finite, then we can find a $c$ such that

$$\mu\left(H\left(|f/c|\right)\right) = 1. \tag{23.19}$$

Then the definition takes the simple form

$$\|f\|_H = c, \tag{23.20}$$

where $c$ is defined by the previous equation.

It is not too difficult to show that this norm defines a Banach space $L_H(\mu)$. The key point is that the convexity of $H$ makes the norm satisfy the triangle inequality.

**Theorem 23.14** *The Orlicz norm satisfies the triangle inequality*

$$\|f + g\|_H \leq \|f\|_H + \|g\|_H. \tag{23.21}$$

Proof: Let $c = \|f\|_H$ and $d = \|g\|_H$. Then by the fact that $H$ is increasing and convex

$$H\left(\left|\frac{f+g}{c+d}\right|\right) \leq H\left(\frac{c}{c+d}\left|\frac{f}{c}\right| + \frac{d}{c+d}\left|\frac{g}{d}\right|\right) \leq \frac{c}{c+d}H\left(\left|\frac{f}{c}\right|\right) + \frac{d}{c+d}H\left(\left|\frac{g}{d}\right|\right). \tag{23.22}$$

Integrate. This gives

$$\mu\left(H\left(\left|\frac{f+g}{c+d}\right|\right)\right) \leq 1. \tag{23.23}$$

Thus $\|f+g\|_H \leq c+d$. □

Notice that this result is a generalization of Minkowski's inequality. So we see that the idea behind $L^p$ spaces is convexity. The convexity is best for $1 < p < \infty$, since then the function $x^p$ has second derivative $p(p-1)x^{p-2} > 0$. (For $p = 1$ the function $x$ is still convex, but the second derivative is zero, so it not strictly convex.)

One can also try to create a duality theory for Orlicz spaces. For this it is convenient to make the additional assumptions that $H'(0) = 0$ and $H''(x) > 0$ and $H'(x)$ increases to infinity.

The dual function to $H(x)$ is a function $K(y)$ called the *Legendre transform.* The definition of $K(y)$ is

$$K(y) = xy - H(x), \tag{23.24}$$

where $x$ is defined implicitly in terms of $y$ by $y = H'(x)$.

This definition is somewhat mysterious until one computes that $K'(y) = x$. Then the secret is revealed: The functions $H'$ and $K'$ are inverse to each other. Furthermore, the Legendre transform of $K(y)$ is $H(x)$.

Examples:

1. Let $H(x) = x^p/p$. Then $K(y) = y^q/q$, where $1/p + 1/q = 1$.
2. Let $H(x) = e^x - 1 - x$. Then $K(y) = (y+1)\log(y+1) - y$.

**Lemma 23.15** *Let $H(x)$ have Legendre transform $K(y)$. Then for all $x \geq 0$ and $y \geq 0$*

$$xy \leq H(x) + K(y). \tag{23.25}$$

Proof: Fix $y$ and consider the function $xy - H(x)$. Since $H'(x)$ is increasing to infinity, the function rises and then dips below zero. It has its maximum where the derivative is equal to zero, that is, where $y - H'(x) = 0$. However by the definition of Legendre transform, the value of $xy - H(x)$ at this point is $K(y)$. □

**Theorem 23.16** *(Hölder's inequality) Suppose that $H$ and $K$ are Legendre transforms of each other. Then*

$$|\mu(fg)| \leq 2\|f\|_H\|g\|_K. \tag{23.26}$$

Proof: It is sufficient to prove this when $\|f\|_H = 1$ and $\|g\|_K = 1$. However by the lemma

$$|f(x)||g(x)| \leq H(|f(x)|) + K(|g(y)|). \tag{23.27}$$

Integrate. □

This is just the usual derivation of Hölder's inequality. However if we take $H(x) = x^p/p$, $K(y) = y^q/q$, then the $H$ and $K$ norms are not quite the usual $L^p$ and $L^q$, but instead multiples of them. This explains the extra factor of 2. In any case we see that the natural context for Hölder's inequality is the Legendre transform for convex functions. For more on this general subject, see Appendix H (Young-Orlicz spaces) in the treatise of Dudley [3].

# Problems

1. Let $X = \mathbb{R}$, let $\mathcal{B}o$ be the real Borel functions on $\mathbb{R}$, and let the integral $\mu$ be defined by
$$\mu(f) = \left(\frac{1}{2}\right)^n \sum_{k=0}^{n} \binom{n}{k} f(k). \tag{23.28}$$
(a) Show that $\mu(1) = 1$. Hint: The number $2^m$ of subsets of an $m$ element set may be written as the sum $\sum_{j=0}^{m} \binom{m}{j} = 2^m$ over $j$ of the number of $j$ element subsets. Here $\binom{m}{j} = \frac{m!}{j!(m-j)!}$.
(b) What is the dimension of the vector space $\mathcal{L}^1(\mathbb{R}, \mathcal{B}o, \mu)$?
(c) What is the dimension of the vector space $L^1(\mathbb{R}, \mathcal{B}o, \mu)$?
(d) Let $f(x) = x$. Is $f$ in $\mathcal{L}^1(\mathbb{R}, \mathcal{B}o, \mu)$? If so, then compute $\mu(f)$. If not, explain the problem.

2. In this problem $\ell^p$ is a space of real sequences indexed by natural numbers with counting measure, and $L^p$ is a space of Borel measurable functions on the unit interval with Lebesgue measure. In each problem give a yes or no answer, together with a proof or counterexample.
(a) Is $\ell^1 \subset \ell^2$?
(b) Is $L^2 \subset L^1$?
(c) Is $\ell^1$ dense in $\ell^\infty$?
(d) Is $L^\infty$ dense in $L^1$?

3. Prove that
$$\left(\int_1^\infty |f(x)|\, dx\right)^5 \le 4 \left(\int_1^\infty x^{\frac{5}{16}} |f(x)|^{\frac{5}{4}}\, dx\right)^4. \tag{23.29}$$

4. Consider complex Borel functions on the unit interval $[0, 1]$ with Lebesgue measure. Let $k(x, y)$ be a complex function such that
$$c^2 = \int_0^1 \int_0^1 |k(x, y)|^2\, dx\, dy < \infty. \tag{23.30}$$
Define a linear transformation $K$ on $L^2$ by
$$(Kf)(x) = \int_0^1 k(x, y) f(y)\, dy. \tag{23.31}$$
(a) Show that $K$ is continuous from $L^2$ to $L^2$.
(b) Show that if $c < 1$, then for each $g$ in $L^2$ the equation
$$Kf + g = f \tag{23.32}$$
has a unique solution $f$ in $L^2$. Hint: Define the map $T$ from $L^2$ to $L^2$ by $Tf = Kf + g$.

5. Consider Lebesgue measure $\lambda$ defined for Borel functions $\mathcal{B}$ defined on $\mathbb{R}$. Let $1 \leq q < r < \infty$. (a) Give an example of a function in $\mathcal{L}^q$ that is not in $\mathcal{L}^r$. (b) Give an example of a function in $\mathcal{L}^r$ that is not in $\mathcal{L}^q$.

6. Consider a finite measure $\mu$ (the measure of the entire set or the integral of the constant function 1 is finite). Show that if $1 \leq q \leq r \leq \infty$, then $\mathcal{L}^r \subset \mathcal{L}^q$.

7. Let $\phi$ be a smooth convex function, so that that for each $a$ and $t$ we have $\phi(t) \geq \phi(a) + \phi'(a)(t-a)$. Let $\mu$ be a probability measure. Let $f$ be a real function in $\mathcal{L}^1$. Show that $\phi(\mu(f)) \leq \mu(\phi(f))$. (This is Jensen's inequality.) Hint: Let $a = \mu(f)$ and $t = f$. Where do you use the fact that $\mu$ is a probability measure?

8. Let $\phi$ be a smooth convex function as above. Deduce from the preceding problem the simple fact that if $0 \leq a$ and $0 \leq b$ with $a + b = 1$, then $\phi(au + bv) \leq a\phi(u) + b\phi(v)$. Describe explicitly the probability measure $\mu$ and the random variable $f$ that you use.

9. Suppose that $f$ is in $\mathcal{L}^r$ for some $r$ with $1 \leq r < \infty$. (a) Show that the limit as $p \to \infty$ of $\|f\|_p$ is equal to $\|f\|_\infty$. Hint: Obtain an upper bound on $\|f\|_p$ in terms of $\|f\|_r$ and $\|f\|_\infty$. Obtain a lower bound on $\|f\|_p$ by using Chebyshev's inequality applied to the set $|f| > a$ for some $a$ with $0 < a < \infty$. Show that this set must have finite measure. For which $a$ does this set have strictly positive measure? (b) Show that the result is not true without the assumption that $f$ belongs to some $\mathcal{L}^r$.

10. Consider $1 \leq p < \infty$. Let $\mathcal{B}$ denote Borel measurable functions on the line. Consider Lebesgue measure $\lambda$ and the corresponding space $L^p(\mathbb{R}, \mathcal{B}, \lambda)$. If $f$ is in this $L^p$ space, the translate $f_a$ is defined by $f_a(x) = f(x-a)$. (a) Show that for each $f$ in $L^p$ the function $a \mapsto f_a$ is continuous from the real line to $L^p$. (b) Show that the corresponding result for $L^\infty$ is false. Hint: Take it as known that the space of step functions is dense in the space $L^p$ for $1 \leq p < \infty$.

11. Define the Fourier transform for $f$ in $L^1(\mathbb{R}, \mathcal{B}, \lambda)$ by

$$\hat{f}(k) = \int_{-\infty}^{\infty} e^{-ikx} f(x)\, dx. \tag{23.33}$$

Show that if $f$ is in $L^1$, then the Fourier transform is in $L^\infty$ and is continuous. Hint: Use the dominated convergence theorem.

12. Show that if $f$ is in $L^1$, then the Fourier transform of $f$ vanishes at infinity. Hint: Take it as known that the space of step functions is dense in the space $L^1$. Compute the Fourier transform of a step function.

13. Minkowski's inequality for integrals (a) Let $1 \leq p < \infty$. Show that the $L^p$ norm of the integral is bounded by the integral of the $L^p$ norm. More

specifically, let $\mu$ be a measure defined for functions on $X$, and let $\nu$ be a measure defined for functions on $Y$. Suppose that $\mu$ and $\nu$ are each $\sigma$-finite. Let $f$ be a product measurable function on $X \times Y$. Then $\nu(f \mid 1)$ denotes the $\nu$ partial integral of $f$ keeping the first variable fixed, and $\|f \mid 2\|_p$ is the $\mathcal{L}^p$ norm with respect to $\mu$ keeping the second variable fixed. The assertion is that

$$\|\nu(f \mid 1)\|_p \leq \nu(\|f \mid 2\|_p). \tag{23.34}$$

That is,

$$\left(\int |\int f(x,y)\, d\nu(y)|^p \, d\mu(x)\right)^{\frac{1}{p}} \leq \int \left(\int |f(x,y)|^p \, d\mu(x)\right)^{\frac{1}{p}} \, d\nu(y). \tag{23.35}$$

(b) What is the special case of this result when $\nu$ is a counting measure on two points? (c) What is the special case of this result when $\mu$ is a counting measure on two points? Hint: For the general inequality it is enough to give the proof when $f$ is a positive function. Write $\alpha(y) = (\int f(x,y)^p \, d\mu(x))^{\frac{1}{p}}$ and set $\alpha = \int \alpha(y)\, d\nu(y)$. Then

$$\left(\frac{1}{\alpha}\int f(x,y)\, d\nu(y)\right)^p = \left(\int \frac{f(x,y)}{\alpha(y)} \frac{\alpha(y)}{\alpha}\, d\nu(y)\right)^p \leq \int \left(\frac{f(x,y)}{\alpha(y)}\right)^p \frac{\alpha(y)}{\alpha}\, d\nu(y). \tag{23.36}$$

Apply the $\mu$ integral and interchange the order of integration.

14. Let $K$ be the Legendre transform of $H$. Thus $K(y) = xy - H(x)$, where $x$ is the solution of $y = H'(x)$. (a) Show that $K'(y) = x$, in other words, $K'$ is the inverse function to $H'$. (b) Show that if $H''(x) > 0$, then also $K''(y) > 0$. What is the relation between these two functions?

# Chapter 24

# Hilbert space

## 24.1 Inner products

A *Hilbert space* $H$ is a vector space with an inner product that is complete. The vector space can have real scalars, in which case the Hilbert space is a real Hilbert space. Or the vector space can have complex scalars; this is the cased of a complex Hilbert space. Both cases are useful. Real Hilbert spaces have a geometry that is easy to visualize, and they arise in applications. However complex Hilbert spaces are better in some contexts. In the following most of the attention will be given to complex Hilbert spaces.

An *inner product* is defined so that it is linear in one variable and conjugate linear in the other variable. The convention adopted here is that for vectors $u, v, w$ and complex scalars $a, b$ we have

$$\begin{aligned} \langle u, av + bw \rangle &= a\langle u, v\rangle + b\langle u, w\rangle \\ \langle au + bw, v\rangle &= \bar{a}\langle u, v\rangle + \bar{b}\langle w, v\rangle. \end{aligned} \tag{24.1}$$

Thus the inner product is conjugate linear in the left variable and linear in the right variable. This is the convention in physics, and it is also the convention in some treatments of elementary matrix algebra. However in advanced mathematics the opposite convention is common.

The inner product also satisfies the condition

$$\langle u, v\rangle = \overline{\langle v, u\rangle}. \tag{24.2}$$

Thus $\langle u, u\rangle$ is real. In fact, we require for an inner product that

$$\langle u, u\rangle \geq 0. \tag{24.3}$$

The final requirement is that

$$\langle u, u\rangle = 0 \Rightarrow u = 0. \tag{24.4}$$

The inner product defines a norm $\|u\| = \sqrt{\langle u, u\rangle}$. It has the basic homogeneity property that $\|au\| = |a|\|u\|$. The most fundamental norm identity is

$$\|u + v\|^2 = \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2. \tag{24.5}$$

Notice that the cross terms are real, and in fact $\langle u, v\rangle + \langle v, u\rangle = 2\Re\langle u, v\rangle$. This leads to the *Schwarz inequality.*

**Theorem 24.1 (Schwarz inequality)**

$$|\langle u, v\rangle| \leq \|u\|\|v\|. \tag{24.6}$$

Proof: If either $u$ or $v$ is the zero vector, then the inequality is trivial. Otherwise let $u_1 = u/\|u\|$ and $v_1 = e^{i\theta}v/\|v\|$. Then these are unit vectors. By the fundamental identity

$$-2\Re(e^{i\theta}\langle u_1, v_1\rangle) \leq 2. \tag{24.7}$$

Pick $\theta$ so that $-\Re(e^{i\theta}\langle u_1, v_1\rangle) = |\langle u_1, v_1\rangle)$. Then the equation gives

$$2|\langle u_1, v_1\rangle| \leq 2. \tag{24.8}$$

This leads immediately to the Schwarz inequality. □

**Theorem 24.2 (Triangle inquality)**

$$\|u + v\| \leq \|u\| + \|v\|. \tag{24.9}$$

Proof: From the fundamental identity and the Schwarz inequality

$$\|u + v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2. \tag{24.10}$$

□

Each element $u$ of $H$ defines an element $u^*$ of the dual space $H^*$ of continuous linear functions from $H$ to $\mathbb{C}$. The definition of $u$ is that it is the linear function $v \mapsto \langle u, v\rangle$. That is, the value

$$u^*(v) = \langle u, v\rangle. \tag{24.11}$$

From the Schwarz inequality we have

$$\|u^*(v)\| \leq \|u\|\|v\|, \tag{24.12}$$

which proves that $u^*$ is Lipschitz, hence continuous.

This notation shows the advantage of the convention that inner products are continuous in the right variable. In fact, this notation is rather close to one that is common in linear algebra. If $u$ is a column vector, then $u^*$ is given by the corresponding row vector with complex conjugate entries. The algebraic properties of this correspondence are just what one would want: $(u + w)^* = u^* + w^*$ and $(au)^* = \bar{a}u^*$.

The notation is also compatible with the standard notation for the adjoint of a linear transformation from one Hilbert space to another. We can identify the vector $u$ with the linear transformation $a \mapsto au$ from $\mathbb{C}$ to $H$. The adjoint of this transformation is the transformation $w \mapsto \langle u, w\rangle$ from $H$ to $\mathbb{C}$. This is just the identity $\overline{\langle u, w\rangle} a = \langle w, au\rangle$. So it is reasonable to denote the transformation $w \mapsto \langle u, w\rangle$ by the usual notation $u^*$ for adjoint.

One particularly convenient aspect of this notation is that one may also form the outer product $vu^*$. This is a linear transformation from $H$ to itself given by $w \mapsto v\langle u, w\rangle$. This following result is well known in the context of integral equations.

**Proposition 24.3** *The linear transformation $vu^*$ has eigenvalues $\langle u, v\rangle$ and 0. If $\mu$ is not equal to either of these two values, then the inverse if $\mu I - vu^*$ is*

$$(\mu I - vu^*)^{-1} = \frac{1}{\mu}\left[I + \frac{1}{\mu - \langle u, v\rangle} vu^*\right]. \tag{24.13}$$

If $\langle u, v\rangle = 1$, then $vu^*$ is a slant projection onto $v$ along the directions perpendicular to $u$. In particular, if $u^*u = 1$, then $uu^*$ is an orthogonal projection onto $u$.

In the physics literature a vector $v$ is called a ket. A dual vector $u^*$ is called a bra. The complex number that results from the pairing of a bra and a ket is $\langle u, v\rangle$, hence a bracket. This rather silly terminology is due to Dirac.

If $u$ and $v$ are vectors, we write $u \perp v$ if $\langle u, v\rangle = 0$. They are said to be perpendicular or orthogonal.

**Theorem 24.4 (Pythagoras)** *If $u \perp v$, then*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2. \tag{24.14}$$

The *theorem of Pythagoras* says that for a right triangle the sums of the squares of the sides is the square of the hypotenuse. It follows immediately from the fundamental identity. The following *parallelogram law* does not need the hypothesis of orthogonality.

**Theorem 24.5 (parallelogram law)** *In an an inner product space*

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2. \tag{24.15}$$

It says that for an arbitrary triangle there is an associated parallelogram consisting of $0, u, v, u + v$, and the sum of the squares of the diagonals is equal to the sum of the square of the four sides.

This parallelogram law is a fundamental convexity result. Write it in the form

$$\|\frac{u - v}{2}\|^2 + \|\frac{u + v}{2}\|^2 = \frac{1}{2}\|u\|^2 + \frac{1}{2}\|v\|^2. \tag{24.16}$$

This immediately shows that the function $u \to \|u\|^2$ has a convexity property, at least with weights 1/2 and 1/2. Furthermore, there is a kind of strict convexity, in that there is an extra term on the left that is strictly positive whenever $u \neq v$.

## 24.2 Closed subspaces

The classic examples of Hilbert spaces are sequence spaces $\mathbb{C}^n$ (finite dimensional) and $\ell^2$ (infinite dimensional). A more general class of infinite dimensional examples is given by $L^2(X, \mathcal{F}, \mu)$, the quotient space formed from $\mathcal{F}$ measurable functions that are square integrable with respect to $\mu$. In this case the inner product comes from

$$\langle f, g \rangle = \mu(\bar{f}g) = \int \overline{f(x)} g(x)\, d\mu(x). \tag{24.17}$$

The norm is

$$\|f\| = \sqrt{\mu(|f|^2)} = \sqrt{\int |f(x)|^2\, d\mu(x)}. \tag{24.18}$$

A *linear subspace* (or vector subspace) $M$ of a Hilbert space $H$ is a subset with the zero vector in it and that is closed under vector addition and scalar multiplication. If it is also a in the topological sense, then it is another Hilbert space.

**Theorem 24.6** *A subspace $M$ of a Hilbert space is a closed subspace if and only if it is itself a Hilbert space (with the same inner product).*

Proof: Suppose that $M$ is a closed subspace. Since it is a closed subset of a complete metric space, it is complete. Therefore $M$ is a Hilbert space.

Suppose on the other hand that $M$ is a Hilbert space. Since it is a complete subset of a metric space, it follows that $M$ is a closed subset. So $M$ is a closed subspace. □

Examples:

1. Consider the space $\ell^2$ of square summable sequences. Fix $n$. The subspace consisting of all sequences $x$ in this space with $x_k = 0$ for $k \leq n$ is a closed infinite dimensional subspace.

2. Consider the space $\ell^2$. The subspace consisting of all sequences $x$ such that there exists an $n$ such that $x_k = 0$ for $k > n$ is an infinite dimensional subspace that is not closed. In fact, its closure is all of $\ell^2$.

3. Consider the space $\ell^2$. Fix $n$. The subspace consisting of all sequences $x$ in this space with $x_k = 0$ for $k > n$ is a closed finite dimensional subspace.

4. Consider the space $L^2(\mathbb{R}, \mathcal{B}, \lambda)$ of Borel functions on the line that are square summable with respect to Lebesgue measure $\lambda$. The subspace consisting of all step functions is not closed. In fact, its closure is the entire Hilbert space.

5. Consider the space $L^2(\mathbb{R}, \mathcal{B}, \lambda)$ of Borel functions on the line that are square summable with respect to Lebesgue measure $\lambda$. The subspace consisting of all continuous functions with compact support is not closed. In fact, its closure is the entire Hilbert space. Notice that the subspaces in the last two examples have only the zero vector in common.

If $M$ is a subspace of a Hilbert space $H$, then we write $w \perp M$ if for all $v$ in $M$ we have $w \perp v$. Then $M^\perp$ consists of all vectors $w$ in $H$ such that $w \perp M$.

**Theorem 24.7** *If $M$ is a subspace, then $M^\perp$ is a closed subspace. Furthermore, $M \subset M^{\perp\perp}$.*

Notice that in general, it is not the case that $M$ is equal to $M^{\perp\perp}$. As an example, let $M$ be the subspace that is the intersection of the subspaces in examples 1 and 2 above. Then $M^\perp$ is the subspace in example 3. However $M^{\perp\perp}$ is the subspace in example 1.

## 24.3 The projection theorem

**Lemma 24.8** *Let $M$ be a subspace of $H$. Let $u$ be a vector in $H$. Then $v$ is a vector in $M$ that is closest to $u$ if and only if $v$ is in $M$ and $u-v$ is perpendicular to $M$.*

Proof: Suppose that $v$ is in $M$ and $u - v$ is perpendicular to $M$. Let $v'$ be another vector in $M$. Then $u - v' = u - v + v - v'$. By the theorem of Pythagoras $\|u - v'\|^2 = \|u - v\|^2 + \|v - v'\|^2$. Hence $\|u - v\| \leq \|u - v'\|$. So $v$ is the vector in $M$ closest to $u$.

Suppose on the other hand that $v$ is the vector in $M$ closest to $u$. Let $w \neq 0$ be another vector in $M$. Let $cw$ be the projection of $u - v$ onto the subspace generated by $w$ given by taking $c = \langle u-v, w\rangle/\langle w, w\rangle$. Let $v' = v+cw$. Then the difference $(u - v') - (v' - v) = u - v'$ is orthogonal to $v' - v$. By the theorem of Pythagoras $\|u-v\|^2 = \|u-v'\|^2 + \|v' - v\|^2$. Since $v$ is closet to $u$, the left hand side must be no larger than the first term on the right hand side. Hence the second term on the right hand side is zero. Thus $v' - v = cw$ is zero, and hence $c = 0$. We conclude that $u - v$ is perpendicular to $w$. Since $w$ is an arbitrary non-zero vector in $M$, this proves that $u - v$ is perpendicular to $M$. □

**Theorem 24.9 (Projection theorem)** *Let $M$ be a closed subspace of the Hilbert space $H$. Let $u$ be a vector in $H$. Then there exists a unique vector $v$ in $M$ that is closest to $u$. In particular, $u - v$ is perpendicular to $M$.*

Proof: Let $a$ be the infimum of the numbers $\|u - v\|^2$ for $v$ in $M$. Let $v_n$ be a sequence of vectors in $M$ such that $\|u - v_n\|^2 \to a$ as $n \to \infty$. Apply the parallelogram identity to two vectors $u - v_m$ and $u - v_n$. This gives

$$\|\frac{1}{2}(v_m - v_n)\|^2 + \|u - \frac{1}{2}(v_m + v_n)\|^2 = \frac{1}{2}\|u - v_m\|^2 + \frac{1}{2}\|u - v_n\|^2. \quad (24.19)$$

Since each $\frac{1}{2}(v_m + v_n)$ is in $M$, it follows that

$$\|\frac{1}{2}(v_m - v_n)\|^2 + a \leq \frac{1}{2}\|u - v_m\|^2 + \frac{1}{2}\|u - v_n\|^2. \tag{24.20}$$

As $m, n$ get large, the right hand side tends to $a$. This proves that the $v_n$ form a Cauchy sequence. Since $M$ is complete, there exists a $v$ in $M$ such that $v_n \to v$ as $n \to \infty$. Since the $\|u - v_n\|^2 \to a$ and the norm is a continuous function on $H$, it follows that $\|u - v\|^2 = a$. This proves that $v$ is the vector in $M$ closest to $u$. □

The vector $v$ in this *projection theorem* is the *orthogonal projection* (or projection) of $u$ on $M$. It is clear that the vector $u - v$ is the orthogonal projection of $u$ on $M^{+}$. Thus every vector in $H$ may be written as the sum of a vector in $M$ and a vector in $M^{\perp}$.

**Corollary 24.10** *If $M$ is a closed subspace, then $M^{\perp\perp} = M$.*

Proof: It is evident that $M \subset M^{\perp\perp}$. The hard part is to show that $M^{\perp\perp} \subset M$. Let $u$ be in $M^{\perp\perp}$. By the projection theorem we can write $u = w + v$, where $v$ is in $M$ and $w$ is in $M^{\perp}$. Then $0 = \langle u, w\rangle = \langle w, w\rangle + \langle v, w\rangle = \langle w, w\rangle$. So $w = 0$ and $u = v$ is in $M$. □

## 24.4 The Riesz-Fréchet theorem

**Theorem 24.11 (Riesz-Fréchet representation theorem)** *Let $H$ be a Hilbert space. Let $L$ be a continuous linear function $L : H \to \mathbb{C}$. Then there exists a unique vector $u$ in $H$ such that $L = u^*$. That is, the vector $u$ satisfies $L(u) = \langle u, v\rangle$ for all $v$ in $H$.*

Proof: Let

$$E[w] = \frac{1}{2}\|w\|^2 - \Re L(w). \tag{24.21}$$

Since $L$ is continuous, it is Lipschitz. That is, there is a constant $c$ such that $|Lw| \leq c\|w\|$. It follows that $E[w] \geq \frac{1}{2}\|w\|^2 - c\|w\| \geq -\frac{1}{2}c^2$. Thus the $E[w]$ are bounded below. Let $a$ be the infimum of the $E[w]$. Let $w_n$ be a sequence such that $E[w_n] \to a$ as $n \to \infty$. By the parallelogram identity and the fact that $L$ is linear we have

$$\|\frac{1}{2}(w_m - w_n)\|^2 + 2E[\frac{1}{2}(w_m + w_n)] = E[w_m] + E[w_n]. \tag{24.22}$$

Hence

$$\|\frac{1}{2}(w_m - w_n)\|^2 + 2a \leq E[w_m] + E[w_n]. \tag{24.23}$$

However, the right hand side tends to $2a$. This proves that the $w_n$ form a Cauchy sequence. Since $H$ is complete, they converges to a vector $u$ in $M$. By continuity $E[u] = a$.

The rest of the proof amounts to taking the derivative of the function $E[w]$ at $w = u$ in the direction $v$. The condition that this derivative is zero should give the result. However it is perhaps worth writing it out explicitly. Let $v$ be a vector in $H$. Then for each real number $t > 0$ we have

$$a = E[u] \leq E[u + tv]. \tag{24.24}$$

This says that

$$0 \leq t\Re\langle u, v\rangle - t\Re L(v) + \frac{1}{2}t^2\|v\|^2. \tag{24.25}$$

Divide both sides by $t > 0$. Since the resulting inequality is true for each $t > 0$, we have $0 \leq \Re\langle u, v\rangle - \Re L(v)$. Thus we have proved that $\Re L(v) \leq \Re\langle u, v\rangle$ for all $v$. The same argument applied to $-v$ shows that $-\Re L(v) \leq -\Re\langle u, v\rangle$. Hence $\Re L(v) = \Re\langle u, v\rangle$. The above reasoning applied to $-iv$ shows that $\Im L(v) = \Im\langle u, v\rangle$. We conclude that $L(v) = \langle u, v\rangle$. □

This *Riesz-Fréchet representation theorem* says that every element of the dual space $H^*$ comes from a vector $u$ in $H$. However it does not quite say that a Hilbert space is naturally isomorphic to its dual space, at least not in the case of complex scalars. In fact, the correspondence $u \mapsto u^*$ from $H$ to $H^*$ is conjugate linear. However it does preserve the norm. (In the real case it is completely accurate to say that a Hilbert space is naturally self-dual.)

Here is a warning. While a Hilbert space is naturally conjugate-isomorphic to its dual space, in general it is a mistake to identify a Hilbert space with its dual space. In some instances this does no harm, but in other contexts (such as the theory of Sobolev spaces) it can cause real confusion. Furthermore, for Banach spaces it is quite clear that there is no natural identification of the space with its dual space; indeed these spaces can be quite different.

## 24.5 Adjoint transformations

Let $T : H \to K$ be a continuous linear transformation from the Banach space $E$ to the Banach space $F$. In this context the value of $T$ on $u$ is often written in the form $Tu$. The *adjoint* transformation $T^* : K \to H$ is defined by $\langle T^*v, u\rangle = \langle v, Tu\rangle$. It is not difficult to prove that $T^{*}* = T$.

If we think of the space as consisting of column vectors, then $T$ is like a matrix on the left acting on column vectors on the right. The adjoint $T^*$ is the conjugate transpose matrix matrix, also acting this way.

The Lipschitz norm of the transformation $T$ is denoted $\|T\|$. Thus it is the smallest constant satisfying $\|Tu\| \leq \|T\|\|u\|$. The norm has the attractive property that $\|T^*\| = \|T\|$.

It may be shown that space of all continuous linear transformations from $H$ to $K$ is a Banach space. However in general it is not a Hilbert space.

Let $w$ be a vector in the Hilbert space $H$. Then it defines continuous linear transformation $a \mapsto aw$ from $\mathbb{C}$ to $H$. The adjoint transformation, denoted $w^*$, maps $H$ continuously to $\mathbb{C}$. In other words, it is in the dual space $H^*$. By definition it satisfies $\overline{w^*(v)}a = \langle v, aw\rangle$. This is equivalent to $w^*(v) = \langle w, v\rangle$. If we think of $v$ as a column vector, then $w^*$ is more like a conjugated row vector. The last equation says that the product of a conjugated row vector on the left with a column vector on the right gives the value of the inner product.

## 24.6 Bases

In the following we need the notion of isomorphism of Hilbert spaces. If $H$ and $K$ are each Hilbert spaces, then an *isomorphism* (or *unitary transformation*) from $H$ to $K$ is a bijection $U : H \to K$ that is linear and preserves the inner product. The latter means that $\langle Uv, Uw \rangle = \langle u, v \rangle$, where the inner product on the left is that of $K$ and the inner product on the right is that of $H$. An isomorphism of Hilbert spaces obviously also preserves the norm. As a consequence, an isomorphism of Hilbert spaces is simultaneously an isomorphism of vector spaces and an isomorphism of metric spaces. Clearly the inverse of an isomorphism is also an isomorphism.

Let $H$ be a Hilbert space. Let $J$ be an index set. An *orthonormal family* is a function $j \mapsto u_j$ from $J$ to $H$ such that $\langle u_j, u_k \rangle = \delta_{jk}$. The key result for an orthonormal family is the *Bessel inequality.*

**Proposition 24.12 (Bessel's inequality)** *Let $f$ be in $H$ and define coefficients $c$ by*

$$c_j = \langle u_j, f \rangle. \tag{24.26}$$

*Then $c$ is in $\ell^2(J)$ and*

$$\|c\|^2 \leq \|f\|^2. \tag{24.27}$$

Proof: Let $J_0$ be a finite subset of $J$. Let $g = \sum_{j \in J_0} c_j u_j$. Then $f - g \perp u_k$ for each $k$ in $J_0$. By the theorem of Pythagoras $\|f\|^2 = \|f - g\|^2 + \|g\|^2$. Hence $\|g\|^2 \leq \|f\|^2$. This says that $\sum_{j \in J_0} |c_j|^2 \leq \|f\|^2$. Since this is true for arbitrary finite $J_0 \subset J$, we have $\sum_{j \in J} |c_j|^2 \leq \|f\|^2$. □

The Bessel inequality and the completeness of Hilbert space lead to the following *Riesz-Fischer theorem.*

**Proposition 24.13 (Riesz-Fischer)** *Let $j \mapsto u_j$ be an orthonormal family. Let $c$ be in $\ell^2(J)$. Then the series*

$$g = \sum_j c_j u_j \tag{24.28}$$

*converges in the Hilbert space sense to a vector $g$ in $H$. Furthermore, we have the Parseval identity*

$$\|g\|^2 = \|c\|^2. \tag{24.29}$$

Proof: Since $c$ is in $\ell^2(J)$, there are only countably many values of $j$ such that $c_j \neq 0$. So we may consider this countable sum indexed by natural numbers. Let $g_n$ be the $n$th partial sum. Then for $m > n$ we have $\|g_m - g_n\|^2 = \sum_{j=n+1}^{m} |c_j|^2$. This approaches zero as $m, n \to \infty$. So the $g_n$ form a Cauchy sequence. Since

$H$ is complete, the $g_n$ converge to some $g$ in $H$. That is, $\|g - \sum_{j=0}^{n} c_j e_j\| \to 0$ as $n \to \infty$. This is the Hilbert space convergence indicated in the statement of the proposition. □

**Theorem 24.14** *Let $j \mapsto u_j$ be an orthonormal family of vectors in the Hilbert space $H$. Then there exists a closed subspace $M$ of $H$ such that the map $c \to \sum_j c_j u_j$ is an isomorphism from $\ell^2$ to $M$. The map that sends $f$ in $H$ to $g = \sum_j \langle u_j, f\rangle u_j$ in $H$ is the orthogonal projection of $f$ onto $M$.*

Sometimes physicists like to write the orthogonal projection onto $M$ with a notation that somewhat resembles

$$E = \sum_j u_j u_j^*. \tag{24.30}$$

This just means that the orthogonal projection of a vector $f$ onto the span of the $e_j$ is $Ef = \sum_j u_j \langle u_j, f\rangle$.

An orthonormal family $j \mapsto u_j$ is a *basis* for the Hilbert space $H$ if every vector $f$ in $H$ has the representation

$$f = \sum_j \langle u_j, f\rangle u_j. \tag{24.31}$$

In this case the correspondence between $H$ and $\ell^2(J)$ given by the basis is an isomorphism of Hilbert spaces. That is, the coefficient vectors $c$ with $c_j = \langle u_j, f\rangle$ give an alternative description of the Hilbert space.

**Proposition 24.15** *Suppose that $j \mapsto e_j$ is a maximal orthonormal family in $H$. Then it is a basis for $H$.*

Proof: Let $M$ be the collection of all linear combinations $\sum_j c_j u_j$ where $c$ is in $\ell^2$. Then $M$ is a closed subspace. If the family does not form a basis, then $M$ is a proper subset of $H$. Let $f$ be in $H$ and not in $M$. Let $g$ be the projection $f$ on $M$. Then $f - g$ is orthogonal to $M$ and is non-zero, and so can be normalized to be a unit vector. The gives a strictly larger orthonormal family, so the original family is not maximal. □

If $H$ is an arbitrary Hilbert space, then it follows from the axiom of choice via Zorn's lemma that there is a maximal orthonormal family $j \mapsto u_j$ defined on some index set $J$. This is a basis for $H$. In other words, every Hilbert space has a basis. It follows that for every Hilbert space there is a set $J$ such that $H$ is isomorphic to $\ell^2(J)$. This method of constructing bases involves lots of arbitrary choices and is not particularly practical. However it is of considerable theoretical interest: it says that Hilbert spaces of the same dimension (cardinality of index set) are isomorphic.

**Proposition 24.16** . *Suppose that $j \mapsto u_j$ is an orthonormal family such that for every vector $f$ in $H$ we have*

$$\|f\|^2 = \sum_j |\langle u_j, f\rangle|^2. \tag{24.32}$$

*Then $u \mapsto u_j$ is a basis for $H$.*

Proof: Let $M$ be the collection of all linear combinations $\sum_j c_j u_j$ where $c$ is in $\ell^2$. Then $M$ is a closed subspace. If the family does not form a basis, then $M$ is a proper subset of $H$. Let $f$ be in $H$ and not in $M$. Let $g$ be the projection $f$ on $M$. By the theorem of Pythagoras, $\|f\|^2 = \|g\|^2 + \|f - g\|^2 > \|g\|^2$. This violates the equality. □

**Proposition 24.17** *Suppose that $j \mapsto e_j$ is an orthonormal family, and suppose that the set of all finite linear combinations of these vectors is dense in $H$. Then it is a basis.*

Proof: Let $M$ be the collection of all linear combinations $\sum_j c_j u_j$ where $c$ is in $\ell^2$. Then $M$ is a closed subspace. Let $f$ be in $H$. Consider $\epsilon > 0$. Then there exists a finite linear combination $h$ such that $\|f - h\| < \epsilon$. Furthermore, $h$ is in $M$. Let $g$ be the projection of $f$ onto $M$. Since $g$ is the element of $M$ that is closest to $f$, it follows that $\|f - g\| \leq \|f - h\|$. Hence $\|f - g\| < \epsilon$. Since $\epsilon > 0$ is arbitrary, we have $f = g$. □

## 24.7 Separable Hilbert spaces

The following *Gram-Schmidt construction* gives a more explicit way of producing an orthonormal family.

**Theorem 24.18 (Gram-Schmidt orthonormalization)** *Let $k \mapsto v_k$ be a linearly independent sequence of vectors in the Hilbert space $H$. Then there exists an orthonormal sequence $k \mapsto u_k$ such that for each $m$ the span of $u_1, \ldots, u_m$ is equal to the span of $v_1, \ldots, v_m$.*

Proof: The proof is by induction on $m$. For $m = 0$ the sequences are empty, and the corresponding spans are both just the subspace with only the zero vector. Given $u_1, \ldots, u_m$ orthonormal with the same span as $v_1, \ldots, v_m$, let $q_{m+1}$ be the orthogonal projection of $v_{m+1}$ on this subspace. Let $p_{m+1} = v_{m+1} - q_{m+1}$. Then $p_{k+1}$ is orthogonal to $u_1, \ldots, u_m$. This vector is non-zero, since otherwise $v_{m+1}$ would depend linearly on the vectors in the subspace. Thus it is possible to define the unit vector $u_{m+1} = p_{m+1}/\|p_{m+1}\|$. □

**Theorem 24.19** *Let $H$ be a separable Hilbert space. Then $H$ has a countable basis.*

Proof: Since $H$ is separable, there is a sequence $s : \mathbf{N} \to H$ such that $n \mapsto s_n$ has dense range. In particular, the set of all finite linear combinations of the vectors $s_n$ is dense in $H$. Define a new subsequence $v_k$ (finite or infinite) by going through the $s_n$ in order and throwing out each element that is a linear combination of the preceding elements. Then the $v_k$ are linearly independent, and the linear span of the $v_k$ is the same as the linear span of the $s_n$. Thus we have a linearly independent sequence $t_k$ whose linear span is dense in $H$. Apply the Gram-Schmidt orthonormalization procedure. This gives an orthonormal sequence $u_k$ with the same linear span. Since the linear span of the $u_k$ is dense in $H$, the $u_k$ form a basis for $H$. □

**Corollary 24.20** *Every separable Hilbert space is isomorphic to some space $\ell^2(J)$, where $J$ is a countable set.*

For most applications separable Hilbert spaces are sufficient. In fact, a separable Hilbert space is either finite dimensional or has countable infinite dimension. All countable infinite dimensional separable Hilbert spaces are isomorphic. In fact, they are all isomorphic to $\ell^2(\mathbf{N})$.

The space $L^2([0,1],\mathcal{B},\lambda)$ of Borel functions on the unit interval with

$$\lambda(|f|^2) = \int_0^1 |f(x)|^2\,dx < +\infty \tag{24.33}$$

is a separable infinite dimensional Hilbert space. An example of an orthonormal basis is given by the Walsh functions.

Consider a natural number $n \geq 1$. Divide the interval from 0 to 1 in $2^n$ equally spaced parts, numbered from 0 to $2^n - 1$. The *Rademacher function* $r_n$ is the function that is 1 on the even numbered intervals and $-1$ on the odd numbered intervals. A Walsh function is a product of Rademacher functions. Let $S \subset \{1, 2, 3, \ldots\}$ be a finite set of strictly positive natural numbers. Let the *Walsh function* be defined by

$$w_S = \prod_{j\in S} r_j. \tag{24.34}$$

Notice that when $S$ is empty the product is 1.

The Walsh functions may be generated from the Rademacher functions in a systematic way. At stage zero start with the function 1. At stage one take also $r_1$. At stage two take $r_2$ times each of the functions from the previous stages. This gives also $r_2$ and $r_1r_2$. At stage three take $r_3$ times each of the functions from the previous stages. This gives also $r_3$ and $r_1r_3$ and $r_2r_3$ and $r_1r_2r_3$. It is clear how to continue. The Walsh functions generated in this way oscillate more and more.

**Theorem 24.21** *The Walsh functions form an orthonormal basis of $L^2([0,1],\mathcal{B},\lambda)$ with respect to the inner product*

$$\langle f, g\rangle = \lambda(\bar{f}g) = \int_0^1 \overline{f(x)}g(x)\,dx. \tag{24.35}$$

*Thus for an arbitrary function $f$ in $L^2([0,1])$ there is an $L^2$ convergent Walsh expansion*

$$f(x) = \sum_S \langle w_S, f\rangle w_S(x). \tag{24.36}$$

Proof: The $2^n$ Walsh functions $w_S$ with $S \subset \{1,\ldots,n\}$ are linearly independent. It follows that they span the $2^n$ dimensional space of binary step functions with step width $1/2^n$. Thus the linear span of the Walsh functions is the space of all binary step functions. These are dense in $L^2$. In fact, they are even uniformly dense in the space of continuous functions. □

Consider a natural number $n \geq 0$. Divide the interval from 0 to 1 in $2^n$ equally spaced parts, numbered from 0 to $2^n - 1$. The binary rectangular function $f_{n;k}$ is the function that is the indicator function of the $k$th interval.

A *Haar function* is a multiple of a product of a binary rectangular function with a Rademacher function. For $n \geq 0$ and $0 \leq k < 2^n$ define the Haar function to be

$$h_{n;k} = c_n f_{n;k} r_{n+1}, \tag{24.37}$$

and define $h_{-1;0} = 1$. For $n \geq 0$ the coefficient $c_n > 0$ is determined by $c_n^2 = 1/2^n$. The function $h_{-1;0}$ together with the other Haar functions $h_{j;k}$ for $j = 0$ to $n-1$ and $0 \leq k < 2^j$ form a basis for the binary step functions with width $1/2^n$. Note that the number of such functions is $1 + \sum_{j=0}^{n-1} 2^j = 2^n$.

The Haar functions may be generated in a systematic way. At stage zero start with the function 1. At stage one take also $r_1$. At stage two take also $f_{1;0}r_2$ and $f_{1;1}r_2$. At stage three take also $f_{2;0}r_3$ and $f_{2;1}r_3$ and $f_{2;2}r_3$ and $f_{2;3}r^3$. The Haar functions generated in this way become more and more concentrated in width.

**Theorem 24.22** *The Haar functions form an orthonormal family of vectors with respect to the inner product*

$$\langle f, g\rangle = \lambda(\bar{f}g) = \int_0^1 \overline{f(x)}g(x)\,dx. \tag{24.38}$$

*For an arbitrary function $f$ in $L^2([0,1])$ there is an $L^2$ convergent Haar expansion*

$$f(x) = \sum_{n=-1}^{\infty} \sum_{0 \leq k < 2^n} \langle h_{n;k}, f\rangle h_{n;k}(x). \tag{24.39}$$

Proof: The $2^n$ partial sum of the Haar series is the same as the $2^n$ partial sum of the Walsh series. Each of these is the projection onto the space of rectangular functions of width $1/2^n$. □

# Problems

1. This problem concerns real Borel measurable functions on the unit interval and Lebesgue measure. Consider the function $x^2$.

   (a) What is the constant function that best approximates $x^2$ in the $L^\infty$ norm?

   (b) What is the constant function that best approximates $x^2$ in the $L^2$ norm?

2. Consider the finite dimensional real Hilbert space $H = \mathbf{R}^n$ with Lebesgue measure $\lambda^n$. Say that a point in $\mathbf{R}^n$ is a "corner point" if each coordinate has absolute value one. Let $A$ be the open ball of radius 2 centered at the origin. Let $B$ be the union of the open balls of radius 1 centered at the corner points.

   (a) What is the ratio of the Lebesgue measure of $B$ to the Lebesgue measure of $A$?

   (b) For which $n$ are the sets $A$ and $B$ disjoint?

3. Consider the Hilbert space $\mathcal{H} = L^2(\mathbb{R}, \mathcal{B}, \lambda)$, where $\lambda(f) = \int_{-\infty}^{\infty} f(x)\,dx$ is Lebesgue measure. Consider the closed subspace $M$ consisting of all functions $g$ in $\mathcal{H}$ satisfying $g(-x) = -g(x)$. Find the orthogonal projection of the function $f(x) = (1+x)e^{-x}/(1+e^{-2x})$ onto this subspace.

4. Let
$$E = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z} \tag{24.40}$$
be the Euler operator. Let
$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{24.41}$$
be the Laplace operator. The space $P_\ell$ of *solid spherical harmonics* of degree $\ell$ consists of all polynomials $p$ in variables $x, y, z$ with $Ep = \ell p$ and $\Delta p = 0$. It may be shown that $P_\ell$ has dimension $2\ell+1$. (a) Prove that $P_\ell$ is invariant under rotations about the origin. (b) Find the spaces $P_\ell$ for $\ell = 0, 1, 2, 3$.

5. Let $S$ be the unit sphere in $\mathbb{R}^3$ defined by the equation $x^2 + y^2 + z^2 = 1$. The space $H_\ell$ of *surface spherical harmonics* consists of the restrictions of the solid spherical harmonics to $S$. Show that $P_\ell$ and $H_\ell$ are isomorphic vector spaces. Give an explicit formula that expresses a solid spherical harmonic in terms of the corresponding surface spherical harmonic.

6. Let $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$ and $z = r\cos\theta$ be expressed in spherical polar coordinates. Then $E = r\partial/\partial r$. Furthermore,
$$\Delta = \frac{1}{r^2}\left(E(E+1) + \Delta_S\right), \tag{24.42}$$
where
$$\Delta_S = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \tag{24.43}$$
Consider the space $L^2(S)$ with the rotation invariant measure $\sin\theta\,d\theta\,d\phi$, where the co-latitude $\theta$ goes from 0 to $\pi$ and the longitude $\phi$ goes from 0 to $2\pi$. (a) Show that for smooth functions in $L^2(S)$ we have the identity $\langle\Delta_S u, v\rangle = \langle u, \Delta_S v\rangle$. When you integrate by parts, be explicit about boundary terms. (b) Show that for $u$ in $H_\ell$ we have $\Delta_S u = -\ell(\ell+1)u$.

7. Show that the subspaces $H_\ell$ of surface spherical harmonics are orthogonal in $L^2(S)$ for different values of $\ell$.

8. Consider the Hilbert space $L^2(\mathbb{R}, \mathcal{B}, \gamma)$, where $\gamma$ is the measure $\gamma(h) = \int_{-\infty}^{\infty} h(x) \exp(-x^2)\, dx$. Thus the inner product is

$$\langle f, g\rangle = \gamma(\bar{f} g) = \int_{-\infty}^{\infty} \overline{f(x)} g(x) e^{-x^2}\, dx. \tag{24.44}$$

Let the polynomial $h_n(x)$ of degree $n$ be defined by

$$h_n(x) = \left(2x - \frac{d}{dx}\right)^n 1. \tag{24.45}$$

Thus $h_0(x) = 1$, $h_1(x) = 2x$, $h_2(x) = 4x^2-2$, and so on. Find $h_3(x), h_4(x), h_5(x)$. Show that the $h_n$ form an orthogonal family of vectors in the Hilbert space. Hint: Integrate by parts.

9. Show that the inner product in a Hilbert space is determined by its norm. Prove that a Banach space isomorphism between Hilbert spaces is actually a Hilbert space isomorphism. (That is, prove that if there is a linear bijective correspondence that preserves the norm, then it preserves the inner product.) Hint: For a real Hilbert space this follows from $4\langle u, v\rangle = \|u + v\|^2 - \|u - v\|^2$. How about for a complex Hilbert space?

10. Let $j \to u_j$ be an orthogonal family of non-zero vectors in $H$ indexed by $j \in J$. Suppose that finite linear combinations of these vectors are dense in $H$. Let $w_j = 1/\langle u_j, u_j\rangle$ be numerical weights associated with these vectors. Let $\ell^2(J, w)$ be the Hilbert space of all sequences $c$ with $\sum_j |c_j|^2 w_j < \infty$. Show that $H$ is isomorphic to $\ell^2(J, w)$ by the isomorphism that sends $f$ to $c$ given by $c_j = \langle u_j, f\rangle$. Find the inverse isomorphism. Check that both isomorphisms preserve the norm.

11. This is a continuation. Let $T = (-\frac{L}{2}, \frac{L}{2}]$ be the circle of length $L$, and consider the Hilbert space $H = L^2(T, \mathcal{B}, \lambda)$ with norm squared equal to

$$\|f\|^2 = \int_T |f(x)|^2\, dx. \tag{24.46}$$

Let $k$ be an integer multiple of $\frac{2\pi}{L}$. Let $u_k$ be the element of $H$ defined by $u_k(x) = \exp(ikx)$. According to the theory of Fourier series the finite linear combinations of these vectors are dense in $H$. These vectors are orthogonal but not orthonormal. Do not normalize them! Instead, find the weights $w_j = 1/\langle u_j, u_j\rangle$. (a) Write explicitly the formula for the coefficients $c$ in the space $\ell^2(\frac{2\pi}{L}\mathbf{Z}, w)$ of a function $f$ in $H$. (b) Write explicitly the formula for $f$ in terms of the coefficients $c$. (c) Write explicitly the equation that expresses the equality of norms (squared) in the two Hilbert spaces.

12. This is a continuation. Fix $f$ smooth with compact support. Let $L \to \infty$ with $k$ fixed. What are the limiting formulas corresponding to (a),(b),(c) above. Why was it important not to normalize the vectors?

# Chapter 25

# Differentiation

## 25.1 The Lebesgue decomposition

Consider a measurable space $X, \mathcal{F}$. Here $\mathcal{F}$ can stand for a $\sigma$-algebra of subsets, or for the corresponding $\sigma$-algebra of real measurable functions. Let $\nu$ be a measure. As usual we write the integral of a measurable function $f \geq 0$ as $\nu(f) \geq 0$. We also write the measure of a measurable subset $E$ as $\nu(E) = \nu(1_E)$.

Consider two such measures $\nu$ and $\mu$. Then $\nu$ is said to be *absolutely continuous* with respect to $\mu$ if every measurable subset $E$ with $\mu(E) = 0$ also has $\nu(E) = 0$. In that case we write $\nu \prec \nu$.

Consider two such measures $\nu$ and $\mu$. Then $\nu$ and $\mu$ are said to be *mutually singular* if there exists a measurable set $A$ with complement $A^c = X \setminus A$ such that $\mu(A) = 0$ and $\nu(A^c) = 0$. In that case we write $\nu \perp \mu$. If $\mu$ is thought of as a reference measure, then sometimes we say that $\nu$ is *singular* with respect to $\mu$. The *Lebesgue decomposition* describes the relation between two finite measures.

**Theorem 25.1 (Lebesgue decomposition)** *Let $X, \mathcal{F}$ be a measurable space. Let $\mu$ and $\nu$ be finite measures. Then $\nu = \nu_{ac} + \nu_s$, where $\nu_{ac} \prec \mu$ and $\nu_s \perp \mu$.*

Proof: This Hilbert space proof is due to von Neumann. The trick is to compare $\nu$ not directly with $\mu$ but with $\mu + \nu$ instead. In fact, it is also possible to compare $\mu$ to $\mu + \nu$ in the same way.

The main technical device is to look at $\nu$ as a linear functional on the real Hilbert space $L^2(X, \mathcal{F}, \mu + \nu)$. We have

$$|\nu(f)| \leq \nu(|f|) \leq (\mu + \nu)(|f|) = \langle 1, |f| \rangle. \tag{25.1}$$

By this Schwarz inequality and the fact that $\mu + \nu$ is a finite measure we have the continuity

$$|\nu(f)| \leq \|1\|_2 \|f\|_2. \tag{25.2}$$

It follows from the Riesz-Fréchet theorem that there exists a function $g$ in $L^2(X, \mathcal{F}, \mu + \nu)$ so that $\nu$ is given by the inner product by $\nu(f) = \langle g, f \rangle$. In other words

$$\nu(f) = (\mu + \nu)(gf). \tag{25.3}$$

A little algebra gives

$$\mu(f) = (\mu + \nu)((1 - g)f). \tag{25.4}$$

Let $E$ be the set where $g < 0$. If $(\mu+\nu)(E) > 0$, then $\nu(E) = (\mu+\nu)(g1_E) < 0$, which is a contradiction. Thus $(\mu + \nu)(E) = 0$ and so $0 \leq g$ almost everywhere with respect to $\mu + \nu$. Let $F$ be the set where $g > 1$. If $(\mu + \nu)(F) > 0$, then $\mu(F) = (\mu+\nu)((1-g)1_F) < 0$, which is a contradiction. So also $(\mu+\nu)(F) = 0$. So we may as well assume that $0 \leq g \leq 1$.

One consequence of this result is that the last two displayed equations hold for all measurable functions $f \geq 0$, by the monotone convergence theorem.

Let $A$ be the set where $g = 1$. Let $A^c = X \setminus A$ be its complement, so $0 \leq g < 1$ on $A^c$. Let $\nu_s(f) = \nu(f1_A)$ and let $\nu_{ac}(f) = \nu(f1_{A^c})$. Since $\mu(A) = 0$ and $\nu_s(A^c) = 0$, we have $\nu_s \perp \mu$. If $\mu(E) = 0$, let $E' = E \cap A^c$. Then $\mu(E') = 0$ and so $(\mu + \nu)((1 - g)1_{E'}) = 0$. Since $1 - g > 0$ on $E'$, we have $(\mu + \nu)(E') = 0$ and so $\nu(E') = 0$. Thus $\nu_{ac}(E) = 0$. This proves that $\nu_{ac} \prec \mu$. □

The Lebesgue decomposition also holds true when $\nu$ and $\mu$ are $\sigma$-finite measures. It also holds true in the context when $\nu$ is a finite signed measure (a difference of two finite measures) and $\mu$ is a $\sigma$-finite measure.

## 25.2 The Radon-Nikodym theorem

Consider a measure $\mu$, regarded as a reference measure. In many cases $\mu$ is Lebesgue measure on some measurable subset of Euclidean space. Let $\nu$ be another measure. Suppose that $h \geq 0$ is in $\mathcal{L}^1$ with respect to the measure $\mu$, and that $\nu(f) = \mu(hf)$ for for all bounded measurable functions $f$. Then $\nu$ is said to be given by the *density* $h$. Actually, this is a relative density, since it depends also on the reference measure $\mu$. The existence of a density is characterized by the *Radon-Nikodym theorem.*

**Theorem 25.2 (Radon-Nikodym)** *Let $X, \mathcal{F}$ be a measurable space. Let $\mu$ and $\nu$ be finite measures. Suppose $\nu \prec \mu$. Then there exists a relative density $h \geq 0$ with $\mu(h) < \infty$ and with $\nu(f) = \mu(hf)$ for all $f \geq 0$.*

Proof: This is the von Neumann method of proof again. Suppose that $\nu \prec \mu$. Then from the Hilbert space argument above we find $g$ with $0 \leq g \leq 1$ such that $\nu(u) = (\mu+\nu)(gu)$ and $\mu(u) = (\mu+\nu)((1-g)u)$. These identities are valid for all measurable functions $u \geq 0$. Let $A$ be the set where $g = 1$. Then $\mu(A) = (\mu + \nu)((1 - g)1_A) = 0$. Hence by absolute continuity $\nu(A) = 0$. Let $h = g/(1 - g)$ on $A^c$ and $h = 0$ on $A$. Thus

$$(1 - g)h = g1_{A^c}. \tag{25.5}$$

This equation is crucial: it shows that $h$ is related to $g$ in a quite non-linear way. All that we know about $h$ is that $0 \leq h < +\infty$.

Let $f \geq 0$. Then we use the first identity with $u = f1_{A_c}$ and the second identity with $u = hf$ to get

$$\nu(f) = \nu(f1_{A^c}) = (\mu + \nu)(g1_{A^c}f) = (\mu + \nu)((1 - g)hf) = \mu(hf). \qquad (25.6)$$

It follows that $\mu(h) = \nu(1) < +\infty$, so $h$ is integrable. □

The Radon-Nikodym theorem is also true when $\nu$ is a finite measure and $\mu$ is a $\sigma$-finite measure. It also holds true in the more general context when $\nu$ is a finite signed measure and $\mu$ is a $\sigma$-finite measure.

The density function $h$ is called the *Radon-Nikodym derivative* of $\nu$ with respect to the reference measure $\mu$. Some justification for this terminology may be found in the problems.

## 25.3 Absolutely continuous functions

**Proposition 25.3** *Let $\nu$ be a finite measure and let $\mu$ be another measure. Then $\nu \prec \mu$ is equivalent to the condition that for every $\epsilon > 0$ there is a $\delta > 0$ such that for every measurable subset $E$ we have $\mu(E) < \delta \Rightarrow \nu(E) < \epsilon$.*

Proof: Suppose that $\nu \prec \mu$. Suppose there exists $\epsilon > 0$ such that for every $\delta > 0$ there is a measurable subset $E$ such that $\mu(E) < \delta$ and $\nu(E) \geq \epsilon$. Consider such an $\epsilon$. For each $n$ choose a measurable subset $E_n$ such that $\mu(E_n) < 1/2^{n+1}$ and $\nu(E_n) \geq \epsilon$. Let $F_k = \bigcup_{n=k}^{\infty} E_n$. Then $\mu(F_k) \leq 1/2^k$. Let $F = \bigcap_{k=1}^{\infty} F_k$. Since for each $k$ we have $F \subset F_k$, we have $\mu(F) \leq \mu(F_k) \leq 1/2^k$. Thus $\mu(F) = 0$. On the other hand, $E_k \subset F_k$, so $\nu(F_k) \geq \nu(E_k) \geq \epsilon$. Since $F_k \downarrow F$ and $\nu$ is a finite measure, it must be that $\nu(F_k) \downarrow \nu(F)$. Hence $\nu(F) \geq \epsilon$. The existence of $F$ with $\mu(F) = 0$ and $\nu(F) \geq \epsilon$ implies that $\nu \prec \mu$ is false. This a contradiction. Thus the $\epsilon - \delta$ condition holds. Thus the implication follows.

The converse is considerably easier. Suppose that the $\epsilon - \delta$ condition is satisfied. Suppose $\mu(E) = 0$. Let $\epsilon > 0$. It follows from the condition that $\nu(E) < \epsilon$. Since $\epsilon > 0$ is arbitrary, we have $\nu(E) = 0$. This is enough to show that $\nu \prec \mu$. □

Consider the real line $\mathbb{R}$ with the notion $\mathcal{B}$ of Borel set and Borel function. Let $F$ be an increasing right continuous real function on $\mathbb{R}$. Then there is a unique measure $\nu_F$ with the property that $\nu_F((a, b]) = F(b) - F(a)$ for all $a < b$. This measure is finite on compact sets. The measure determines the function up to an additive constant. Clearly $\nu_F$ is a finite measure precisely when $F$ is bounded.

One other fact that we need is that the measure of a Borel measurable set is determined from the function $F$ by a two-stage process. The first stage is to extend the measure from intervals $(a, b]$ to countable unions of such intervals. The second stage is to approximate an arbitrary measurable subset from outside by such countable unions.

It is not very difficult to show that the same result may be obtained by using open intervals $(a, b)$ instead of half-open intervals $(a, b]$. The most general open subset $U$ of the line is a countable union of open intervals. So the second stage of the approximation process gives the condition of outer regularity:

$$\nu_F(E) = \inf\{\nu_F(U) \mid U \text{ open }, E \subset U\}. \tag{25.7}$$

An increasing function $F$ is said to be absolutely continuous increasing if for every $\epsilon > 0$ there is $\delta > 0$ such that whenever $V$ is a finite union of disjoint open intervals with total length $\lambda(V) < \delta$ the corresponding sum of increments of $F$ is $< \epsilon$.

A Lipschitz increasing function from the real line to itself is an absolutely continuous increasing function. The converse is false.

An absolutely continuous increasing function is a uniformly continuous function from the real line to itself. However again the converse is false: not every uniformly continuous increasing function from the line to itself is absolutely continuous. The Cantor function provides an example.

**Theorem 25.4** *Suppose $F$ is bounded, so $\nu_F$ is finite. The measure $\nu_F \prec \lambda$ if and only if $F$ is an absolutely continuous increasing function.*

Proof: Suppose that $\nu_F \prec \lambda$. Then the fact that $F$ is absolutely continuous increasing follows from the proposition above.

Suppose on the other hand that $F$ is an absolutely continuous increasing function. Consider $\epsilon > 0$. Choose $\epsilon' < \epsilon$ with $\epsilon' > 0$. Then there exists a $\delta > 0$ such that whenever $U$ is a finite union of disjoint open sets, then the sum of the corresponding increases of $F$ is $< \epsilon'$. Suppose that $E$ is a Borel measurable subset with $\lambda(E) < \delta$. Since $\lambda$ is outer regular, there exists an open set $U$ with $E \subset U$ and $\lambda(U) < \delta$. There is a sequence $U_k$ of finite disjoint unions of $k$ open intervals such that $U_k \uparrow U$ as $k \to \infty$. Since $\lambda(U_k) < \delta$, it follows that the sum of the increases $\nu_F(U_k) < \epsilon'$. However the sequence $\nu_F(U_k) \uparrow \nu_F(U)$ as $k \to \infty$. Hence $\nu_F(U) \leq \epsilon'$. Hence $\nu_F(E) \leq \epsilon' < \epsilon$. This establishes the $\epsilon - \delta$ condition that is equivalent to absolute continuity. $\square$

Suppose that $F$ is an absolutely continuous increasing function that is bounded. Then the corresponding finite measure $\nu_F$ is absolutely continuous with respect to Lebesgue measure, and hence there is a measurable function $h \geq 0$ with finite integral such that

$$\nu_F(f) = \lambda(hf). \tag{25.8}$$

Explicitly, this says that

$$\int_{\infty}^{\infty} f(x)\, dF(x) = \int_{-\infty}^{\infty} f(x)h(x)\, dx. \tag{25.9}$$

Take $f(x)$ to be the indicator function of the interval from $a$ to $b$. Then we obtain

$$F(b) - F(a) = \int_{-a}^{b} h(x)\, dx. \tag{25.10}$$

So the absolutely continuous increasing functions are precisely those functions that can be written as indefinite integrals of positive functions.

There is a more general concept of *absolutely continuous function* that corresponds to a signed measure that is absolutely continuous with respect to Lebesgue measure. These absolutely continuous functions are the indefinite integrals of integrable functions. See for instance Folland [5] for a detailed discussion of this important topic.

It is not true that the derivative of an absolutely continuous function exists at every point. However a famous theorem of Lebesgue says that it exists at almost every point and that the function can be recovered from its derivative by integration.

# Problems

1. Let $\lambda^2$ be Lebesgue measure on the square $[0,1] \times [0,1]$. Let $g \geq 0$ be integrable with respect to $\lambda^2$. Define a measure $\mu$ on the interval $[0,1]$ by
$$\mu(f) = \lambda^2(gf) = \int_0^1 \int_0^1 g(x,y) f(x)\, dx\, dy. \tag{25.11}$$
Find the function $h(x)$ that is the Radon-Nikodym derivative of $\mu$ with respect to the Lebesgue measure $\lambda$ on $[0,1]$.

2. Let $\lambda$ denote Lebesgue measure on the Borel subsets of the closed interval $[-1,1]$.

   (a) Let $\phi : [-1,1] \to [-1,1]$ be defined by $\phi(x) = x^2$. Find the image measure $\mu = \phi[\lambda]$.

   (b) Is $\mu$ absolutely continuous with respect to $\lambda$? Prove or disprove. If so, find its Radon-Nikodym derivative.

   (c) Is $\lambda$ absolutely continuous with respect to $\mu$? Prove or disprove. If so, find its Radon-Nikodym derivative.

   (d) Let $\psi : [-1,1] \to [-1,1]$ be defined by $\psi(x) = \text{sign}(x)$. (Here $\text{sign}(x) = x/|x|$ for $x \neq 0$ and $\text{sign}(0) = 0$.) Find the image measure $\nu = \psi[\lambda]$.

   (e) Is $\nu$ absolutely continuous with respect to $\lambda$? Prove or disprove. If so, find its Radon-Nikodym derivative.

   (f) Is $\lambda$ absolutely continuous with respect to $\nu$? Prove or disprove. If so, find its Radon-Nikodym derivative.

3. Consider real Borel functions $f$ on the interval $[0,2]$. Define the integral $\mu$ by
$$\mu(f) = \int_0^1 \int_0^1 f(x+y)\, dx\, dy. \tag{25.12}$$
Find the Radon-Nikodym derivative of $\mu$ with respect to Lebesgue measure on $[0,2]$.

4. Say that $\mu$ is a finite measure and $h \geq 0$ is a measurable function. Find the function $g$ that minimizes the quantity $\frac{1}{2}\mu((1+h)g^2) - \mu(hg)$.

5. Show that if $\nu \prec \mu$, then the derivative of $\nu$ with respect to $\mu$ is $h$. That is, show that if there is no division by zero, then

$$\lim_{\epsilon \downarrow 0} \frac{\nu(t \leq h < t + \epsilon)}{\mu(t \leq h < t + \epsilon)} = t. \tag{25.13}$$

Hint: Prove in fact the bounds

$$t \leq \frac{\nu(t \leq h < t + \epsilon)}{\mu(t \leq h < t + \epsilon)} \leq t + \epsilon. \tag{25.14}$$

# Chapter 26

# Conditional Expectation

## 26.1 Hilbert space ideas in probability

Consider a probability space $\Omega, \mathcal{S}, \mu$. Here $\mu$ will denote the expectation or *mean* defined for $\mathcal{S}$ measurable real functions. In particular $\mu(1) = 1$.

Recall that the set $\Omega$ is the set of outcomes of an experiment. A real measurable function $f$ on $\Omega$ is called a random variable, since it is a real number that depends on the outcome of the experiment. If $\omega \in \Omega$ is an outcome, then $f(\omega)$ is the corresponding experimental number.

A measurable subset $A \subset \Omega$ is called an event. The probability of the event $A$ is written $\mu(A)$. If $\omega \in \Omega$ is an outcome, then the event $A$ happens when $\omega \in A$.

A real measurable function $f$ is in $L^1(\Omega, \mathcal{S}, \mu)$ if $\mu(|f|) < +\infty$. In this case the function is called a random variable with finite first moment. The expectation $\mu(f)$ is a well-defined real number. The random variable $f - \mu(f)$ is called the centered version of $f$.

A real measurable function $f$ is in $L^2(\Omega, \mathcal{S}, \mu)$ if $\mu(f^2) < +\infty$. It this case it is called a random variable with finite second moment, or with finite variance. The second moment is $\mu(f^2)$. The *variance* is the second moment of the centered version. In the Hilbert space language this is

$$\mu((f - \mu(f))^2) = \|f - \mu(f)\|^2. \tag{26.1}$$

This equation may be thought of in terms of projections. The projection of $f$ onto the constant functions is $\mu(f)$. Thus the variance is the square of the length of the projection of $f$ onto the orthogonal complement of the constant functions. It is a quantity that tells how non-constant the function is.

In probability a common notion for variance is

$$\mathrm{Var}(f) = \mu((f - \mu(f))^2). \tag{26.2}$$

As mentioned before, this is squared length of the component orthogonal to the constant functions. There is a corresponding notion of *covariance*

$$\mathrm{Cov}(f,g) = \mu((f - \mu(f))(g - \mu(g))). \tag{26.3}$$

This is the inner product of the components orthogonal to the constant functions. Clearly $\mathrm{Var}(f) = \mathrm{Cov}(f, f)$.

Another quantity encountered in probability is the *correlation*

$$\rho(f,g) = \frac{\mathrm{Cov}(f,g)}{\sqrt{\mathrm{Var}(f)}\sqrt{\mathrm{Var}(g)}}. \tag{26.4}$$

The Hilbert space interpretation of this is the cosine of the angle between the vectors (in the subspace orthogonal to constants). This explains why $-1 \leq \rho(f,g) \leq 1$.

In statistics there are similar formulas for quantities like mean, variance, covariance, and correlation. Consider, for instance, a sample vector $f$ of $n$ experimental numbers. Construct a probability model where each index has probability $1/n$. This is called the empirical distribution. Then $f$ is a random variable, and so its mean and variance can be computed in the usual way. These are called the sample mean and sample variance. Or consider instead a sample of $n$ ordered pairs. This can be regarded as an ordered pair $f, g$, where $f$ and $g$ are each a vector of $n$ experimental numbers. Then $f$ and $g$ are each random variables with respect to the empirical distribution on the $n$ index points, and the covariance and correlation is computed as before. These are called the sample covariance and sample correlation. (Warning: Statisticians often use a slightly different definition for the sample variance or sample covariance, in which they divide by $n - 1$ instead of $n$. This does not matter for the sample correlation.)

The simplest (and perhaps most useful) case of the *weak law of large numbers* is pure Hilbert space theory. It says that averaging $n$ uncorrelated random variables makes the variance get small at the rate $1/n$.

**Proposition 26.1 (Weak law of large numbers)** *Let $f_1, \ldots, f_n$ be random variables with means $\mu(f_j) = m$ and covariances $\mathrm{Cov}(f_j, f_k) = \sigma^2\delta_{jk}$. Then their average (sample mean) satisfies*

$$\mu\left(\frac{f_1 + \cdots + f_n}{n}\right) = m \tag{26.5}$$

*and*

$$\mathrm{Var}\left(\frac{f_1 + \cdots + f_n}{n}\right) = \frac{\sigma^2}{n}. \tag{26.6}$$

## 26.2 Elementary notions of conditional expectation

In probability there is an elementary notion of conditional expectation given an event $B$ with probability $\mu(B) > 0$. It is

$$\mu(f \mid B) = \frac{\mu(f 1_B)}{\mu(B)}. \tag{26.7}$$

This defines a new expectation, corresponding to a world in which it is known that the event $B$ has happened. There is also the special case of conditional probability

$$\mu(A \mid B) = \frac{\mu(A \cap B)}{\mu(B)}. \tag{26.8}$$

Even these elementary notions can be confusing. Here is a famous problem, a variant of the shell game.

> Suppose you're on a game show and you're given the choice of three doors. Behind one is a car, behind each of the others is a goat. You pick a door, say door a, and the host, who knows what's behind the other doors, opens another door, say b, which has a goat. He then says : "Do you want to switch to door c?" Is it to your advantage to take the switch?'

Here is a simple probability model for the game show. Let $X$ be the door with the car. Then $P[X = a] = P[X = b] = P[X = c] = 1/3$. Suppose the contestant always initially chooses door $a$.

Solution 1: The host always opens door $b$. Then we are looking at conditional probabilities given $X \neq b$. Then $P[X = a \mid X \neq b] = (1/3)/(2/3) = 1/2$ and $P[X = c \mid X \neq b] = (1/3)/(2/3) = 1/2$. There is no advantage to switching. However this careless reading of the problem overlooks the hint that the host knows $X$.

Solution 2. The host always opens a door without a car. The door he opens is $g(X)$, where $g(b) = c$ and $g(c) = b$ and where for definiteness $g(a) = b$. Let $f$ be defined by $f(b) = c$ and $f(c) = b$. Then the contestant can choose $a$ or can switch and choose $f(g(X))$. There is no need to condition on $g(X) \neq X$, since it is automatically satisfied. The probabilities are then $P[X = a] = 1/3$ and $P[X = f(g(X))] = 2/3$. It pays to switch. This is the solution that surprised so many people.

## 26.3 The $L^2$ theory of conditional expectation

The idea of conditional expectation is that there is a smaller $\sigma$-algebra of measurable functions $\mathcal{F}$ with random variables that convey partial information about the result of the experiment.

For instance, suppose that $g$ is a random variable that may be regarded as already measured. Then every function $\phi(g)$ is computable from $g$, so one may think of $\phi(g)$ as measured. The $\sigma$-algebra of functions $\sigma(g)$ generated by $g$ consists of all $\phi(g)$, where $\phi$ is a Borel function.

Given a $\sigma$-algebra $\mathcal{F} \subset \mathcal{S}$, we have the closed subspace

$$L^2(\Omega, \mathcal{F}, \mu) \subset L^2(\Omega, \mathcal{S}, \mu). \tag{26.9}$$

Suppose $f$ is in $L^2(\Omega, \mathcal{S}, \mu)$. The *conditional expectation* $\mu(f \mid \mathcal{F})$ is defined to be the orthogonal projection of $f$ onto the closed subspace $L^2(\Omega, \mathcal{F}, \mu)$.

The conditional expectation satisfies the usual properties of orthogonal projection. Thus $\mu(f \mid \mathcal{F})$ is a random variable in $\mathcal{F}$, and $f - \mu(f \mid \mathcal{F})$ is orthogonal to $L^2(\Omega, \mathcal{F}, \mu)$. This says that for all $g$ in $L^2(\Omega, \mathcal{F}, \mu)$ we have $\langle \mu(f \mid \mathcal{F}), g\rangle = \langle f, g\rangle$, that is,

$$\mu(\mu(f \mid \mathcal{F})g) = \mu(fg). \tag{26.10}$$

If we take $g = 1$, then we get the important equation

$$\mu(\mu(f \mid \mathcal{F})) = \mu(f). \tag{26.11}$$

This says that we can compute the expectation $\mu(f)$ in two stages: first compute the conditional expectation random variable $\mu(f \mid \mathcal{F})$, then compute its expectation. In other words, work out the prediction based on the first stage of the experiment, then use these results to compute the prediction for the total experiment.

**Proposition 26.2** *The conditional expectation is order-preserving. If $f \leq g$, then $\mu(f \mid \mathcal{F}) \leq \mu(g \mid \mathcal{F})$.*

Proof: First we prove that if $h \geq 0$, then $\mu(h \mid \mathcal{F}) \geq 0$. Consider $h \geq 0$. Let $E$ be the set where $\mu(h \mid \mathcal{F}) < 0$. Then $1_E$ is in $\mathcal{F}$, so $\mu(\mu(f \mid \mathcal{F})1_E) = \mu(f1_E) \geq 0$. This can only happen if $\mu(f \mid \mathcal{F}) = 0$ almost everywhere on $E$. We can then apply this to $h = g - f$. □

**Corollary 26.3** *The random variables $\mu(f \mid \mathcal{F})$ and $\mu(|f| \mid \mathcal{F})$ satisfy $|\mu(f \mid \mathcal{F})| \leq \mu(|f| \mid \mathcal{F})$.*

Proof: Since $\pm f \leq |f|$, we have $\pm\mu(f \mid \mathcal{F}) \leq \mu(|f| \mid \mathcal{F})$. □

**Corollary 26.4** *The expectations satisfy $\mu(|\mu(f \mid \mathcal{F})|) \leq \mu(|f|)$.*

Here is the easiest example of a conditional expectation. Suppose that there is a partition of $\Omega$ into a countable family of disjoint measurable sets $B_j$ with union $\Omega$. Suppose that the probability of each $B_j$ is strictly positive, that is, $\mu(B_j) > 0$. Let $\mathcal{B}$ be the $\sigma$-algebra of measurable functions generated by the indicator functions $1_{B_j}$. The functions in $\mathcal{B}$ are constant on each set $B_j$. Then the conditional expectation of $f$ with finite variance is the projection

$$\mu(f \mid \mathcal{B}) = \sum_j \frac{\langle 1_{B_j}, f\rangle}{\langle 1_{B_j}, 1_{B_j}\rangle} 1_{B_j}. \tag{26.12}$$

Explicitly, this is

$$\mu(f \mid \mathcal{B}) = \sum_j \mu(f \mid B_j) 1_{B_j}. \tag{26.13}$$

Now specialize to the case when $f = 1_A$. Then this is the conditional probability random variable

$$\mu(A \mid \mathcal{B}) = \sum_j \mu(A \mid B_j) 1_{B_j}. \tag{26.14}$$

This is the usual formula for conditional probability. It says that the conditional probability of $A$ given which of the events $B_j$ happened depends on the outcome of the experiment. If the outcome is such that a particulary $B_j$ happened, then the value of the conditional probability is $\mu(A \mid B_j)$.

In this example the formula that the expectation of the conditional expectation is the expectation takes the form

$$\mu(f) = \sum_j \mu(f \mid B_j)\mu(B_j). \tag{26.15}$$

The corresponding formula for probability is

$$\mu(A) = \sum_j \mu(A \mid B_j)\mu(B_j). \tag{26.16}$$

Sometimes the notation $\mu(f \mid g)$ is used to mean $\mu(f \mid \sigma(g))$, where $\sigma(g)$ is the $\sigma$-algebra of measurable random variables generated by the random variable $g$. Since $\mu(f \mid g)$ belongs to this $\sigma$-algebra of functions, we have $\mu(f \mid g) = \phi(g)$ for some Borel function $\phi$. Thus the conditional expectation of $f$ given $g$ consists of the function $\phi(g)$ of $g$ that best predicts $f$ based on the knowledge of the value of $g$. Notice that the special feature of probability is not the projection operation, which is pure Hilbert space, but the nonlinear way of generating the closed subspace on which one projects.

## 26.4 The $L^1$ theory of conditional expectation

Consider again a probability space $\Omega, \mathcal{S}, \mu$. The conditional expectation may be defined for $f$ in $L^1(\Omega, \mathcal{S}, \mu)$. Let $\mathcal{F} \subset \mathcal{S}$ be a $\sigma$-algebra of functions. Let $f_n = f$ where $|f| \leq n$ and let $f_n = 0$ elsewhere. Then $f_n \to f$ in $L^1(\Omega, \mathcal{S}, \mu)$, by the dominated convergence theorem. So $L^2(\Omega, \mathcal{S}, \mu)$ is dense in $L^1(\Omega, \mathcal{S}, \mu)$. Furthermore, by a previous corollary the map $f \mapsto \mu(f \mid \mathcal{F})$ (defined as a projection in Hilbert space) is uniformly continuous with respect to the $L^1(\Omega, \mathcal{S}, \mu)$ norm. Therefore it extends by continuity to all of $L^1(\Omega, \mathcal{S}, \mu)$. It other words, for each $f$ in $L^1(\Omega, \mathcal{S}, \mu)$ the conditional expectation is defined and is an element of $L^1(\Omega, \mathcal{S}, \mu)$.

It is not hard to see that for $f$ in $L^1(\Omega, \mathcal{S}, \mu)$ the conditional expectation $\mu(f \mid \mathcal{F})$ is the element of $L^1(\Omega, \mathcal{S}, \mu)$ characterized by the following two properties. The first is that $\mu(f \mid \mathcal{F})$ is in $\mathcal{F}$, or equivalently, that $\mu(f \mid \mathcal{F})$ is in $L^1(\Omega, \mathcal{F}, \mu)$. The second is that for all $g$ in $L^\infty(\Omega, \mathcal{F}, \mu)$ we have $\mu(\mu(f \mid \mathcal{F})g) = \mu(fg)$.

Here is a technical remark. There is another way to construct the $L^1$ conditional expectation by means of the Radon-Nikodym theorem. Say that $f \geq 0$ is in $L^1(\Omega, \mathcal{S}, \mu)$. The idea is to look at the finite measure $\nu$ defined on $\mathcal{F}$ measurable functions $g \geq 0$ by $\nu(g) = \mu(fg)$. Suppose that $\mu(g) = 0$. Then the set where $g > 0$ is in $\mathcal{F}$ with $\mu$ measure zero, and so the set where $fg > 0$ is in $\mathcal{S}$ with $\mu$ measure zero. So $\nu(g) = 0$. This shows that $\nu \prec \mu$ as measures defined for $\mathcal{F}$ measurable functions. By the Radon-Nikodym theorem there exists an $h \geq 0$ in $L^1(\Omega, \mathcal{F}, \mu)$ such that $\nu(g) = \mu(hg)$ for all $g \geq 0$ that are $\mathcal{F}$ measurable. This $h$ is the desired conditional expectation $h = \mu(f \mid \mathcal{F})$.

Here is an example where conditional expectation calculations are simple. Say that $\Omega = \Omega_1 \times \Omega_2$ is a product space. The $\sigma$-algebra $\mathcal{S} = \mathcal{S}_1 \bigotimes \mathcal{S}_2$. There is a product reference measure $\nu = \nu_1 \times \nu_2$. The actual probability measure $\mu$ has a density $w$ with respect to this product measure:

$$\mu(f) = (\nu_1 \times \nu_2)(fw) = \int \int f(x, y) w(x, y) \, d\nu_1(x) \, d\nu_2(y). \tag{26.17}$$

Thus the experiment is carried on in two stages. What prediction can we make if we know the result for the first stage? Let $\mathcal{F}_1 = \mathcal{S}_1 \bigotimes \mathbb{R}$ consist of the functions $g(x, y) = h(x)$ where $h$ is in $\mathcal{S}_1$. This is the information given by the first stage. It is easy to compute the prediction $\mu(f \mid \mathcal{F}_1)$ for the second stage. The answer is

$$\mu(f \mid \mathcal{F}_1)(x, y) = \frac{\int f(x, y') w(x, y') \, d\nu_2(y')}{\int w(x, y') \, d\nu_2(y')}. \tag{26.18}$$

This is easy to check from the definition. Notice that the conditional expectation only depends on the first variable, so it is in $\mathcal{F}_1$. For those who like to express such results without the use of bound variables, the answer may also be written as

$$\mu(f \mid \mathcal{F}_1) = \frac{\nu_2 \circ (fw)^{|1}}{\nu_2 \circ w^{|1}}. \tag{26.19}$$

A notation such as $w^{|1}$ means the function that assigns to each $x$ the function $y \mapsto w(x, y)$ of the second variable. Thus $\nu_2 \circ w^{|1}$ means the composite function that assigns to each $x$ the integral $\nu_2(w^{|1}(x)) = \int w(x, y) \, d\nu_2(y)$ of this function of the second variable.

# Problems

1. Deduce the weak law of large numbers as a consequence of Hilbert space theory.

2. Consider the game show problem with the three doors $a, b, c$ and prize $X = a$, $b$, or $c$ with probability $1/3$ for each. Recall that the host chooses $g(X)$, where $g(c) = b$ and $g(b) = c$ and also $g(a) = b$, though this is not known to the contestant. (i) Find $P[X = a \mid g(X) = b]$ and $P[X = f(g(X)) \mid g(X) = b]$. If the game show host chooses $b$, does the contestant

gain by switching? (ii) Find $P[X = a \mid g(X) = c]$ and $P[X = f(g(X)) \mid g(X) = c]$. If the game show host chooses $c$, does the contestant gain by switching? (iii) Find the probabilities $P[g(X) = b]$ and $P[g(X) = c]$. (iv) Consider the random variable with value $P[X = f(g(X)) \mid g(X) = b]$ provided that $g(X) = b$ and with value $P[X = f(g(X)) \mid g(X) = c]$ provided that $g(X) = c$. Find the expectation of this random variable.

3. Say that $f$ is a random variable with finite variance, and $g$ is another random variable. How can one choose the function $\phi$ to make the expectation $\mu((f - \phi(g))^2)$ as small as possible?

4. Let $\lambda > 0$ be a parameter describing the rate at which accidents occur. Let $W_1$ be the time to wait for the first accident, and let $W_2$ be the time to wait from then until the second accident. These are each exponentially distributed random variables, and their joint distribution is given by a product measure. Thus
$$\mu(f(W_1, W_2)) = \int_0^\infty \int_0^\infty f(w_1, w_2)\lambda \exp(-\lambda w_1)\lambda \exp(-\lambda w_2)\, dw_1\, dw_2. \tag{26.20}$$
Let $T_1 = W_1$ be the time of the first accident, and let $T_2 = W_1 + W_2$ be the time of the second accident. Show that
$$\mu(h(T_1) \mid T_2) = \frac{1}{T_2} \int_0^{T_2} h(u)\, du. \tag{26.21}$$
That is, show that given the time $T_2$ of the second accident, the time $T_1$ of the first accident is uniformly distributed over the interval $[0, T_2]$. Hint: Make the change of variable $t_1 = w_1$ and $t_2 = w_1 + w_2$ and integrate with respect to $dt_1 dt_2$. Be careful about the limits of integration.

5. Can a $\sigma$-algebra of measurable functions (closed under pointwise operations of addition, multiplication, sup, inf, limits) be a finite dimensional vector space? Describe all such examples.

6. Say that $\mu$ is a probability measure or the corresponding expectation. Let $f$ be in $L^1$ and let $\mathcal{F}$ be a smaller $\sigma$-algebra of measurable functions. Then the conditional expectation $\mu(f \mid \mathcal{F})$ is defined by the properties that it is an $L^1$ function that is $\mathcal{F}$ measurable, and for all positive $L^\infty$ functions $g$ that are $\mathcal{F}$ measurable there is an identity
$$\mu(fg) = \mu(\mu(f \mid \mathcal{F})g). \tag{26.22}$$
Prove the monotone convergence theorem for conditional expectation. That is, prove that if $0 \leq f_n \uparrow f$, then $0 \leq \mu(f_n \mid \mathcal{F}) \uparrow \mu(f|\mathcal{F})$ almost everywhere and in $L^1$.

# Chapter 27

# Fourier series

## 27.1 Periodic functions

Let $T$ be the circle parameterized by $[0, 2\pi)$ or by $[-\pi, \pi)$. Let $f$ be a complex function in $L^2(T)$. The $n$th *Fourier coefficient* is

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} e^{-inx} f(x)\, dx. \tag{27.1}$$

The goal is to show that $f$ has a representation as a *Fourier series*

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}. \tag{27.2}$$

Another goal is to establish the equality

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2\, dx = \sum_{n=-\infty}^{\infty} |c_n|^2. \tag{27.3}$$

There are two problems with the Fourier series representation. One is to interpret the sense in which the series converges. The second is to show that it actually converges to $f$.

Before turning to these issues, it is worth looking at the intuitive significance of these formulas. Write $e^{inx} = \cos(nx) + i \sin(nx)$. Then

$$f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)], \tag{27.4}$$

where $a_n = c_n + c_{-n}$ and $b_n = i(c_n - c_{-n})$ for $n \geq 0$. Note that $b_0 = 0$. Also $2c_n = a_n - ib_n$ and $2c_{-n} = a_n + ib_n$ for $n \geq 0$. Furthermore, $|a_n|^2 + |b_n|^2 = 2(|c_n|^2 + |c_{-n}|^2)$.

In some applications $f(x)$ is real and the coefficients $a_n$ and $b_n$ are real. This is equivalent to $c_{-n} = \bar{c}_n$. In this case for $n \geq 0$ we can write $a_n = r_n \cos(\phi_n)$

and $b_n = r_n \sin(\phi_n)$, where $r_n = \sqrt{a_n^2 + b_n^2} \geq 0$. Thus $\phi_0$ is an integer multiple of $\pi$. Then the series becomes for real $f(x)$

$$f(x) = \frac{1}{2} r_0 \cos(-\phi_0) + \sum_{n=1}^{\infty} r_n \cos(nx - \phi_n). \tag{27.5}$$

We see that the $r_n$ determines the *amplitude* of the wave at angular frequency $n$, while the $\phi_n$ is a *phase*. Notice that the complex form coefficients are then $2c_n = r_n e^{-i\phi_n}$ and $c_{-n} = r_n e^{i\phi_n}$. Thus the coefficients in the complex expansion carry both the amplitude and phase information.

## 27.2 Convolution

It is possible to define Fourier coefficients for $f$ in $L^1(T)$. The formula is

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} e^{-inx} f(x)\, dx. \tag{27.6}$$

It is clear that the sequence of coefficients is in $\ell^\infty$.

If $f$ and $g$ are in $L^1(T)$, then we may define their *convolution* $f * g$ by

$$(f * g)(x) = \frac{1}{2\pi} \int_T f(x - y) g(y)\, dy = \frac{1}{2\pi} \int_T f(z) g(x - z)\, dz. \tag{27.7}$$

All integrals are over the circle continued periodically.

**Proposition 27.1** *If $f$ has Fourier coefficients $c_n$ and $g$ has Fourier coefficients $d_n$, then $f * g$ has Fourier coefficients $c_n d_n$.*

Proof: This is an elementary calculation. □

Another useful operation is the *adjoint (function)adjoint* of a function in $L^1(T)$. The adjoint function $f^*$ is defined by $f^*(x) = \overline{f(-x)}$.

**Proposition 27.2** *If $f$ has Fourier coefficients $c_n$, then its adjoint $f^*$ has Fourier coefficients $\bar{c}_n$.*

## 27.3 Approximate delta functions

An *approximate delta function* is a sequence of functions $\delta_a$ for $a > 0$ with the following properties.

1. For each $a > 0$ the integral $\int_{-\infty}^{\infty} \delta_a(x)\, dx = 1$.
2. The function $\delta_a(x) \geq 0$ is positive.
3. For each $c > 0$ the integrals satisfy $\lim_{a \to 0} \int_{|x| \geq c} \delta_a(x)\, dx = 0$.

**Theorem 27.3** *Let $\delta_a$ for $a > 0$ be an approximate $\delta$ function. Then for each bounded continuous function $f$ we have*

$$\lim_{a \to 0} \int_{-\infty}^{\infty} f(x-y)\delta_a(y)\,dy = f(x). \tag{27.8}$$

Proof: By the first property

$$\int_{-\infty}^{\infty} f(x-y)\delta_a(y)\,dy - f(x) = \int_{-\infty}^{\infty} [f(x-y) - f(x)]\delta_a(y)\,dy. \tag{27.9}$$

By the second property

$$|\int_{-\infty}^{\infty} f(x-y)\delta_a(y)\,dy - f(x)| \leq \int_{-\infty}^{\infty} |f(x-y) - f(x)|\delta_a(y)\,dy. \tag{27.10}$$

Consider $\epsilon > 0$. Then by the continuity of $f$ at $x$ there exists $c > 0$ such that $|y| < c$ implies $|f(x-y) - f(x)| < \epsilon/2$. Suppose $|f(x)| \leq M$ for all $x$. Break up the integral into the parts with $|y| \geq c$ and $|y| < c$. Then using the first property on the second term we get

$$|\int_{-\infty}^{\infty} f(x-y)\delta_a(y)\,dy - f(x)| \leq 2M \int_{|y|>c} \delta_a(y)\,dy + \epsilon/2 \tag{27.11}$$

The third property says that for sufficiently small $a > 0$ we can get the first term also bounded by $\epsilon/2$. □

There is also a concept of approximate delta function for functions on the circle $T$. This is what we need for the application to Fourier series. In fact, there is an explicit formula given by the *Poisson kernel* for the circle. For $0 \leq r < 1$ let

$$P_r(x) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{inx} = \frac{1 - r^2}{1 - 2r\cos(x) + r^2}. \tag{27.12}$$

The identity is proved by summing a geometric series. Then the functions $\frac{1}{2\pi} P_r(x)$ have the properties of an approximate delta function as $r$ approaches 1. Each such function is positive and has integral 1 over the periodic interval. Furthermore,

$$P_r(x) \leq \frac{1 - r^2}{2r(1 - \cos(x))}, \tag{27.13}$$

which approaches zero as $r \to 1$ away from points where $\cos(x) = 1$.

## 27.4 Abel summability

The following theorem shows that the Fourier series of a continuous function on the circle is always *Abel summable*. This means that one multiplies the coefficients by $r^{|n|}$ with $0 < r < 1$, performs the resulting sum, and then takes the limit as $r$ increases to 1. This is formally the same as taking the usual sum, but there is no guarantee that this usual sum is absolutely convergent. The most important consequence of Abel summability is that the Fourier coefficients uniquely determine the function.

**Theorem 27.4** *Let $f$ in $C(T)$ be a continuous function on the circle. Then*

$$f(x) = \lim_{r \uparrow 1} \sum_{n=-\infty}^{\infty} r^{|n|} c_n e^{inx}. \tag{27.14}$$

Proof: Proof: It is easy to compute that

$$\frac{1}{2\pi} \int_0^{2\pi} P_r(y) f(x-y)\, dy = \sum_{n=-\infty}^{\infty} r^{|n|} c_n e^{inx}. \tag{27.15}$$

Let $r \uparrow 1$. Then by the theorem on approximate delta functions

$$f(x) = \lim_{r \uparrow 1} \sum_{n=-\infty}^{\infty} r^{|n|} c_n e^{inx}. \tag{27.16}$$

□

**Corollary 27.5** *Let $f$ in $C(T)$ be a continuous function on the circle. Suppose that the Fourier coefficients $c$ of $f$ are in $\ell^1$. Then*

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}. \tag{27.17}$$

*The convergence is uniform.*

Proof: If in addition $c$ is in $\ell^1$, then the dominated convergence theorem for sums says it is possible to interchange the limit and the sum. □

## 27.5 $L^2$ convergence

The simplest and most useful theory is in the context of Hilbert space. The result of this section shows that a square-integrable $2\pi$-periodic function may be specified by giving its Fourier coefficients at all frequencies, and conversely, every $\ell^2$ sequence of coefficients gives rise to such a function. There is a perfect equivalence between the two descriptions.

Let $L^2(T)$ be the space of all (Borel measurable) functions such that

$$\|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2\, dx < \infty. \tag{27.18}$$

Then $L^2(T)$ is a Hilbert space with inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x)\, dx. \tag{27.19}$$

Here $T$ is the circle, regarded as parameterized by an angle that goes from 0 to $2\pi$.

Let

$$\phi_n(x) = \exp(inx). \tag{27.20}$$

Then the $\phi_n$ form an orthonormal family in $L^2(T)$. It follows from general Hilbert space theory (theorem of Pythagoras) that

$$\|f\|_2^2 = \sum_{|n|\leq N} |c_n|^2 + \|f - \sum_{|n|\leq N} c_n\phi_n\|_2^2. \tag{27.21}$$

In particular, Bessel's inequality says that

$$\|f\|_2^2 \geq \sum_{|n|\leq N} |c_n|^2. \tag{27.22}$$

This shows that

$$\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty. \tag{27.23}$$

The space of sequences satisfying this identity is $\ell^2$. Thus we have proved the following proposition.

**Proposition 27.6** *If $f$ is in $L^2(T)$, then its sequence of Fourier coefficients is in $\ell^2$.*

**Theorem 27.7** *If $f$ is in $L^2(T)$, then*

$$\|f\|_2^2 = \sum_n |c_n|^2. \tag{27.24}$$

Proof: The function $f$ has Fourier coefficients $c_n$. The adjoint function $f^*$ defined by $f^*(x) = \overline{f(-x)}$ has complex conjugate Fourier coefficients $\bar{c}_n$. The coefficients of a convolution are the product of the coefficients. Hence $g = f^* * f$ has coefficients $\bar{c}_n c_n = |c_n|^2$.

Suppose that $f$ is in $L^2(T)$. Then $g = f^* * f$ is in $C(T)$. In fact,

$$g(x) = \frac{1}{2\pi}\int_T \overline{f(y-x)}f(y)\,dy = \langle f_x, f\rangle, \tag{27.25}$$

where $f_x$ is $f$ translated by $x$. Since translation is continuous in $L^2(T)$, it follows that $g$ is a continuous function. Furthermore, since $f$ is in $L^2(T)$, it follows that $c$ is in $\ell^2$, and so $|c|^2$ is in $\ell^1$. Thus the theorem applies, and

$$g(x) = \frac{1}{2\pi}\int_T \overline{f(y-x)}f(y)\,dy = \sum_n |c_n|^2 e^{inx}. \tag{27.26}$$

The conclusion follows by taking $x = 0$. □

**Theorem 27.8** *If $f$ is in $L^2(T)$, then*

$$f = \sum_{n=-\infty}^{\infty} c_n \phi_n \tag{27.27}$$

*in the sense that*

$$\lim_{N\to\infty} \|f - \sum_{|n|\leq N} c_n\phi_n\|_2^2 = 0. \tag{27.28}$$

Proof: Use the identity

$$\|f\|_2^2 = \sum_{|n|\leq N} |c_n|^2 + \|f - \sum_{|n|\leq N} c_n\phi_n\|_2^2. \tag{27.29}$$

The first term on the right hand side converges to the left hand side, so the second term on the right hand side must converge to zero. □

## 27.6 C(T) convergence

Define the function spaces

$$C(T) \subset L^\infty(T) \subset L^2(T) \subset L^1(T). \tag{27.30}$$

The norms $\|f\|_\infty$ on the first two spaces are the same, the smallest number $M$ such that $|f(x)| \leq M$ (with the possible exception of a set of $x$ of measure zero). The space $C(T)$ consists of continuous functions; the space $L^\infty(T)$ consists of all bounded functions. The norm on $L^2(T)$ is given by $\|f\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\,dx$. The norm on $L^1(T)$ is given by $\|f\|_1 = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|\,dx$. Since the integral is a probability average, their relation is

$$\|f\|_1 \leq \|f\|_2 \leq \|f\|_\infty. \tag{27.31}$$

Also define the sequence spaces

$$\ell^1 \subset \ell^2 \subset c_0 \subset \ell^\infty. \tag{27.32}$$

The norm on $\ell^1$ is $\|c\|_1 = \sum_n |c_n|$. Then norm on $\ell^2$ is given by $\|c\|_2^2 = \sum_n |c_n|^2$. The norms on the last two spaces are the same, that is, $\|c\|_\infty$ is the smallest $M$ such that $|c_n| \leq M$. The space $c_0$ consists of all sequences with limit 0 at infinity. The relation between these norms is

$$\|c\|_\infty \leq \|c\|_2 \leq \|c\|_1. \tag{27.33}$$

We have seen that the Fourier series theorem gives a perfect correspondence between $L^2(T)$ and $\ell^2$. For the other spaces the situation is more complicated. Some useful information is expressed in the *Riemann-Lebesgue lemma.*

**Lemma 27.9 (Riemann-Lebesgue)** *If $f$ is in $L^1(T)$, then the Fourier coefficients of $f$ are in $c_0$, that is, they approaches 0 at infinity.*

Proof: Each function in $L^2(T)$ has Fourier coefficients in $\ell^2$, so each function in $L^2(T)$ has Fourier coefficients that vanish at infinity. The map from a function to its Fourier coefficients gives a continuous map from $L^1(T)$ to $\ell^\infty$. However every function in $L^1(T)$ may be approximated arbitrarily closely in $L^1(T)$ norm by a function in $L^2(T)$. Hence its coefficients may be approximated arbitrarily well in $\ell^\infty$ norm by coefficients that vanish at infinity. Therefore the coefficients vanish at infinity. □

In summary, the map from a function to its Fourier coefficients gives a continuous map from $L^1(T)$ to $c_0$. That is, the Fourier coefficients of an integrable function are bounded (this is obvious) and approach zero (Riemann-Lebesgue lemma). Furthermore, it may be shown that the Fourier coefficients determine the function uniquely.

The map from Fourier coefficients to functions gives a continuous map from $\ell^1$ to $C(T)$. An sequence that is absolutely summable defines a Fourier series that converges absolutely and uniformly to a continuous function.

For the next result the following lemma will be useful.

**Lemma 27.10** *Say that*

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{27.34}$$

*with $L^2(T)$ convergence. Then the identity*

$$f'(x) = \sum_{n=-\infty}^{\infty} inc_n e^{inx} \tag{27.35}$$

*obtained by differentiating holds, again in the sense of $L^2(T)$ convergence. Here the relation between $f$ and $f'$ is that $f$ is an indefinite integral of $f'$. Furthermore $f'$ has integral zero and $f$ is periodic.*

Proof: Let $h(x) = \sum_{n\neq 0} b_n e^{inx}$ be in $L^2(T)$. Define the integral $Vh$ by $(Vh)(x) = \int_0^x h(y)\,dy$. Then

$$\|Vh\|_2 \leq \|Vh\|_\infty \leq (2\pi)\|h\|_1 \leq (2\pi)\|h\|_2. \tag{27.36}$$

This shows that $V$ is continuous from $L^2(T)$ to $L^2(T)$. Thus we can apply $V$ to the series term by term. This gives

$$(Vh)(x) = \sum_{n\neq 0} b_n \frac{e^{inx}-1}{in} = C + \sum_{n\neq 0} \frac{b_n}{in} e^{inx}. \tag{27.37}$$

Thus the effect of integrating is to divide the coefficient by $in$. Since differentiation has been defined in this context to be the inverse of integration, the effect of differentiation is to multiply the coefficient by $in$. □

**Theorem 27.11** *If $f$ is in $L^2(T)$ and if $f'$ exists (in the sense that $f$ is an integral of $f$) and if $f'$ is also in $L^2(T)$, then the Fourier coefficients are in $\ell^1$. Therefore the Fourier series converges in the $C(T)$ norm.*

Proof: The hypothesis of the theorem means that there is a function $f'$ in $L^2(T)$ with $\int_0^{2\pi} f'(y)\,dy = 0$. Then $f$ is a function defined by

$$f(x) = c_0 + \int_0^x f'(y)\,dy \tag{27.38}$$

with an arbitrary constant of integration. This $f$ is an absolutely continuous function. It is also periodic, because of the condition on the integral of $f'$.

The proof is completed by noting that

$$\sum_{n\neq 0} |c_n| = \sum_{n\neq 0} \frac{1}{|n|}|nc_n| \leq \sqrt{\sum_{n\neq 0}\frac{1}{n^2}}\sqrt{\sum_{n\neq 0} n^2|c_n|^2}. \tag{27.39}$$

In other words,

$$\sum_{n\neq 0} |c_n| \leq \sqrt{\frac{\pi^2}{3}}\|f'\|_2. \tag{27.40}$$

□

## 27.7 Pointwise convergence

There remains one slightly unsatisfying point. The convergence in the $L^2$ sense does not imply convergence at a particular point. Of course, if the derivative is in $L^2$ then we have uniform convergence, and in particular convergence at each point. But what if the function is differentiable at one point but has discontinuities at other points? What can we say about convergence at that one point? Fortunately, we can find something about that case by a closer examination of the partial sums.

One looks at the partial sum

$$\sum_{|n|\leq N} c_n e^{inx} = \frac{1}{2\pi}\int_0^{2\pi} D_N(x-y)f(y)\,dy. \tag{27.41}$$

Here

$$D_N(x) = \sum_{|n|\leq N} e^{inx} = \frac{\sin((N+\frac{1}{2})x)}{\sin(\frac{1}{2}x)}. \tag{27.42}$$

This *Dirichlet kernel* $D_N(x)$ has at least some of the properties of an approximate delta function. Unfortunately, it is not positive; instead it oscillates wildly for large $N$ at points away from where $\sin(x/2) = 0$. However the function $1/(2\pi)D_N(x)$ does have integral 1.

**Theorem 27.12** *If for some $x$ the function*

$$d_x(z) = \frac{f(x+z) - f(x)}{2\sin(z/2)} \tag{27.43}$$

*is in $L^1(T)$, then at that point*

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \phi_n(x). \tag{27.44}$$

Note that if $d_x(z)$ is continuous at $z = 0$, then its value at $z = 0$ is $d_x(0) = f'(x)$. So the hypothesis of the theorem is a condition related to differentiability of $f$ at the point $x$. The conclusion of the theorem is pointwise convergence of the Fourier series at that point. Since $f$ may be discontinuous at other points, it is possible that this Fourier series is not absolutely convergent. Thus the series must be interpreted as the limit of the partial sums over $|n| \leq N$, taken as $N \to \infty$.

Proof: We have

$$f(x) - \sum_{|n|\leq N} c_n e^{inx} = \frac{1}{2\pi}\int_0^{2\pi} D_N(z)(f(x) - f(x-z))\,dz. \tag{27.45}$$

We can write this as

$$f(x) - \sum_{|n|\leq N} c_n e^{inx} = \frac{1}{2\pi}\int_0^{2\pi} 2\sin((N+\frac{1}{2})z)d_x(-z)\,dz. \tag{27.46}$$

This goes to zero as $N \to \infty$, by the Riemann-Lebesgue lemma. □

## 27.8 Supplement: Ergodic actions

Consider the case where a group acts on a measurable space $X$. This defines an equivalence relation, where the equivalence classes are the orbits of the group. A probability measure $\mu$ on $X$ is *ergodic* if every measurable invariant subset (union of equivalence classes) has $\mu$ probability zero or one.

The historical origin of this subject is the case when the group action consists of time translation acting on some space of configurations. There is an invariant probability measure $\mu$ on the space of configurations, and the ergodic condition implies that long time averages of various quantities may be obtained by taking the expectation with respect to the probability measure. See the book by Sinai [19] for an introduction to the theory.

The next theorem gives the simplest example, where the group is $\mathbb{Z}$ and it acts on the circle $T$ by rotation by an irrational angle.

**Theorem 27.13** *Let $T$ be the circle of circumference one with the rotationally invariant probability measure. Let $\alpha$ be an irrational number. Then rotation by $\alpha$ is ergodic.*

Proof: The group action is given by $n \cdot x = x + n\alpha$ modulo 1. Two points $x, x'$ are in the same orbit if there are $n, n'$ with $n \cdot x = n' \cdot x'$. This says that $x - x' = (n' - n)\alpha$ modulo 1.

The proof that this is an ergodic action follows easily from Fourier analysis. Let $f$ be an $L^2$ function on $T$ that is invariant under the irrational rotation. Thus $F$ could be the indicator function of a measurable subset that is invariant under irrational rotation. Expand $f$ in a Fourier series

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi i k x}. \tag{27.47}$$

The condition that $f$ is invariant under the irrational rotation by $\alpha$ translates to the condition that

$$c_k = e^{2\pi i k \alpha} c_k. \tag{27.48}$$

So either $c_k = 0$ or the phase $e^{2\pi i k\alpha} = 1$. This last is true if and only if $k\alpha = m$ for some integer $m$. Since $\alpha$ is irrational, this can be true only if $k = 0$. We conclude that $k \neq 0$ implies $c_k = 0$. Thus $f(x) = c_0$ is constant. If $f$ is an indicator function, then the corresponding measurable subset is either empty or the whole circle, up to sets of measure zero. $\square$

This result has a generalization to the torus $T^2$ that is the product of two circles. What is interesting here is the condition on the action of the group $\mathbb{Z}$. Contrast this with the condition in the following theorem, where the group action is by $\mathbb{R}$.

**Theorem 27.14** *Let $T^2$ be the torus that is the product of two circles each of circumference one. The measure is Lebesgue measure. Let $\alpha$ and $\beta$ be two numbers. Suppose that whenever $p$ and $q$ and $m$ are integers with $p\alpha + q\beta = m$, then $p = q = 0$. Then rotation by $\alpha, \beta$ is ergodic.*

Proof: The group action is given by $n \cdot (x, y) = (x + n\alpha, y + n\beta)$, with the addition taken modulo 1. The proof of ergodicity is left to the reader. It uses the Fourier series expansion

$$f(x, y) = \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} c_{p,q} e^{2\pi i [px + qy]}. \tag{27.49}$$

The rest of the proof is much as in the previous result. $\square$

The group does not have to be discrete. Consider the case is where the space $X = T^2$ is the torus, and the group consists of the reals $\mathbb{R}$.

**Theorem 27.15** *Let $T^2$ be the that is the product of two circles, each of circumference one. The measure is Lebesgue measure. Let $\alpha$ and $\beta$ be numbers such that whenever $p$ and $q$ are integers with $p\alpha + q\beta = 0$, then $p = q = 0$. Then rotation by $\alpha, \beta$ is ergodic.*

Proof: The group action is given by $t \cdot (x, y) = (x + t\alpha, y + t\beta)$, where the sums are taken modulo 1. The proof that this is an ergodic action again follows from Fourier analysis. Let $f$ be an $L^2$ function on $T^2$ that is invariant under the irrational rotation. Expand $f$ in a Fourier series

$$f(x, y) = \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} c_{p,q} e^{2\pi i[px+qy]}. \tag{27.50}$$

The condition that $f$ is invariant under the rotations translates to the condition that for each real $t$

$$c_{p,q} = e^{2\pi i t[p\alpha+q\beta]} c_{p,q}. \tag{27.51}$$

So either $c_{p,q} = 0$ or all the phases $e^{2\pi i t[p\alpha+q\beta]} = 1$. This last is true only if $p\alpha + q\beta = 0$, that is, only if $p = q = 0$. □

# Problems

1. Let $f(x) = x$ defined for $-\pi \leq x < \pi$. Find the $L^1(T)$, $L^2(T)$, and $L^\infty(T)$ norms of $f$, and compare them.

2. Find the Fourier coefficients $c_n$ of $f$ for all $n$ in $Z$.

3. Find the $\ell^\infty$, $\ell^2$, and $\ell^1$ norms of these Fourier coefficients, and compare them.

4. Use the equality of $L^2$ and $\ell^2$ norms to compute

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2}. \tag{27.52}$$

5. Compare the $\ell^\infty$ and $L^1$ norms for this problem. Compare the $L^\infty$ and $\ell^1$ norms for this problem.

6. Use the pointwise convergence at $x = \pi/2$ to evaluate the infinite sum

$$\sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1}, \tag{27.53}$$

regarded as a limit of partial sums. Does this sum converge absolutely?

7. Let $F(x) = \frac{1}{2}x^2$ defined for $-\pi \leq x < \pi$. Find the Fourier coefficients of this function.

8. Use the equality of $L^2$ and $\ell^2$ norms to compute

$$\zeta(4) = \sum_{n=1}^{\infty} \frac{1}{n^4}. \tag{27.54}$$

9. Compare the $\ell^\infty$ and $L^1$ norms for this problem. Compare the $L^\infty$ and $\ell^1$ norms for this problem.

10. At which points $x$ of $T$ is $F(x)$ continuous? Differentiable? At which points $x$ of $T$ is $f(x)$ continuous? Differentiable? At which $x$ does $F'(x) = f(x)$? Can the Fourier series of $f(x)$ be obtained by differentiating the Fourier series of $F(x)$ pointwise? (This last question can be answered by inspecting the explicit form of the Fourier series for the two problems.)

11. (a) Evaluate
$$\lim_{n\to\infty} \int_0^1 (\sin^2(2\pi x))^n \, dx. \tag{27.55}$$
(b) Evaluate
$$\lim_{n\to\infty} \int_0^1 \sin^2(2\pi n x) \, dx. \tag{27.56}$$
(c) Evaluate
$$\lim_{n\to\infty} \int_0^1 \sin(2\pi n x) \frac{1}{\sqrt{x}} \, dx. \tag{27.57}$$

12. Ergodic actions. Give an example of $\alpha, \beta$ such that $p\alpha + q\beta = 0$ implies $p = q = 0$ but $p\alpha + q\beta = m$ does not imply $p = q = 0$.

13. Ergodic actions. Give an example of $\alpha, \beta$ such that $p\alpha + q\beta = m$ does not imply $p = q = 0$.

# Chapter 28

# Fourier transforms

## 28.1 Fourier analysis

The general context of Fourier analysis is an abelian group and its dual group. The elements $x$ of the abelian group are thought of as space (or time) variables, while the elements of the dual group are thought of as *wave number* (or *angular frequency*) variables.

Examples:

1. Let $\Delta x > 0$. The finite group consists of all $x = j\Delta x$ for $j = 0, \ldots N-1$ with addition mod $N\Delta x$. The dual group is the finite group $k = \ell\Delta k$ with $\ell = 0, \ldots N-1$ with addition mod $N\Delta k$. Here $N\,\Delta x\Delta k = 2\pi$.

2. Let $L > 0$. The compact group $T_L$ consists of all $x$ in the circle of circumference $L$ with addition mod $L$. The dual group is the discrete group $\mathbb{Z}\,\Delta k$ consisting of all $k = \ell\Delta k$ with $\ell \in \mathbf{Z}$. Here $L\,\Delta k = 2\pi$.

3. The discrete group $\mathbb{Z}\,\Delta x$ consists of all $x = j\Delta x$ with $j \in \mathbf{Z}$. The dual group $T_B$ is the compact group of $k$ in the circle of circumference $B$ with addition mod $B$, where $\Delta x\, B = 2\pi$.

4. The group $\mathbb{R}$ consists of all $x$ in the real line. The dual group $\mathbb{R}$ is all $k$ in the (dual) line.

In each case the formula are the essentially the same. Let $\lambda > 0$ be an arbitrary constant. We have dual measures $dx/\lambda$ and $\lambda dk/(2\pi)$. Their product is $dx\,dk/(2\pi)$. The Fourier transform is

$$\hat{f}(k) = \int e^{-ikx} f(x)\,\frac{1}{\lambda} dx. \tag{28.1}$$

The integral is over the group. For a finite or discrete group it is a sum, and the $dx$ is replaced by $\Delta x$. The Fourier representation is then given by the inversion formula

$$f(x) = \int e^{ikx} \hat{f}(k)\, \lambda \frac{dk}{2\pi}. \tag{28.2}$$

The integral is over the dual group. For a finite or discrete dual group it is a sum, and the $dk$ is replaced by $\Delta k$.

The constant $\lambda > 0$ is chosen for convenience. There is a lot to be said for standardizing on $\lambda = 1$. In the case of the circle one variant is $\lambda = L$. This choice makes $dx/\lambda$ a probability measure and $\lambda\Delta k/(2\pi) = 1$. Similarly, in the case of the discrete group $\lambda = \Delta x$ makes $\Delta x/\lambda = 1$ and $\lambda\, dk/(2\pi)$ a probability measure. In the case of the line $\lambda = 1$ is most common, though some people prefer the ugly choice $\lambda = \sqrt{2\pi}$ in a misguided attempt at symmetry.

## 28.2 $L^1$ theory

Let $f$ be a complex function on the line that is in $L^1$. The *Fourier transform* $\hat{f}$ is the function defined by

$$\hat{f}(k) = \int_{-\infty}^{\infty} e^{-ikx} f(x)\, dx. \tag{28.3}$$

Note that if $f$ is in $L^1$, then its Fourier transform $\hat{f}$ is in $L^\infty$ and satisfies $\|\hat{f}\|_\infty \leq \|f\|_1$. Furthermore, it is a continuous function.

Similarly, let $g$ be a function on the (dual) line that is in $L^1$. Then the *inverse Fourier transform* $\check{g}$ is defined by

$$\check{g}(x) = \int_{-\infty}^{\infty} e^{ikx} g(k)\, \frac{dk}{2\pi}. \tag{28.4}$$

If $g$ is a function and $y$ is a real number, then the function $x \mapsto g(x-y)$ is called the *translate* (or shift) of $g$ by $y$. The purpose of Fourier analysis is to analyze operations that are built out of translation, such as convolution and differentiation. The ultimate reason that this succeeds is that the effect of translation on the Fourier transform is simple. It replaces $\hat{g}$ by the function $k \mapsto e^{iky}\hat{g}(k)$. In other words, it is just pointwise multiplication by a phase factor.

We can look at the Fourier transform from a more abstract point of view. The space $L^1$ is a Banach space. Its dual space is $L^\infty$, the space of essentially bounded functions. An example of a function in the dual space is the exponential function $\phi_k(x) = e^{ikx}$. The Fourier transform is then

$$\hat{f}(k) = \langle \phi_k, f\rangle = \int_{-\infty}^{\infty} \overline{\phi_k(x)} f(x)\, dx, \tag{28.5}$$

where $\phi_k$ is in $L^\infty$ and $f$ is in $L^1$.

**Proposition 28.1** *$f, g$ are in $L^1(\mathbb{R}, dx)$, then the* convolution *$f * g$ is another function in $L^1(\mathbb{R}, dx)$ defined by*

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\,dy = \int_{-\infty}^{\infty} f(y)g(x-y)\,dy. \tag{28.6}$$

**Proposition 28.2** *If $f, g$ are in $L^1(\mathbb{R}, dx)$, then the Fourier transform of the convolution is the product of the Fourier transforms:*

$$\widehat{(f * g)}(k) = \hat{f}(k)\hat{g}(k). \tag{28.7}$$

Notice that convolution is defined in terms of translation. As a consequence, the Fourier transform performs a great simplification, turning convolution into pointwise multiplication.

As before, we define the *adjoint!function)adjoint* of a function $f$ by $f^*(x) = \overline{f(-x)}$. We shall see the reason for the term adjoint in the context of the $L^2$ theory.

**Proposition 28.3** *Let $f^*$ be the adjoint of $f$. Then the Fourier transform of $f^*$ is the complex conjugate of $\hat{f}$.*

**Theorem 28.4** *If $f$ is in $L^1$ and is also continuous and bounded, we have the inversion formula in the form*

$$f(x) = \lim_{\epsilon \downarrow 0} \int_{-\infty}^{\infty} e^{ikx} \hat{\delta}_\epsilon(k) \hat{f}(k) \frac{dk}{2\pi}, \tag{28.8}$$

*where*

$$\hat{\delta}_\epsilon(k) = \exp(-\epsilon|k|). \tag{28.9}$$

Proof: The inverse Fourier transform of this is

$$\delta_\epsilon(x) = \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2}. \tag{28.10}$$

It is easy to calculate that

$$\int_{-\infty}^{\infty} e^{ikx} \hat{\delta}_\epsilon(k) \hat{f}(k) \frac{dk}{2\pi} = (\delta_\epsilon * f)(x). \tag{28.11}$$

However $\delta_\epsilon$ is an approximate delta function. The result follows by taking $\epsilon \to 0$. □

## 28.3 $L^2$ theory

The space $L^2$ is its own dual space, and it is a Hilbert space. It is the setting for the most elegant and simple theory of the Fourier transform. This is the *Plancherel theorem* that says that the Fourier transform is an isomorphism of Hilbert spaces. In order words, there is a complete equivalence between the time description and the frequency description of a function.

**Lemma 28.5** *If $f$ is in $L^1(\mathbb{R}, dx)$ and in $L^2(\mathbb{R}, dx)$, then $\hat{f}$ is in $L^2(\mathbb{R}, dk/(2\pi))$, and $\|f\|_2^2 = \|\hat{f}\|_2^2$.*

Proof: Let $g = f^* * f$. Then $g$ is in $L^1$, since it is the convolution of two $L^1$ functions. Furthermore, it is continuous and bounded. This follows from the representation $g(x) = \langle f_x, f\rangle$, where $f_x$ is translation by $x$. Since $x \to f_x$ is continuous from $\mathbb{R}$ to $L^2$, the result follows from the Hilbert space continuity of the inner product. Finally, the Fourier transform of $g$ is $|\hat{f}(k)|^2$. Thus

$$\|f\|_2^2 = g(0) = \lim_{\epsilon \downarrow 0} \int_{-\infty}^{\infty} \hat{\delta}_\epsilon(k)|\hat{f}(k)|^2 \frac{dk}{2\pi} = \int_{-\infty}^{\infty} |\hat{f}(k)|^2 \frac{dk}{2\pi} \tag{28.12}$$

by the monotone convergence theorem. □

**Theorem 28.6 (Plancherel theorem)** *The Fourier transform $F$ initially defined on $L^1(\mathbb{R}, dx) \cap L^2(\mathbb{R}, dx)$ extends by uniform continuity to $F : L^2(\mathbb{R}, dx) \to L^2(\mathbb{R}, dk/(2\pi))$. The inverse Fourier transform $F^*$ initially defined on $L^1(\mathbb{R}, dk/(2\pi)) \cap L^2(\mathbb{R}, dk/(2\pi))$ extends by uniform continuity to $F^* : L^2(\mathbb{R}, dk/(2\pi)) \to L^2(\mathbb{R}, dx)$. These are linear transformations that preserve $L^2$ norm and preserve inner product. Furthermore, $F^*$ is the inverse of $F$.*

Proof: It is easy to see that $L^1 \cap L^2$ is dense in $L^2$. Here is the proof. Take $f$ in $L^2$ and let $A_n$ be a sequence of sets of finite measure that increase to all of $\mathbb{R}$. Then $1_{A_n} f$ is in $L^1$ for each $n$, by the Schwarz inequality. Furthermore, $1_{A_n} f \to f$ in $L^2$, by the $L^2$ dominated convergence theorem.

The lemma shows that $F$ is an isometry, hence uniformly continuous. Furthermore, the target space $L^2$ is a complete metric space. Thus $F$ extends by uniform continuity to the entire domain space $L^2$. It is easy to see that this extension is also an isometry.

The same reasoning shows that the inverse Fourier transform $F^*$ also maps $L^2$ onto $L^2$ and preserves norm.

Now it is easy to check that $(F^*h, f) = (h, Ff)$ for $f$ and $h$ in $L^1 \cap L^2$. This identity extends to all of $L^2$. Take $h = Fg$. Then $\langle F^*Fg, f)\rangle = \langle Fg, Ff\rangle = \langle g, f\rangle$. That is $F^*Fg = g$. Similarly, one may show that $FF^*u = u$. These equations show that $F^* = F^{-1}$ is the inverse of $F$. □

**Corollary 28.7** *Let $f$ be in $L^2$. Let $A_n$ be a sequences subsets of finite measure that increase to all of $\mathbb{R}$. Then $1_{A_n} f$ is in $L^1 \cap L^2$ and $F(1_{A_n} f) \to F(f)$ in $L^2$ as $n \to \infty$. That is, for fixed $n$ the function with values given by*

$$F(1_{A_n} f)(k) = \int_{A_n} e^{-ikx} f(x)\, dx \tag{28.13}$$

*is well defined for each $k$, and the sequence of such functions converges in the $L^2$ sense to the Fourier transform $Ff$, where the function $(Ff)(k)$ is defined for almost every $k$. Explicitly, this says that the Fourier transform of an $L^2$ function $f$ is the $L^2$ function $\hat{f} = Ff$ characterized by*

$$\int_{-\infty}^{\infty} |\hat{f}(k) - \int_{A_n} e^{-ikx} f(x)\, dx|^2 \frac{dk}{2\pi} \to 0 \tag{28.14}$$

*as* $n \to \infty$.

Another interesting result about convolution is *Young's inequality.* The case of interest for us is the Hilbert space case $p = 2$.

**Theorem 28.8 (Young's inequality)** *If $f$ is in $L^1$ and $g$ is in $L^p$, $1 \le p \le \infty$, then*

$$\|f * g\|_p \le \|f\|_1 \|g\|_p. \tag{28.15}$$

.

Proof: Here is a proof for the case when $1 \le p < \infty$. Consider the function $|g(x, y)|$. The left hand side of the inequality is bounded by the $L^p$ norm with respect to $dx$ of the integral of this function with the finite measure $|f(y)|\,dy$. Minkowski's inequality for integrals says that the $L^p$ norm of the integral is bounded by the integral of the $L^p$ norms. So the left hand side is bounded by the integral with respect to $|f(y)|\,dy$ of the $L^p$ norm of $|g(x - y)|$ with respect to $dx$. A miracle occurs: due to translation invariance this is independent of $y$ and in fact equal to $\|g\|_p$. The remaining integral with respect to $dy$ gives the other factor $\|f\|_1$. □

A special case of Young's inequality is that if $f$ is in $L^1$ and $g$ is in $L^2$, then the convolution $f * g$ is in $L^2$. In this context $\hat{f}$ is in $L^\infty$ and $\hat{g}$ is in $L^2$, so the Fourier transform of the convolution $f * g$ in $L^2$ is the pointwise product $\hat{f}\hat{g}$ in $L^2$.

This sheds light on the role of the adjoint function $f^*$. It is not difficult to verify that $\langle f^* * h, g\rangle = \langle h, f * g\rangle$. In other words, convolution by $f^*$ is the adjoint in the usual Hilbert space sense of convolution by $f$.

Since $f^* * h$ has Fourier transform $\overline{\hat{f}}\hat{f}$, the Fourier transform takes convolution by the adjoint function into pointwise multiplication by the complex conjugate function.

## 28.4 Absolute convergence

We have seen that the Fourier transform gives a perfect correspondence between $L^2(\mathbb{R}, dx)$ and $L^2(\mathbb{R}, dk/(2\pi))$. For the other spaces the situation is more complicated. It is difficult to characterize the image of $L^1(\mathbb{R}, dx)$, but the *Riemann-Lebesgue lemma* gives some information about it.

**Theorem 28.9 (Riemann-Lebesgue lemma)** *The map from a function to its Fourier transform gives a continuous map from $L^1(\mathbb{R}, dx)$ to $C_0(\mathbb{R})$. That is, the Fourier transform of an integrable function is continuous and bounded and approaches zero at infinity.*

Proof: We have seen that the Fourier transform of an $L^1$ function is bounded and continuous. The main content of the Riemann-Lebesgue lemma is that is also goes to zero at infinity. This can be proved by checking it on a dense subset, such as the space of step functions. □

One other useful fact is that if $f$ is in $L^1(\mathbb{R}, dx)$ and $g$ is in $L^2(\mathbb{R}, dx)$, then the convolution $f * g$ is in $L^2(\mathbb{R}, dx)$. Furthermore, $\widehat{f * g}(k) = \hat{f}(k)\hat{g}(k)$ is the product of a bounded function with an $L^2(\mathbb{R}, dk/(2\pi))$ function and therefore is in $L^2(\mathbb{R}, dk/(2\pi))$.

However the same pattern of the product of a bounded function with an $L^2(\mathbb{R}, dk/(2\pi))$ function can arise in other ways. For instance, consider the translate $f_a$ of a function $f$ in $L^2(\mathbb{R}, dx)$ defined by $f_a(x) = f(x - a)$. Then $\widehat{f_a}(k) = \exp(-ika)\hat{f}(k)$. This is also the product of a bounded function with an $L^2(\mathbb{R}, dk/(2\pi))$ function.

One can think of this last example as a limiting case of a convolution. Let $\delta_\epsilon$ be an approximate $\delta$ function. Then $(\delta_\epsilon)_a * f$ has Fourier transform $\exp(-ika)\hat{\delta}_\epsilon(k)\hat{f}(k)$. Now let $\epsilon \to 0$. Then $(\delta_\epsilon)_a * f \to f_a$, while $\exp(-ika)\hat{\delta}_\epsilon(k)\hat{f}(k) \to \exp(-ika)\hat{f}(k)$.

**Theorem 28.10** *If $f$ is in $L^2(\mathbb{R}, dx)$ and if $f'$ exists (in the sense that $f$ is an integral of $f$) and if $f'$ is also in $L^2(\mathbb{R}, dx)$, then the Fourier transform is in $L^1(\mathbb{R}, dk/(2\pi))$. As a consequence $f$ is is $C_0(\mathbb{R})$.*

Proof: $\hat{f}(k) = (1/\sqrt{1+k^2}) \cdot \sqrt{1+k^2}\hat{f}(k)$. Since $f$ is in $L^2$, it follows that $\hat{f}(k)$ is in $L^2$. Since $f'$ is in $L^2$, it follows that $k\hat{f}(k)$ is in $L^2$. Hence $\sqrt{1+k^2}\hat{f}(k)$ is in $L^2$. Since $1/\sqrt{1+k^2}$ is also in $L^2$, it follows from the Schwarz inequality that $\hat{f}(k)$ is in $L^1$. □

## 28.5 Fourier transform pairs

There are some famous Fourier transforms. Fix $\sigma > 0$, and consider first the Gaussian

$$g_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2}{2\sigma^2}). \tag{28.16}$$

Its Fourier transform is similar; it is the *Gauss kernel*

$$\hat{g}_\sigma(k) = \exp(-\frac{\sigma^2 k^2}{2}). \tag{28.17}$$

Here is a proof of this Gaussian formula. Define the Fourier transform $\hat{g}_\sigma(k)$ by the usual formula. Check that

$$\left(\frac{d}{dk} + \sigma^2 k\right)\hat{g}_\sigma(k) = 0. \tag{28.18}$$

This proves that

$$\hat{g}_\sigma(k) = C\exp(-\frac{\sigma^2 k^2}{2}). \tag{28.19}$$

Now apply the equality of $L^2$ norms. This implies that $C^2 = 1$. By looking at the case $k = 0$ it becomes obvious that $C = 1$.

Let $\epsilon > 0$. Introduce the Heaviside function $H(k)$ that is 1 for $k > 0$ and 0 for $k < 0$. The two basic Fourier transform pairs are

$$f_\epsilon(x) = \frac{1}{x - i\epsilon} \tag{28.20}$$

with Fourier transform

$$\hat{f}_\epsilon(k) = 2\pi i H(-k) e^{\epsilon k}. \tag{28.21}$$

and its complex conjugate

$$\overline{f_\epsilon(x)} = \frac{1}{x + i\epsilon} \tag{28.22}$$

with Fourier transform

$$\overline{\hat{f}_\epsilon(-k)} = -2\pi i H(k) e^{-\epsilon k}. \tag{28.23}$$

These may be checked by computing the inverse Fourier transform. Notice that $f_\epsilon$ and its conjugate are not in $L^1(R)$.

Take $1/\pi$ times the imaginary part. This gives the approximate delta function given by the *Poisson kernel* on the line given by

$$\delta_\epsilon(x) = \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2}. \tag{28.24}$$

with Fourier transform

$$\hat{\delta}_\epsilon(k) = e^{-\epsilon|k|}. \tag{28.25}$$

Instead take the real part. This gives the approximate *principal value* of $1/x$ function

$$p_\epsilon(x) = \frac{x}{x^2 + \epsilon^2} \tag{28.26}$$

with Fourier transform

$$\hat{p}_\epsilon(k) = -\pi i [H(k) e^{-\epsilon k} - H(-k) e^{\epsilon k}]. \tag{28.27}$$

## 28.6 Supplement: Poisson summation formula

The classical setting for the *Poisson summation formula* begins with the group $\mathbb{R}$ and its dual group, which is also $\mathbb{R}$. However there is also a specified discrete discrete subgroup $\mathbb{Z}\,\Delta x$ and a corresponding quotient group, which is a circle of circumference $\Delta x$. Finally, there is the dual group of this quotient group, which is $\mathbb{Z}\,\Delta k$, where $\Delta k\,\Delta x = 2\pi$.

**Theorem 28.11 (Poisson summation formula)** *Let $f$ be in $L^1(\mathbb{R}, dx)$ with $\hat{f}$ in $L^1(\mathbb{R}, dk/(2\pi))$ and such that $\sum_\ell |\hat{f}(\ell\,\Delta k)| < \infty$. Then*

$$\sum_{j \in \mathbf{Z}} f(j\,\Delta x)\,\Delta x = \sum_{\ell \in \mathbf{Z}} \hat{f}(\ell\,\Delta k). \tag{28.28}$$

Proof: Let

$$h(x) = \sum_j f(x + j\,\Delta x). \tag{28.29}$$

Since $h(x)$ is periodic with period $\Delta x$, we can expand

$$h(x) = \sum_\ell a_\ell e^{i\ell\, Deltakx}. \tag{28.30}$$

It is easy to compute that

$$a_\ell = \frac{1}{\Delta x} \int_0^{\Delta x} e^{-i\ell\,\Delta k\,y} f(y)\,dy = \hat{f}(\ell \Delta k). \tag{28.31}$$

So the Fourier series of $h(x)$ is absolutely summable, and hence converges pointwise to a continuous function. This gives the representation

$$\sum_j f(x + j\,\Delta x)\,\Delta x = \sum \hat{f}(\ell \Delta k) e^{i\ell\,\Delta k\,x} \tag{28.32}$$

Take $x = 0$ to get the formula as stated. □

Say that one is interested in the Fourier transform $\hat{f}(\alpha)$ at angular frequency $\alpha$. One might use the Poisson summation formula with a given $\Delta x$ to compute a discrete Riemann sum approximation to this Fourier transform. The following corollary shows that the result is a sum of the Fourier transforms not only at $\alpha$, but also at all "alias" angular frequencies $\alpha + \ell\Delta k$.

**Corollary 28.12** *If the Poisson summation formula is applied to $f(x)e^{-i\alpha x}$ with Fourier transform $\hat{f}(k + \alpha)$, then it becomes*

$$\sum_{j \in \mathbf{Z}} f(j\; Deltax) e^{-i\alpha j \Delta x}\,\Delta x = \sum_{\ell \in \mathbf{Z}} \hat{f}(\alpha + \ell\,\Delta k). \tag{28.33}$$

# Problems

1. Let $f(x) = 1/(2a)$ for $-a \le x \le a$ and be zero elsewhere. Find the $L^1(\mathbb{R}, dx)$, $L^2(\mathbb{R}, dx)$, and $L^\infty(\mathbb{R}, dx)$ norms of $f$, and compare them.

2. Find the Fourier transform of $f$.

3. Find the $L^\infty(\mathbb{R}, dk/(2\pi))$, $L^2(\mathbb{R}, dk/(2\pi))$, and $L^1(\mathbb{R}, dk/(2\pi))$ norms of the Fourier transform, and compare them.

4. Compare the $L^\infty(\mathbb{R}, dk/(2\pi))$ and $L^1(\mathbb{R}, dx)$ norms for this problem. Compare the $L^\infty(\mathbb{R}, dx)$ and $L^1(\mathbb{R}, dk/(2\pi))$ norms for this problem.

5. Use the pointwise convergence at $x = 0$ to evaluate an improper integral.

6. Calculate the convolution of $f$ with itself.

7. Find the Fourier transform of the convolution of $f$ with itself. Verify in this case that the Fourier transform of the convolution is the product of the Fourier transforms.

8. In this problem the Fourier transform is band-limited, that is, only waves with $|k| \leq a$ have non-zero amplitude. Make the assumption that $|k| > a$ implies $\hat{f}(k) = 0$. That is, the Fourier transform of $f$ vanishes outside of the interval $[-a, a]$.

   Let
   $$g(x) = \frac{\sin(ax)}{ax}. \tag{28.34}$$
   The problem is to prove that
   $$f(x) = \sum_{m=-\infty}^{\infty} f(\frac{m\pi}{a})g(x - \frac{m\pi}{a}). \tag{28.35}$$
   This says that if you know $f$ at multiples of $\pi/a$, then you know $f$ at all points.

   Hint: Let $g_m(x) = g(x - m\pi/a)$. The task is to prove that $f(x) = \sum_m c_m g_m(x)$ with $c_m = f(m\pi/a)$. It helps to use the Fourier transform of these functions. First prove that the Fourier transform of $g(x)$ is given by $\hat{g}(k) = \pi/a$ for $|k| \leq a$ and $\hat{g}(k) = 0$ for $|k| > a$. (Actually, it may be easier to deal with the inverse Fourier transform.) Then prove that $\hat{g}_m(k) = \exp(-im\pi k/a)\hat{g}(k)$. Finally, note that the functions $\hat{g}_m(k)$ are orthogonal.

9. In the theory of neural networks one wants to synthesize an arbitrary function from linear combinations of translates of a fixed function. Let $f$ be a function in $L^2$. Suppose that the Fourier transform $\hat{f}(k) \neq 0$ for all $k$. Define the translate $f_a$ by $f_a(x) = f(x - a)$. The task is to show that the set of all linear combinations of the functions $f_a$, where $a$ ranges over all real numbers, is dense in $L^2$.

   Hint: It is sufficient to show that if $g$ is in $L^2$ with $(g, f_a) = 0$ for all $a$, then $g = 0$. (Why is this sufficient?) This can be done using Fourier analysis.

10. Let $\epsilon > 0$. Prove that
   $$\frac{1}{\pi}\sum_{n=-\infty}^{\infty} \frac{\epsilon}{\epsilon^2 + n^2} = \frac{2}{1 - e^{-2\pi\epsilon}} - 1. \tag{28.36}$$
   Hint: Apply the Poisson summation formula to $(1/\pi)\epsilon/(\epsilon^2 + x^2)$ with $\Delta x = 1$. Sum explicitly over angular frequencies.

11. Prove that
   $$1 = \sum_{m=-\infty}^{\infty} \frac{\sin^2(\pi\alpha)}{\pi^2(m + \alpha)^2}. \tag{28.37}$$

Does this formula make sense when $\alpha$ is an integer? Hint: Let $g = 1/(2\pi)$ on $[-\pi, \pi]$ and be zero otherwise. Apply the corollary to the Poisson summation formula to $f = g * g$ with $\Delta x = 2\pi$. Perform an explicit sum on one side.

# Part VII

# Topology and Measure

# Chapter 29

# Topology

## 29.1 Topological spaces

Let $X$ be a set. The power set $P(X)$ consists of all subsets of $X$. In the following we shall fix attention on a universe $X$ of points and certain subsets $U \subset X$. Often we shall want to speak of sets of subsets. For clarity, we shall often speak instead of collections of subsets. Thus a collection is a subset of the power set $P(X)$.

Let $\Gamma \subset P(X)$ be a collection of sets. Recall the definitions of union and intersection:

$$\bigcup \Gamma = \{x \in X \mid \exists U (U \in \Gamma \wedge x \in U)\} \tag{29.1}$$

and

$$\bigcap \Gamma = \{x \in X \mid \forall U (U \in \Gamma \Rightarrow x \in U)\}. \tag{29.2}$$

Thus the union and intersection are each a subset of $X$.

A *topology* on $X$ is a subcollection $\mathcal{T}$ of $P(X)$ with the following two properties:

1. If $\Gamma \subset \mathcal{T}$, then $\bigcup \Gamma \in \mathcal{T}$.

2. If $\Gamma \subset \mathcal{T}$ is finite, then $\bigcap \Gamma \in \mathcal{T}$.

The structure $X, \mathcal{T}$ consisting of a set $X$ with a given topology $\mathcal{T}$ is called a *topological space*. When the topology under consideration is clear from context, then the topological space is often referred to by its underlying set $X$.

It follows from the first property that $\bigcup \emptyset = \emptyset \in \mathcal{T}$. It follows from the second property that $\bigcap \emptyset = X \in \mathcal{T}$. (The fact that $\bigcap \emptyset = X$ follows from the convention that for $\Gamma \subset P(X)$ the universe is $X$.) An *open subset* is a subset that is in the topology. A *closed subset* is a subset that is the complement of an open set.

The *interior* $\operatorname{int} S$ of a subset $S$ is the union of all open subsets of it. It is the largest open subset of $S$. A point is in the interior of $S$ iff it belongs to an open subset of $S$. The *closure* $\bar{S}$ of a subset $S$ is the intersection of all closed supersets. It is the smallest closed superset of $S$. A point is in the closure of $S$ if and only if every open set to which it belongs intersects $S$ in at least one point.

Let $X$ and $Y$ each have a topology. A *continuous map* $f : X \to Y$ is a function such that for every open subset $V$ of $Y$ the inverse image $f^{-1}[V]$ is an open subset of $X$.

There is an alternate characterization of continuous maps that is often useful. A function $f : X \to Y$ is continuous if and only if for every closed subset $F$ of $Y$ the inverse image $f^{-1}[F]$ is a closed subset of $X$. This is a useful fact; it often used to show that the solutions of an equation form a closed set.

If $f : X \to Y$ is a continuous bijection with continuous inverse, then $f$ is a *topological isomorphism* or *topological equivalence* or *homeomorphism*.

Examples:

1. The open unit interval $(0, 1)$ is homeomorphic to $\mathbb{R}$.
2. The open unit interval $(0, 1)$ is not homeomorphic to the circle $S_1$. (One is compact; the other is not.)
3. The closed unit interval $[0, 1]$ is not homeomorphic to the circle $S_1$. (One can be disconnected by removing a point; the other not.)
4. The sphere $S_{n-1}$ is not homeomorphic to $\mathbb{R}^m$.
5. There can be surprises. The unit sphere in the Hilbert space $\ell^2$ is homeomorphic to $\ell^2$ [2].

It is also useful to have a definition of continuity at a point. We say that $f$ is *continuous at the point* $x$ if for each open subset $V$ with $f(x) \in V$ there is an open subset $U$ with $x \in U$ and $f[U] \subset V$. In using this definition it may be helpful to recall that $f[U] \subset V$ is equivalent to $U \subset f^{-1}[V]$.

Two topologies on the same space may sometimes be compared. If every open set in the first topology is an open set in the second topology, then the first topology is said to be *coarser* (or smaller) and the second topology is said to be *finer* (or larger). The finest possible topology on a set $X$ is the *discrete topology*, for which every subset is open. The coarsest possible topology on a set is the *indiscrete topology*, for which only the empty subset $\emptyset$ and $X$ itself are open subsets.

If $\Gamma \subset P(Y)$ is an arbitrary collection of subsets of $X$, then there is a least (coarsest) topology $\mathcal{T}$ with $\Gamma \subset \mathcal{T}$. This is the topology *generated* by $\Gamma$. It may be denoted by $\mathrm{top}(\Gamma)$.

For example, say that $\Gamma = \{U, V\}$, where $U \subset Y$ and $V \subset Y$. Then the topology generated by $\Gamma$ is $\mathcal{T} = \mathrm{top}(\Gamma) = \{\emptyset, U \cap V, U, V, U \cup V, Y\}$ and can have up to 6 subsets in it.

**Proposition 29.1** *Say that $X$ has topology $\mathcal{S}$ and $Y$ has topology $\mathcal{T}$. Suppose also that $\Gamma$ generates $\mathcal{T}$. Suppose that $f : X \to Y$ and that for every $V$ in the* *generating set* $\Gamma$ *the inverse image* $f^{-1}[V]$ *is an open set in* $\mathcal{S}$*. Then* $f$ *is a continuous map.*

An example where this applies is when $Y$ is a metric space. It says that in this case it is enough to check that the inverse images of open balls are open.

If $Y$ is a topological space, and if $Z$ is a subset of $Y$, then there is a *relative topology* induced on $Z$, so that $Z$ becomes a *subspace* of $X$. It is defined as the collection of all $U \cap Z$ for $U \subset Y$ open. If we define the map $i : Z \to Y$ to be the injection, then the topology on $Z$ is the coarsest topology that makes $i$ continuous.

The most common way of defining a topology space is to take some well known space, such as $Y = \mathbb{R}^n$ and then indicate some subset $Z \subset Y$. Even though the topology of $Z$ is the relative topology derived from $Y$, it is important that one can forget about this and think of $Z$ as a topological space that is a universe with its own topology.

As an example, take the case when $Y = \mathbb{R}$ and $\mathbf{Z} = [0, 1]$. Then a set like $[0, 1/2)$ is an open subset of $\mathbf{Z}$, even though it is not an open subset of $\mathbb{R}$. The reason is that $[0, 1/2)$ is the intersection of $(-2, 1/2)$ with $Z$, and $(-2, 1/2)$ is an open subset of $\mathbb{R}$.

If $f : X \to Y$ is a continuous injection that gives a homeomorphism of $f$ with $Z \subset Y$, where $Z$ has the relative topology, then $f$ is said to be an *embedding* of $X$ into $Y$. Thus $f = i \circ h$, where $h : X \to Z$ is a homeomorphism.

Examples:

1. There is an embedding of the open interval $(0, 1)$ into the circle $S_1$. The range of the embedding is the circle with a single point removed.

2. There is no embedding of the circle into the open interval. A continuous image of the circle in the open interval is compact and connected, and thus is a closed subinterval. But a circle is not homeomorphic to a closed interval.

If $X$ is a topological space, and if $\Gamma$ is a partition of $X$, then there is a *quotient topology* induced on $\Gamma$, so that $\Gamma$ becomes a *quotient space* of $X$. It is defined as the collection of all $V \subset \Gamma$ such that $\bigcup V$ is open in $X$. If we define the map $p : X \to \Gamma$ to map each point onto the subset to which it belongs, then the quotient topology is the finest topology that makes $p$ continuous.

If $f : X \to W$ is a surjection, then the inverse images of points in $W$ produce a partition $\Gamma$ of $X$. If $f : X \to W$ is a continuous surjection that comes from a homeomorphism $h : \Gamma \to W$ with $f = h \circ p$, then $f$ is a way of continuously classifying $X$ into parts, where $W$ is the classification space that indexes the parts.

## 29.2 Comparison of topologies

There are situations in analysis when it is quite natural that there is more than one topology on the same space. A standard example is an infinite dimensional

Hilbert space $H$. The *strong topology* is the topology consisting of all open sets in the usual metric sense. The *weak topology* is the topology generated by all unions of sets of the form $\{u + w \mid u \in U, w \in M^\perp\}$, where $U$ is an open subset of a finite dimensional subspace $M$. Such a set is restricted in finitely many dimensions. It is not hard to see that every open set in the weak topology is an open set in the strong topology. The weak topology is the coarser or smaller topology.

Let $n \mapsto s_n$ be a sequence in the Hilbert space $H$. Then $s_n \to w$ weakly as $n \to \infty$ if and only if for each finite dimensional subspace $M$ with associated orthogonal projection $P_M$ the function $P_M s_n \to P_M w$ as $n \to \infty$. Since finite dimensional projections are given by finite sums involving inner products, this is the same as saying that for each vector $v$ in $H$ the numerical sequence $\langle v, s_n\rangle \to \langle v, w\rangle$ as $n \to \infty$.

It is clear that strong convergence of a sequence implies weak convergence of the sequence. This is because $|\langle v, s_n\rangle - \langle v, w\rangle| = |\langle v, s_n - w\rangle| \leq \|v\|\|s_n - w\|$, by the Schwarz inequality.

The converse is not true. For example, let $n \mapsto e_n$ be a countable orthonormal family. Then $e_n \to 0$ weakly as $n \to \infty$. This is because $\sum_n |\langle v, e_n\rangle|^2 \leq \|v\|^2$. It follows from convergence of this sum that $\langle v, e_n\rangle \to 0$ as $n \to \infty$. However $\|e_n - 0\| = 1$ for all $n$, so there is certainly not strong convergence to zero.

If $\mathcal{T}' \subset \mathcal{T}$, then our terminology is that the topology on $\mathcal{T}'$ is coarser (or smaller), while the topology $\mathcal{T}$ is relatively finer (or larger). Sometime the terms weak and strong are used, but this takes some care, as is shown by the following two propositions.

**Proposition 29.2** *Let $X$ have topology $\mathcal{S}$. If $f : X \to Y$ and $\mathcal{T}' \subset \mathcal{T}$ are topologies on $Y$, then $f : (X, \mathcal{S}) \to (Y, \mathcal{T})$ continuous implies $f : (X, \mathcal{S}) \to (Y, \mathcal{T}')$ continuous.*

The above proposition justifies the use of the word weak to describe the coarser topology on $Y$. Thus strong continuity implies weak continuity for maps into a space.

**Proposition 29.3** *Let $Z$ have topology $\mathcal{U}$. If $f : Y \to Z$ and $\mathcal{T}' \subset \mathcal{T}$ are topologies on $Y$, then $f : (Y, \mathcal{T}') \to (Z, \mathcal{U})$ continuous implies $f : (Y, \mathcal{T}') \to (Z, \mathcal{U})$ continuous.*

The above proposition gives a context when the coarser topology imposes the stronger continuity condition. If we want to continue to use the word weak to describe a coarser topology, then we need to recognize that weak continuity for maps from a space is a more restrictive condition.

As an example, consider again infinite dimensional Hilbert space $H$ with the weak topology. Let $f : H \to \mathbb{R}$ be given by $f(u) = \|u\|^2$. Then $f$ is continuous when $H$ is given the strong topology. However $f$ is not continuous when $H$ is given the weak topology. This may be seen by looking at a sequence $e_n$ that is an orthonormal family. Then $e_n \to 0$ weakly, but $f(e_n) = 1$ for all $n$.

## 29.3 Bases and subbases

A *base* for a topology $\mathcal{T}$ is a collection $\Gamma$ of open sets such that every open set $V$ in $\mathcal{T}$ is the union of some subcollection of $\Gamma$.

Let $X$ be a topological space and let $\Gamma$ be a collection of open subsets. Then $\Gamma$ is a *subbase* if the collection $\tilde{\Gamma}$ of all intersections of finite subcollections of $\Gamma$ is a base. Notice that according to the convention $\bigcap \emptyset = X$, the set $X$ automatically belongs to $\tilde{\Gamma}$.

**Theorem 29.4** *Let $X$ be a set. Let $\Gamma$ be a collection of subsets. Then there is a coarsest topology $\mathcal{T}$ including $\Gamma$, and $\Gamma$ is a subbase for $\mathcal{T}$.*

Proof: Let $\Gamma$ be a collection of subsets of $X$. Let $\tilde{\Gamma}$ be the collection consisting of intersections of finite subsets of $\Gamma$. Let $\mathcal{T}$ be the collection consisting of unions of subsets of $\tilde{\Gamma}$. The task is to show that $\mathcal{T}$ is a topology.

It is clear that the union of a subcollection of $\mathcal{T}$ is in $\mathcal{T}$. The problem is to show that the intersection of a finite subcollection $\Delta$ of $\mathcal{T}$ is in $\mathcal{T}$. Each $W$ in $\Delta$ is a union of a collection of sets $\mathcal{A}_W \subset \tilde{\Gamma}$. By the distributive law the intersection is

$$\bigcap_{W \in \Delta} \bigcup \mathcal{A}_W = \bigcup_s \bigcap_{W \in \Delta} s(W). \tag{29.3}$$

Here $s$ is summed over all possible selection functions with the property that $s(W)$ is in $\mathcal{A}_W$ for each $W$. Since each $s(W)$ is a finite intersection of sets in $\Gamma$, it follows that each $\bigcap_{W \in \Delta} s(W)$ is a finite intersection of sets in $\Gamma$. Thus the finite intersection is a union of such finite intersections. Thus it is in $\mathcal{T}$. $\square$

**Proposition 29.5** *Let $f : X \to Y$. Suppose that for each $V \subset Y$ in a subbase the set $f^{-1}[V]$ is open in $X$. Then $f$ is continuous.*

A *neighborhood base* for a point $x$ in a topological space is a family $\Gamma_x$ of open sets $V$ with $x \in V$ such that for every open set $U$ with $x \in U$ is there is a $V$ in $\Gamma_x$ with $V \subset U$.

A *neighborhood subbase* for a point $x$ in a topological space is a family $\Gamma_x$ of open sets $V$ with $x \in V$ such that for every open set $U$ with $x \in U$ is there is a $V$ that is a finite intersection of sets in $\Gamma_x$ with $V \subset U$.

**Proposition 29.6** *Let $f : X \to Y$. Suppose that for each $V$ in a neighborhood subbase of $f(x)$ there is an open subset $U$ with $x \in U$ and $f[U] \subset V$. Then $f$ is continuous at $x$. On the other hand, suppose that $f$ is continuous at $x$ and that $\Gamma_x$ is a neighborhood base for $x$. Then for each open subset $V$ with $f(x) \in V$ there is a $U$ in $\Gamma_x$ with $x \in U$ and $f[U] \subset V$.*

Again it is good to recall that $f[U] \subset V$ is always equivalent to $U \subset f^{-1}[V]$.

A topological space is *first countable* if every point has a countable neighborhood base. This is the same as having a countable neighborhood subbase. A metric space is first countable. A neighborhood base at $x$ consists of the open balls centered at $x$ of radius $1/n$, for $n = 1, 2, 3, \ldots$. So being close to $x$ is determined by countably many conditions.

A topological space is *second countable* provided that it has a countable base. This is the same as having a countable subbase.

If $X$ is a topological space and $S$ is a subset, then $S$ is *dense* in $X$ if its closure is $X$.

A topological space $X$ is *separable* provided that there is a countable subset $S$ with closure $\bar{S} = X$. In other words, $X$ is separable if it has a countable dense subset.

**Theorem 29.7** *If $X$ is second countable, then $X$ is separable.*

Proof: Let $\Gamma$ be a countable base for $X$. Let $\Gamma' = \Gamma \setminus \{\emptyset\}$. Then $\Gamma'$ consists of non-empty sets. For each $U$ in $\Gamma'$ choose $x$ in $U$. Let $S$ be the set of all such $x$. Let $V = X \setminus \bar{S}$. Since $V$ is open, it is the union of those of its subsets that belong to $\Gamma$. Either there are no such subsets, or there is only the empty set. In either case, it must be that $V = \emptyset$. This proves that $\bar{S} = X$. □

It is also not very difficult to prove that a separable metric space is second countable. It is not true in general that a separable topological space is second countable. For a topological space the more useful notion is that of being second countable.

## 29.4 Compact spaces

A topological space $K$ is *compact* if whenever $\Gamma$ is a collection of open sets with $K = \bigcup \Gamma$, then there is a finite subcollection $\Gamma_0 \subset \Gamma$ with $K = \bigcup \Gamma_0$. This can be summarized in a slogan: Every open cover has a finite subcover.

Sometimes one wants to apply this definition to a subset $K$ of a topological space $X$. Then it is customary to say that $K$ is compact if and only if whenever $\Gamma$ is a collection of open subsets of $X$ with $K \subset \bigcup \Gamma$, then there is a finite subcollection $\Gamma_0 \subset \Gamma$ with $K \subset \bigcup \Gamma_0$. Again: Every open cover has a finite subcover. However this is just the same as saying that $K$ itself is compact with the relative topology.

There is a dual formulation in terms of closed subsets. A topological space $X$ is compact if whenever $\Gamma$ is a collection of closed sets with $\bigcap \Gamma = \emptyset$, then there is a finite subcollection $\Gamma_0 \subset \Gamma$ with $\bigcap \Gamma_0 = \emptyset$.

A collection of sets $\Gamma$ has the *finite intersection property* provided that for every finite subcollection $\Gamma_0 \subset \Gamma$ we have $\bigcap \Gamma_0 \neq \emptyset$.

**Proposition 29.8** *A topological space is compact if and only if every collection $\Gamma$ of closed subsets with the finite intersection property has $\bigcap \Gamma \neq 0$.*

Again there could be a compactness slogan: Every collection of closed sets with the finite intersection property has a common point.

**Corollary 29.9** *A topological space is compact if and only if for every collection $\Gamma$ of subsets with the finite intersection property there is a point $x$ that is in the closure of each of the sets in $\Gamma$.*

Perhaps the compactness slogan could be: For every collection of sets with the finite intersection property there exists a point near each set.

**Proposition 29.10** *If $K$ is compact, $F \subset K$, and $F$ is closed, then $F$ is compact.*

Proof: Let $\Gamma$ be a collection of closed subsets of $F$ with the finite intersection property. Since $F$ is closed, each set in $\Gamma$ is also a closed subset of $K$. Since $K$ is compact, there is a point $p$ in each set in $\Gamma$. This is enough to prove that $F$ is compact. $\Box$

**Theorem 29.11** *Let $f : K \to L$ be a continuous surjection from $K$ onto $L$. If $K$ is compact, then $L$ is compact.*

Proof: Let $\Delta$ be an open cover of $L$. Then the inverse images under $f$ of the sets in $\Delta$ form an open cover of $K$. However $K$ is compact. Therefore there exists a finite subset $\Delta_0$ of $\Delta$ such that the inverse images of the sets in $\Delta_0$ form an open cover of $K$. Since $f$ is a surjection, every point $y$ in $L$ is the image of a point $x$ in $K$. There is an open set $V$ in $\Delta_0$ such that that $x$ is in the inverse image of $V$. It follows that $y$ is in $V$. This proves that $\Delta_0$ is an open cover of $V$. It follows that $L$ is compact. $\Box$

A topological space is *Hausdorff* provided that for each pair of points $x, y$ in the space there are open subsets $U, V$ with $x \in U$, $y \in V$, and $U \cap V = \emptyset$.

**Proposition 29.12** *Each compact subset $K$ of a Hausdorff space is closed.*

Proof: Let $X$ be a Hausdorff space and $K \subset X$ a compact subset.

Fix $y \notin K$. For each $x \in K$ choose $U_x$ and $V_x$ with $x \in U_x$ and $y \in V_x$ and $U_x \cap V_x = \emptyset$. The union of the $U_x$ for $x$ in $K$ includes $K$, so there is a finite subset $S$ of $K$ such that the union of the $U_x$ for $x$ in $S$ included $K$. Let $V$ be the intersection of the $V_x$ for $x$ in $S$. Then $V$ is open and $y \in V$ and $V \cap K = \emptyset$.

Choose for every $y$ in $X \setminus K$ an open set $V_y$ with $y \in V_y$ and $V \cap K = \emptyset$. Then $X \setminus K$ is the union of the $V_y$ with $y$ in $X \setminus K$. This proves that $X \setminus K$ is open, and so $K$ is closed. $\Box$

## 29.5 The one-point compactification

The *one-point compactification* is a construction that works for arbitrary topological spaces, but it will turn out that it is only useful for locally compact Hausdorff spaces.

**Theorem 29.13 (one-point compactification)** *Let $X$ be a topological space that is not compact. Then there exists a topological space $X^*$ with one extra point that is compact and such that $X$ is a subspace of $X^*$ with the induced topology.*

Proof: Let $\infty$ be a point that is not in $X$, and let $X^* = X \cup \{\infty\}$. The topology for $X^*$ is defined as follows. There are two kinds of open sets of $X^*$. If $\infty \notin U$, then $U$ is open if and only if $U$ is an open set in the topology of $X$. If $\infty \in U$, then $U$ is open if and only if $U$ is the complement of a closed compact subset $K$ of $X$. It is clear that the topology of $X$ is the relative topology as a subspace of $X^*$.

Consider an open cover of $X^*$. This is a collection of open subsets of $X^*$ whose union is $X^*$, so there must be at least one open subset that is the complement of a compact closed subset $K$ of $X$. The union of the remaining open subsets in the cover includes $K$. These open sets can be of two kinds. Some of them may be open subsets of $X$. The other are complements of closed subsets of $X$, so their intersections with $X$ are open subsets of $X$. These subsets provide an open cover of $K$, so they have a finite subcover of $K$. This shows that the one open set whose complement is $K$ together with the remaining finite collection of open sets that cover $K$ form an open cover of $X^*$. This proves that $X^*$ is compact. $\square$

A topological space $X$ is *locally compact* if and only if for each point $p$ in $X$ there exists an open set $U$ and a compact set $K$ with $p \in U \subset K$.

**Theorem 29.14 (one-point Hausdorff compactification)** *Let $X$ be a topological space. Then its one-point compactification $X^*$ is Hausdorff if and only if $X$ is both locally compact and Hausdorff.*

Proof: Here is a sketch of the fact that $X$ locally compact Hausdorff implies $X^*$ Hausdorff. The Hausdorff property says that each pair of distinct points are separated by open sets. The separation is clear for two points that are subsets of $X$. The interesting case is when one-point is $p \in X$ and the other point is $\infty$. Then by local compactness there exists an open subset $U$ of $X$ such that $p \in U \subset K$, where $K$ is a compact subset of $X$. Since $X$ is Hausdorff, $K$ is also closed. Let $V$ be the complement of $K$ in $X^*$. Then $\infty \in V$ and $p \in U$, both $U$ and $V$ are open in $X^*$, and the two open sets are disjoint. Thus $p$ and $\infty$ are separated by open sets. $\square$

Examples:

1. $\mathbb{R}^n$ is locally compact and Hausdorff. Its one-point compactification is homeomorphic to a sphere $S_n$.

2. Consider an infinite dimensional real Hilbert space $H$, for example $\ell^2$. It is a metric space and so is Hausdorff. However it is not locally compact. In fact, an open ball is not totally bounded, so it cannot be a subset of a compact set. The one-point compactification of $H$ is not even Hausdorff.

3. What if we give $H$ the weak topology? As we shall see, then each closed ball is compact. But the space is still not locally compact, since there are no non-empty weakly open sets inside the closed unit ball. The non-empty weakly open sets are all unbounded.

The case of $\mathbb{R}$ and $\mathbb{C}$ are particularly interesting. From the point of analytic function theory, the one-point compactification is quite natural. The one-point compactification of $\mathbb{R}$ is a circle, and the one-point compactification of $\mathbb{C}$ is the Riemann sphere.

On the other hand, $\mathbb{R}$ has an order structure, and in this context it is more natural to look at a two point compactification $[-\infty, +\infty]$. This of course is not homeomorphic to a circle, but instead to an interval such as $[0, 1]$.

## 29.6 Metric spaces and topological spaces

Among topological spaces metric spaces are particularly nice. This section is an attempt to explain the special role of metric spaces. This begins with the $T_1$, $T_2$, $T_3$, and $T_4$ properties.

A topological space is $T_1$ if for every pair of points there is an open set with the first point not in it and the second point in it. This is equivalent to the condition that single point sets are closed sets.

A topological space is *Hausdorff* or $T_2$ if every pair of points is separated by a pair of disjoint open sets.

A topological space is *regular* if it is $T_1$ and also satisfies condition $T_3$: every pair consisting of a closed set and a point not in the set is separated by a pair of disjoint open sets.

A topological space is *normal* if it is $T_1$ and satisfies condition $T_4$: every pair of disjoint closed sets is separated by a pair of disjoint open sets.

It the chapter on metric spaces, it was shown that given two disjoint closed sets, there is a continuous real function with values in $[0, 1]$ that is zero on one set and one on the other set. As a consequence, every metric space is normal.

The following theorems clarify the question of when a topological space is metrizable. They are stated here without proof.

**Theorem 29.15** *A compact Hausdorff space is normal.*

**Theorem 29.16** *A locally compact Hausdorff space is regular.*

Recall that in a topological space second countable implies separable. The converse is true for metric spaces.

**Theorem 29.17 (Urysohn)** *A second countable regular space is metrizable.*

**Corollary 29.18** *A second countable locally compact Hausdorff space is metrizable.*

## 29.7 Topological spaces and measurable spaces

The interaction of measure and topology can encounter technical difficulties. Their origin is the following. A topological space is characterized by open sets allowing uncountable unions and finite intersections. A measurable space is characterized by measurable sets allowing countable unions, countable intersections, and complements. The tension arises from a situation when the uncountable operations for a topological space enter the measure theory. We review some of these issues, mainly to point out that there are many situations when they do not arise.

First, we recall that for every $\sigma$-algebra of subsets, there is a corresponding $\sigma$-algebra of real functions, consisting of all real functions measurable with respect to the $\sigma$-algebra of subsets. Conversely, for every $\sigma$-algebra of real functions, there is a corresponding $\sigma$-algebra of subsets. So we can think of either kind of $\sigma$-algebra; they are equivalent.

If $X$ is a topological space, then it determines a measurable space by taking the *Borel $\sigma$-algebra of subsets*. This is the smallest $\sigma$-algebra $\mathcal{B}o$ that contains all the open sets of the topological space. Since it is closed under complements, it also contains all the closed sets.

If $X$ is a measurable space and $Z \subset X$ is a subset, then there is a natural structure of measurable space on $Z$. This is the relative $\sigma$-algebra consisting of all the intersections of measurable subsets of $X$ with $Z$. Furthermore, if $X$ is a topological space, and $Z \subset X$ is a subset, then there is a natural structure of topological space on $Z$. This is the relative topology consisting of all the intersections of open sets of $X$ with $Z$. The relative $\sigma$-algebra on $Z$ induced by the Borel $\sigma$-algebra on $X$ is the same as the Borel $\sigma$-algebra on $Z$ generated by the relative topology on $Z$.

The situation for the product of two topological spaces is the following. The product $\sigma$-algebra of two Borel $\sigma$-algebras is always contained in the Borel $\sigma$-algebra of the product space. However when the original two topological spaces are second-countable, then the product $\sigma$-algebra coincides with the Borel $\sigma$-algebra.

Next we look at the $\sigma$-algebra generated by $C(X)$. This is the same as the $\sigma$-algebra generated by $BC(X)$. (Every function is $C(X)$ is a continuous function of a function in $BC(X)$.) However, in general this can be smaller than the Borel $\sigma$-algebra.

**Theorem 29.19** *If $X$ is a metrizable topological space, then the space $C(X)$ generates the Borel $\sigma$-algebra $\mathcal{B}o$.*

Proof: To prove this, it is sufficient to show that every closed set is in the inverse image of a Borel subset under some continuous function. Let $F$ be a closed subset. Then the function $f(x) = d(x, F)$ is a continuous function that vanishes precisely on $F$. That is, the inverse image of $\{0\}$ is $F$. □

## 29.8 Supplement: Ordered sets and topological spaces

The following topic is optional; it gives a brief introduction to nets, which are maps from a certain kind of ordered set to a topological space. It also illustrates a way of associating a topology to an ordered set, so that a convergent net turns out to be a map that is continuous at infinity.

In a metric space the notion of sequence is important, because most topological properties may be characterized in terms of convergence of sequences. In more general spaces sequences are not enough to characterize convergence. However the more general notion of net does the job.

A *directed set* is an ordered set $I$ with the property that every finite non-empty subset has an upper bound. For general topological spaces it is important that $I$ is not required to be a countable set. Note: Some authors give a definition of directed set that omits the antisymmetry condition on the order.

A *net* in $X$ is a function $w : I \to X$. If $X$ is a topological space, then a net $w$ *converges* to $x$ provided that for every open set $U$ with $x \in U$ there is a $j$ in $I$ such that for all $k$ with $j \leq k$ we have $w_j \in U$.

Examples:

1. A sequence in $X$ is a net in $X$. This is the special case when the directed set is the set of natural numbers.

2. Let $S$ be a set. Let $I$ consist of all finite subsets of $S$. Notice that if $S$ is uncountable, then the index set $I$ is also uncountable. For $H$ and $H'$ in $I$ we write $H \leq H'$ provided that $H \subset H'$. Fix a function $f : S \to \mathbb{R}$. Define the net $H \mapsto \sum_{s \in H} f(s)$ with values in $\mathbb{R}$. If this net converges to a limit, then this limit is a number $\sum f$ that deserves to be called the unordered sum of $f$. The set of $f$ for which such an unordered sum exists is of course just $\ell^1(S)$. It turns out that this example is not so interesting after all, since each $f$ in $\ell^1(S)$ vanishes outside of some countable subset.

3. Here is an example that shows how one can construct a directed set to describe convergence in a general topological setting. Let $x$ be a point in $X$, and define the directed set $I$ to consists of all open sets with $x \in U$. Let $U \leq U'$ mean that $U' \subset U$. Then $I$ is a directed set, since the intersection of finitely many such open sets is open. We shall see in the proof of the next two theorem that this kind of directed set is a rather natural domain for a net.

**Theorem 29.20** *Let $E$ be a subset of $X$. A point $x$ is in the closure $\bar{E}$ if and only if there is a net $w$ with values in $E$ that converges to $x$.*

Proof: First note that the complement of $\bar{F}$ is the largest open set disjoint from $F$. It follows that $x \notin \bar{F}$ is equivalent to $\exists U\, (x \in U \wedge U \cap F = \emptyset$. Here $U$ ranges over open subsets. As a consequence, $x \in \bar{F}$ is equivalent to $\forall U\, (x \in U \Rightarrow U \cap F \neq \emptyset)$.

Suppose $w$ is a net in $E$ that converges to $x$. Let $U$ be an open set in $X$ such that $x \in U$. Then there exists a $j$ so that $w_j \in U$. Hence $U \cap E \neq \emptyset$. Since $U$ is arbitrary, it follows that $x \in \bar{E}$.

Suppose on the other hand that $x$ is in $\bar{E}$. Then for every open set $U$ with $x \in U$ we have $U \cap E \neq \emptyset$. By the axiom of choice there is a point $w_U$ with $w_U \in E$ and $w_U$ in $U$.

Let $I$ consist of the open sets $U$ with $x \in U$. Let $U \leq U'$ provided that $U' \subset U$. Then $I$ is a directed set, since the intersection of finitely many such open sets is an open set. Thus $U \mapsto w_U$ is a net in $E$ that converges to $x$. □

**Theorem 29.21** *A function $f : X \to Y$ is continuous if and only if it maps convergent nets into convergent nets.*

Proof: Suppose $f$ is continuous. Let $w$ be a net in $X$ that converges to $x$. Let $V$ be an open set in $Y$ such that $f(x) \in V$. Let $U = f^{-1}[V]$. Then $x \in U$, so there exists a $j$ such that $j \leq k$ implies $w_k \in U$. Hence $f(w_k) \in V$. This shows that the net $j \mapsto f(w_j)$ converges to $f(x$.

Suppose on the other hand that $f$ maps convergent nets to convergent nets. Suppose that $f$ is not continuous. Then there is an open set $V$ in $Y$ such that $f^{-1}[V]$ is not open.

Next, notice that a set $G$ is open if and only if $\forall x\, x \in G \Rightarrow (\exists U(x \in U \wedge U \subset G)$. Here $U$ ranges over open subsets. This is simply because the a union of open sets is always open. Hence a set $G$ is not open if and only if $\exists x\, x \in G \wedge (\forall U(x \in U \Rightarrow U \setminus G \neq \emptyset)$.

Apply this to the case $G = f^{-1}[V]$. Then there exists an $x$ with $f(x) \in V$ but with the property that for every open set $U$ with $x \in U$ there exists a point $w_U \in U$ with $f(w_U) \notin V$. The existence this function $U \mapsto w_U$ is guaranteed by the axiom of choice.

Let $I$ consist of the open sets $U$ with $x \in U$. Let $U \leq U'$ provided that $U' \subset U$. Then $I$ is a directed set, since the intersection of finitely many such open sets is an open set. Thus $U \mapsto w_U$ is a net in $X$ that converges to $x$. However $U \mapsto f(w_U)$ does not converge to $f(x)$. This is a contradiction. Thus $f$ must be continuous. □

The net language also sheds light on the Hausdorff separation property. It may be shown that a topological space is Hausdorff if and only if every net converges to at most one point.

There is a natural topology on every ordered set, generated by the intervals of the form $I_j = \{k \in I \mid j \leq k\}$. These intervals form a base for the topology. If the ordered set is a directed set, the intersection of two such intervals is never empty; it always includes another interval.

Augment the directed set $I$ with an additional maximal element $\infty$. Define the topology so that the non-empty open subsets of the augmented set are the unions of open intervals of $I$ each augmented with $\{\infty\}$. The fact that $I$ is a directed set implies that this topology has the required finite intersection property.

This gives a topological interpretation to the concept of convergent net. Then a net $w : I \to X$ converges to $x$ if and only if when we augment $w$ by $w(\infty) = x$ we have that $w$ is continuous at $\infty$.

# Problems

1. Let $X = \mathbb{R}$. Show that the sets $(a, +\infty)$ for $a \in \mathbb{R}$, together with the empty set and the whole space, form a topology.

2. Give an example of a continuous function $f : \mathbb{R} \to X$, where $\mathbb{R}$ has the usual metric topology and $X$ has the topology of the preceding problem. Make the example such that $f$ is not a continuous function in the usual sense.

3. What are the compact subsets of $X$ in the example of the first problem?

4. Let $X = \mathbb{R}$. Show that the sets $(a, +\infty)$ for $a \in \mathbb{R}$ together with $(-\infty, b)$ for $b \in \mathbb{R}$ are not a base for a topology. They are a subbase for a topology. Describe this topology.

5. Let $S$ be an infinite set. Consider the topology where the closed subsets are the finite subsets and the set $S$. What are the compact subsets of $S$? Prove that your answer is correct.

6. Consider an infinite-dimensional Hilbert space and a sequence $n \mapsto f_n$ of vectors. Show that if $f_n \to f$ weakly and $\|f_n\| \to \|f\|$ as $n \to \infty$, then $f_n \to f$ strongly as $n \to \infty$. Hint: Look at $\|f_n - f\|^2$.

7. Let $H$ be an infinite dimensional real Hilbert space. Let $n \mapsto e_n$ be a countable orthonormal family indexed by $n = 1, 2, 3, \ldots$.

   (a) Show that $e_k \to 0$ weakly as $k \to \infty$.

   (b) Show that it is false that $ke_k$ converges weakly to zero as $k \to \infty$. Hint: The vector $v = \sum_{n=1}^{\infty} (1/n)e_n$ is in $H$.

   (c) For each $m < n$ let $x_{mn} = e_m + me_n$. Let $X$ be the set of all the vectors $x_{mn}$ for $m < n$. Show that there is no sequence $k \mapsto s_k$ of points in $X$ with $s_k \to 0$ weakly as $k \to \infty$. Hint: If there is such a sequence $s_k = e_{m_k} + m_k e_{n_k}$, then there is one for which $k \mapsto e_{m_k}$ and $k \mapsto e_{n_k}$ are orthonormal families and also $k \leq m_k$.

   (d) Show that 0 is in the weak closure of $X$.

8. Show that a collection $\Gamma$ of subsets of $X$ is a base for a topology if and only if it has the following property: if $U$ and $V$ are in $\Gamma$ with $x \in U \cap V$, then there is a $W$ in $\Gamma$ with $x \in W \subset U \cap V$.

9. Let $s : \mathbf{N} \to X$ be a sequence in a metric space $X$. Let $T_n = \{s_k \mid k \geq n\}$. Show that the sets $T_n$ for $n \in \mathbf{N}$ have the finite intersection property. Show that $x$ is in the closure of every set $T_n$ if and only if there is a subsequence that converges to $x$.

10. Lindelöf theorem. Let $X$ be a topological space with a second countable topology $\mathcal{T}$. (Thus there is a countable collection $\Delta \subset \mathcal{T}$ such that every open set in $\mathcal{T}$ is a union of sets in $\Delta$.) If $A$ is a subset of $X$, then every

open cover of $A$ has a countable subcover. (That is, if $\Gamma$ is a collection of open sets with $A \subset \bigcup \Gamma$, then there is a countable subcollection $\Gamma_0$ with $A \subset \bigcup \Gamma_0$.) Prove this result. Hint: Let $\Sigma$ be the collection of all sets in $\Delta$ that are used in a union forming one of the sets in $\Gamma$. Then by definition for each $S$ in $\Sigma$ the collection of sets in $\Gamma$ that use $S$ is non-empty.

11. Generating a topology. Let $X$ be a set and let $\Gamma$ be a collection of subsets of $X$. Prove that the set of all topologies $\mathcal{T}$ with $\Gamma \subset \mathcal{T}$ has a least element. This is the topology $\mathrm{top}(\Gamma)$ generated by $\Gamma$.

12. Generating a $\sigma$-algebra of subsets. Let $X$ be a set and let $\Gamma$ be a collection of subsets of $X$. Prove that the set of all $\sigma$-algebras $\mathcal{F}_X$ of subsets of $X$ with $\Gamma \subset \mathcal{F}_X$ has a least element. This is the $\sigma$-algebra of subsets $\sigma(\Gamma)$ generated by $\Gamma$.

13. Compatibility of topology and measurability. Let $X$ be a set and let $\Gamma$ be a collection of subsets of $X$. Show that if $\mathcal{T} = \mathrm{top}(\Gamma)$ is second countable, then $\mathcal{T} \subset \sigma(\Gamma)$ and hence $\sigma(\mathcal{T}) = \sigma(\Gamma)$. Hint: Let $\mathcal{T}$ be the smallest collection of sets closed under finite intersection and union including $\Gamma$. Let $\mathcal{T}'$ be the smallest collection of sets closed under finite intersection and countable union including $\Gamma$. Show that $\mathcal{T}'$ is closed under finite intersection and union.

14. Give an example to show that in the previous result one cannot dispense with the assumption that the topology is second countable.

15. A topological space is called separable if it has a countable dense subset. Every subspace of a separable metric space is separable.

    Let the unit interval $[0, 1]$ have its usual topology. Consider the product space $X = [0, 1]^{\mathbf{R}}$ consisting of all functions from $\mathbf{R}$ to $[0, 1]$ with the product topology. Then $X$ is a compact Hausdorff space.

    (a) Show that $X$ is separable. Hint: Consider functions each of which has finitely many values. Find a countable set of such functions that is dense in $X$.

    (b) Consider the subspace $A$ of $X$ consisting of indicator functions of single points. Show that $A$ is not separable.

16. Let $w : I \to X$ be a net. If $z : J \to X$ satisfies $z = w \circ \alpha$, then $z$ is called a *subnet* of $w$ provided that the map $\alpha : J \to I$ has the property that $\forall i \exists j\, f[J_j] \subset I_i$. Formulate this property of $\alpha$ in terms of continuity at infinity.

# Chapter 30

# Product and weak* topologies

## 30.1 Introduction

The following sections deal with important compactness theorems. The proofs of these theorems make use of Zorn's lemma. We have seen that the axiom of choice implies Zorn's lemma. It is quite easy to show that Zorn's lemma implies the axiom of choice.

Here is a quick review of Zorn's lemma. Consider a non-empty partially ordered set. Suppose that every non-empty totally ordered subset has an upper bound. Zorn's lemma is the assertion that the set must have a maximal element.

In a sense, Zorn's lemma is an obvious result. Start at some element of the partially ordered set. Take a strictly larger element, then another, then another, and so on. Of course it may be impossible to go on, in which case one already has a maximal element. Otherwise one can go through an infinite sequence of elements. These are totally ordered, so there is an upper bound. Take a strictly larger element, then another, then another, and so on. Again this may generate a continuation of the totally ordered subset, so again there is an upper bound. Continue in this way infinitely many times, if necessary. Then there is again an upper bound. This process is continued as many times as necessary. Eventually one runs out of set. Either one has reached an element from a previous element and there is not a larger element after that. In that case the element that was reached is maximal. Or one runs at some stage through an infinite sequence, and this has an upper bound, and there is nothing larger than this upper bound. In this case the upper bound is maximal.

Notice that this argument involves an incredible number of arbitrary choices. But the basic idea is simple: construct a generalized orbit that is totally ordered. Keep the construction going until a maximal element is reached, either as the result of a previous point in the orbit, or as the result of an previous sequence in the orbit.

## 30.2 The Tychonoff product theorem

Let $A = \prod_{t \in T} A_t$ be a product space. An element $x$ of $A$ is a function from the index set $T$ with the property that for all $t \in T$ we have $x_t \in A_t$.

Suppose that each $A_t$ is a topological space. For each $t$ there is projection $p_t : A \to A_t$ defined by $p_t(x) = x_t$. The *product topology* or *pointwise convergence topology* is the coarsest topology on $A$ such that each individual projection $p_t$ is continuous. Thus if $U \subset A_t$ is open, then $p_t^{-1}[U]$ is an open set in $A$ that has its $t$ component restricted. Furthermore, a finite intersection of such sets is open. So there are open sets that are restricted in finitely many components.

Write the projection of $x$ on the $t$ coordinate as the value $x(t)$ of the function $x$. A net $j \mapsto w_j$ in $A = \prod_{t \in T} A_t$ converges to a point $x$ in $A$ if and only if for each $t$ the net $j \mapsto w_j(t)$ converges to $x(t)$. For this reason the product topology is also called the topology of pointwise convergence. The fundamental result about the product topology is the *Tychonoff product theorem.*

**Theorem 30.1 (Tychonoff product theorem)** *The product of a family of compact spaces is compact.*

Before starting the proof, it is worth looking at an attempt at a proof that does not work. Suppose that for each $t \in T$ the space $A_t$ is compact. Let $\Gamma$ be a collection of closed subsets of $A$ with the finite intersection property. We want to show that $\bigcap \Gamma \neq \emptyset$. This will prove that $A$ is compact.

Fix $t$. Let $\Gamma_t$ be the collection of all projected subsets $F_t = p_t[F]$ for $F \in \Gamma$. Then $\Gamma_t$ has the finite intersection property. Since $A_t$ is compact, there exists an element that belongs to the closure of each $F_t$ in $\Gamma_t$. Choose such an element $x_t \in A_t$ arbitrarily.

Let $x \in A$ be the vector which has components $x_t$. If we could show that $x$ is in each of the $F$ in $\Gamma$, then this would complete the proof. However this where the attempt fails; there is no guarantee that this is so.

If fact, we could take a simple example in the unit square where this does not work. Let the set $\Gamma$ consist of the single set $F = \{(0,1),(1,0)\}$. Then the projection on the first axis is the set $\{0,1\}$ and the second projection is also $\{0,1\}$. If we take $x_1 = 0$ and $x_2 = 0$, then $x = (0,0)$ is far from belonging to $F$. The trouble is that the projections of a set do not do enough to specify the set. The solution is to specify the point in the product space more closely by taking a larger collection of sets with the finite intersection property. For instance, in the example one could take $\Gamma'$ to consist of the set $F$ together with the smaller set $\{(0,1)\}$. Then the projected sets on the first axis have only 0 in their intersection, and the projected sets on the second axis have only 1 in their intersection. So from these one can reconstruct that point $(0,1)$ in the product space that belongs to all the sets in $\Gamma'$ and hence of $\Gamma$.

Notice that this enlarged collection of sets is somewhat arbitrary; one could have made another choice and gotten another point in the product space. However the goal is to single out a point, and the way to do this is to make a maximal specification of the point. One means to accomplish this is to take a maximal collection of sets with the finite intersection property. For instance, one could take all subsets of which $(0, 1)$ is an element. This is an inefficient but sure way of specifying the point $(0, 1)$.

Proof: Suppose that for each $t \in T$ the space $A_t$ is compact. Let $\Gamma$ be a collection of closed subsets of $A$ with the finite intersection property. We want to show that $\bigcap \Gamma \neq \emptyset$. This will prove that $A$ is compact.

Consider all collections of sets with the finite intersection property that include $\Gamma$. By Zorn's lemma, there is a maximal such collection $\Gamma'$.

Fix $t$. Let $\Gamma'_t$ be the set of all projected subsets $F_t = p_t[F]$ for $F \in \Gamma$. Then $\Gamma'_t$ has the finite intersection property. Since $A_t$ is compact, the intersection of the closures of the $F_t$ in $\Gamma'_t$ is non-empty. Choose an element $x_t$ in the closure of each $F_t$ for $F$ in $\Gamma'$.

Let $x$ be the vector which has components $x_t$. Let $U$ be an open set with $x \in U$. Then there is a finite subset $T_0 \subset T$ and an open set $U_t \subset A_t$ for each $t$ in $T_0$ such that the intersection of the sets $p_t^{-1}[U_t]$ for $t \in T_0$ is an open subset of $U$ with $x$ in it.

Consider $t$ in $T_0$. It is clear that $x_t \in U_t$. Since $x_t$ is in the closure of each $F_t$ for each $F$ in $\Gamma'$, it follows that $U_t \cap F_t \neq \emptyset$ for each $F$ in $\Gamma'$. Thus $p_t^{-1}[U_t] \cap F \neq \emptyset$ for each $F$ in $\Gamma'$. Since $\Gamma'$ is maximal with respect to the finite intersection property, it follows that $p_t^{-1}[U_t]$ is in $\Gamma'$.

Now use the fact that $\Gamma'$ has the finite intersection property. Consider $F$ in $\Gamma'$. Since each of the $p_t^{-1}[U_t]$ for $t$ in the finite set $T_0$ is in $\Gamma'$, it follows that the intersection of the $p_t^{-1}[U_t]$ for $t$ in $T_0$ with $F$ is non-empty.

This shows that $U$ has non-empty intersection with each element $F$ of $\Gamma'$. Since $U$ is arbitrary, this proves that $x$ is in the closure of each element $F$ of $\Gamma'$. In particular, $x$ is in the closure of each element $F$ of $\Gamma$. Since $\Gamma$ consists of closed sets, $x$ is in each element $F$ of $\Gamma$. $\square$

## 30.3 Banach spaces and dual Banach spaces

This section is a quick review of the most commonly encountered Banach spaces of functions and of their dual spaces. Let $E$ be a Banach space. Then its *dual space* $E^*$ consists of the continuous linear functions from $E$ to the field of scalars (real or complex). It is also a Banach space. There is a natural injection from $E$ to $E^{**}$. The Banach space $E$ is said to be *reflexive* if this is a bijection.

Let $X$ be a set and $\mathcal{F}$ be a $\sigma$-algebra, so that $X$ is a measurable space. Fix a measure $\mu$. The first examples consist of the Banach spaces $E = L^p(X, \mathcal{F}, \mu)$ for $1 \leq p < \infty$. (In the case $p = 1$ we require that $\mu$ be a $\sigma$-finite measure.) Then the dual space $E^*$ may be identified with $L^q(X, \mathcal{F}, \mu)$, where $1 < q \leq \infty$. Here $1/p + 1/q = 1$. If $u$ is in $E$ and $f$ is in $E^*$, the value of $f$ on $u$ is the integral $\mu(fu)$ of the product of the two functions. If $1 < p < \infty$, then the Banach space $L^p(X, \mathcal{F}, \mu)$ is reflexive. Thus if $1 < q \leq \infty$ the Banach space $L^q(X, \mathcal{F}, \mu)$ is a dual space. In general $L^1(X, \mathcal{F}, \mu)$ is not the dual of another Banach space. This is because the dual of $L^\infty(X, \mathcal{F}, \mu)$ is considerably larger than $L^1(X, \mathcal{F}, \mu)$. These facts are discussed in standard references, such as the text by Dudley [4].

The following examples introduce topology in order to get a Banach space $E$ with a dual space $E^*$ that can play the role of an enlargement of $L^1$. The space $E$ will be a space of continuous functions, while the space $E^*$ is identified with a space of finite signed measures. In order to avoid some measure theoretic technicalities, we shall deal only with continuous functions defined on metric spaces.

Let $X$ be a compact metric space. Then $E = C(X)$ is a real Banach space. The norm of a function in $C(X)$ is the maximum value of its absolute value. Thus convergence in $C(X)$ is uniform convergence on $X$.

The space $C(X)$ of continuous real functions generates a $\sigma$-algebra of functions. These are called Borel functions, and there is a corresponding $\sigma$-algebra of Borel sets. The integrals or measures under consideration are defined on Borel functions or on Borel sets.

There is a concept of *signed measure* and a corresponding theorem. This theorem says that a signed measure always has a as a difference of two measures that live on disjoint subsets, where at least one of the measures must be finite. Sometimes, in the context of signed measures, a measure of the usual kind is called a positive measure. Thus a signed measure is the difference of two positive measures.

A *finite signed measure* is the difference of two finite measures that live on disjoint sets. That is, there are finite measures $\mu_+$ and $\mu_-$ and measurable sets $B_+$ and $B_-$ such that $\mu_+(B_-) = 0$ and $\mu_-(B_+) = -$. The finite signed measure is then $\mu = \mu_+ - \mu_-$.

There is a natural norm for finite signed measures. If $\mu$ is a finite signed measure, then $\|\mu\| = \mu_+(X) + \mu_-(X)$ is the norm.

The term *Riesz representation theorem* is used in several contexts for a theorem that identifies the dual of a Banach space of functions. For instance, the theorem that identifies the dual of $L^p$ as $L^q$ for $1 \leq p < \infty$ and $1/p + 1/q = 1$ is sometimes called a Riesz representation theorem. The theorem for spaces of continuous functions is particularly important.

**Proposition 30.2 (Riesz representation theorem (compact case))** *Let $X$ be a a compact metrizable space. Let $E = C(X)$ be space of continuous real functions on $X$. Then the dual space $E^*$ may be identified with the space of finite signed Borel measures on $X$. That is, each continuous real linear function on $C(X)$ is of the form $f \mapsto \mu(f)$ for a unique finite signed Borel measure on the compact space $X$.*

Remark: A compact metrizable space the same as a second countable compact Hausdorff space.

This result gives an example of a Banach space that is far from being reflexive. That is, the dual of the space $E^*$ of finite signed measures is much larger than the original space $E = C(X)$. In fact, consider an arbitrary bounded Borel function $f$. Then the map $\mu \mapsto \mu(f)$ is continuous on $E^*$ and hence is in $E^{**}$. However it is not necessarily given by an element of $E$.

The results have a useful generalization. Let $X$ be a separable locally compact metric space. Then $C_0(X)$ is a real Banach space. Here $f$ is in $C_0(X)$ provided that for every $\epsilon > 0$ there is a compact subset $K$ of $X$ such that $|f| < \epsilon$ outside of $K$. Such an $f$ is said to vanish at infinity. The space $C_0(X)$ of continuous real functions that vanish at infinity generates the $\sigma$-algebra of Borel functions.

A *Polish space* is a separable completely metrizable space.

**Proposition 30.3 (Riesz representation theorem (locally compact case))** *Let $X$ be a locally compact Polish space. Let $E = C_0(X)$ be the space of continuous real functions on $X$ that vanish at infinity. Then the dual space $E^*$ may be identified with the space of finite signed Borel measures on $X$. That is, each continuous real linear function on $C_0(X)$ is of the form $f \mapsto \mu(f)$ for a unique finite signed Borel measure on the locally compact space $X$.*

Remark: A locally compact Polish space is the same as a second countable locally compact Hausdorff space.

This proposition is only a slight variant on the preceding proposition. Let $X^*$ be the one-point compactification of $X$. Then the space $C_0(X)$ may be thought of as the functions in $C(X^*)$ that vanish at the point $\infty$. Similarly, the finite signed measures $\mu$ on $X$ may be identified with the measures on $X^*$ that assign mass zero to the the set $\{\infty\}$.

## 30.4 Adjoint transformations

If $u$ is in $E$ and $\alpha$ is in the dual space $E^*$, then we sometimes write $\langle \alpha, u \rangle$ instead of $\alpha(u)$. This is not intended to denote an inner product as in the case of a Hilbert space. Rather, it indicates the pairing between $E^*$ and $E$.

Let $T : E \to F$ be a continuous linear transformation from the Banach space $E$ to the Banach space $F$. In this context the value of $T$ on $u$ is often written in the form $Tu$. The Lipschitz norm of $T$ is the smallest number $\|T\|$ such that $\|Tu\| \leq \|T\|\|u\|$. A well-known theorem says that the space of all such mappings $T$ is itself a Banach space. Furthermore, the dual space $E^*$ is just the special case when the mappings are from $E$ to $\mathbb{R}$.

The *adjoint* transformation $T^* : F^* \to E^*$ is defined by $\langle T^*\alpha, u \rangle = \langle \alpha, Tu \rangle$.

If we think of the space as consisting of column vectors and the dual space as consisting of row vectors, then $T$ is like a matrix that acts from the left on column vectors on the right and $T^*$ is the same matrix acting from the right on column vectors on the left.

One can write the definition of adjoint in the more cryptic form $T^*(\alpha) = \alpha \circ T$. This is the same thing, since this just says that $\langle T^*\alpha, u \rangle = \alpha(T(u)) = \langle \alpha, Tu \rangle$. This way of writing reveals that the adjoint is just a special kind of pullback.

## 30.5 Weak* topologies on dual Banach spaces

The *weak topology* on $E$ is the coarsest topology such that every element of $E^*$ is continuous. As a special case, the weak topology on $E^*$ is the coarsest topology such that every element of $E^{**}$ is continuous. The sets $W(f, V) = \{u \in E \mid f(u) \in V\}$, where $f$ is in $E^*$ and $V$ is an open set of scalars, form a subbase for the weak topology of $E$.

The *weak* topology* on $E^*$ is the coarsest topology such that every element of $E$ defines a continuous function on $E^*$. This is the topology of pointwise convergence for the functions in $E^*$. The sets $W(u, V) = \{f \in E^* \mid f(u) \in V\}$, where $u$ is in $E$ and $V$ is an open set of scalars, form a subbase for the weak topology* of $E$.

**Proposition 30.4** *Let $E$ be a Banach space, and let $E^*$ be its dual space. The weak* topology on $E^*$ is coarser than the weak topology on $E^*$. If $E$ is reflexive, then the weak* topology on $E^*$ is the same as the weak topology on $E^*$.*

Proof: Since each element of $E$ defines an element of $E^{**}$, the weak* topology is the coarsest topology that makes all these elements of $E^{**}$ that come from $E$ continuous. The weak topology is defined by requiring that all the elements of $E^{**}$ are continuous. Since more functions have to be continuous, the weak topology is a finer topology. □

Examples:

1. Fix a $\sigma$-finite measure $\mu$. Let $E = L^1$ be the corresponding space of real integrable functions. Then $E^* = L^\infty$. A sequence $f_n$ in $L^\infty$ converges weak* to $f$ if for every $u$ in $L^1$ the integrals $\mu(f_n u) \to \mu(fu)$.

2. Fix a measure $\mu$. Let $E = L^p$ with $1 < p < \infty$. Then $E^* = L^q$ with $1 < q < \infty$. Here $1/p + 1/q = 1$. A sequence $f_n$ in $L^q$ converges weak* to $f$ if for every $u$ in $L^p$ the integrals $\mu(f_n u) \to \mu(fu)$. Since $L^p$ for $1 < p < \infty$ is reflexive, this is the same as weak convergence in $L^q$.

3. Fix a measure $\mu$. Let $E = L^\infty$. Then $E^*$ is an unpleasant space that includes $L^1$ but also has a huge number of unpleasant measure-like objects in it. Notice that there is no weak* topology on $L^1$, since it is not the dual of another Banach space.

4. Consider a compact metric space $X$. Let $E = C(X)$, the space of all continuous real functions on $X$. The norm on $X$ is the supremum norm that describes uniform convergence. Then $E^*$ consists of signed measures. These are of the form $\mu = \mu_+ - \mu_-$, where $\mu_+$ and $\mu_-$ are finite measures. These are the Radon measures that will be described in more detail in a following chapter. The norm of $\mu$ is $\mu_+(X) + \mu_-(X)$. A sequence $\mu_n \to \mu$ in the weak* topology provided that $\mu_n(u) \to \mu(u)$ for each continuous function $u$.

5. Consider a locally compact metric space $X$. Let $E = C_0(X)$, the space of all continuous real functions on $X$ that vanish at infinity. Then $E^*$ again consists of signed measures. In fact, we can think of $E$ as the space of all continuous functions on the one point compactification of $X$ that vanish at the point $\infty$. Then the measures in $E^*$ are those measures on the compactification that assign measure zero to the set $\{\infty\}$. A sequence $\mu_n \to \mu$ in the weak$^*$ topology provided that $\mu_n(u) \to \mu(u)$ for each continuous function $u$ that vanishes at infinity.

Such examples give an idea of the significance of the weak$^*$ topology. The idea is that for $\mu_n$ to be close to $\mu$ in this sense, it is enough that for each observable quantity $u$ the numbers $\mu_n(u)$ get close to $\mu(u)$. The observation is a kind of blurred observation that does not make too many fine distinctions. In the case of measures it is the requirement of continuity that provides the blurring. This allows measures that are absolutely continuous with respect to Lebesgue measure to approach a discrete measure, and it also allows measures that are discrete to approach a measure that is absolutely continuous with respect to Lebesgue measure.

Examples:

1. The measures with density $n1_{[0,1/n]}$ approaches the point mass $\delta_0$. This is absolutely continuous to singular.

2. The singular measures $\frac{1}{n}\sum_{j=1}^{n}\delta_{j/n}$ approach the measure with density $1_{[0,1]}$. This is singular to absolutely continuous.

## 30.6 The Alaoglu theorem

The weak$^*$ topology is the natural setting for main theorem of this section, a compactness result called the *Alaoglu theorem.* The name in this theorem is Turkish; it is pronounced A-la-ō-lu.

**Theorem 30.5 (Alaoglu)** *Let $E$ be a Banach space. Let $B^*$ be the closed unit ball in the dual space $E^*$. Then $B^*$ is compact with respect to the weak$^*$ topology.*

Proof: This theorem applies to either a real or a complex Banach space. Define for each $u$ in $E$ the set $I_u$ of all scalars $a$ such that $|a| \leq \|u\|$. This is an closed interval in the real case or a closed disk in the complex case. In either case each $I_u$ is a compact space. Let $P = \prod_{u\in E} I_u$. By the Tychonoff product theorem, this product space is compact. An element $f$ of $P$ is a scalar function on $E$ with the property that $|f(u)| \leq \|u\|$ for all $u$ in $D$. The product space topology on $P$ is just the topology of pointwise convergence for such functions.

The unit ball $B^*$ in the dual space $E^*$ consists of all elements of $P$ that are linear. The topology on $B^*$ inherited from $P$ is the topology of pointwise convergence. The topology on $B$ inherited from the weak$^*$ topology on $E^*$ is also the topology of pointwise convergence. So the task is to show that $B^*$ is compact in this topology. For this, it suffices to show that $B^*$ is a closed subset of $P$.

For each $u \in E$ the mapping $f \mapsto f(u)$ is continuous on $P$. Therefore, for each pair of scalars $a, b$ and vectors $u, v$ the mapping $f \mapsto f(au+bv) - af(u) - bf(v)$ is continuous on $P$. It follows that the set of all $f$ with $f(au + bv) - af(u) - bf(v) = 0$ is a closed subset of $P$. The intersection of these closed sets for all $a, b$ and all $u, v$ is also a closed subset of $P$. However this intersection is just $B^*$. Since $B^*$ is a closed subset of a compact space $P$, it must be compact. □

If $E$ is an infinite dimensional Banach space, then its dual space $E^*$ with the weak* topology is not metrizable. This fact should be contrasted with the following important result.

**Theorem 30.6** *If $E$ is a separable Banach space, then the unit ball $B^*$ in the dual space with the weak* topology is metrizable.*

Proof: Suppose $E$ is separable. Let $S$ be a countable dense subset of the unit ball $B$ of $E$. Let $I$ be the closed unit ball in the field of scalars. For each $f$ in $B^*$ there is a corresponding element $u \mapsto f(u)$ in $I^S$. Denote this element by $j(f)$. Thus $j(f)$ is just the restriction of $f$ to $S$. From the fact that $S$ is dense in $B$ it is easy to see $j : B^* \to I^S$ is injective. Give $I^S$ the product topology. Since for each $u$ in $S$ the map $f \mapsto f(u)$ is continuous, it follows that $j$ is continuous.

The remaining task is to prove that the inverse $j(f) \mapsto f$ is continuous. To do this, consider a closed subset $F$ of $B^*$. Since it is a closed subset of a compact space, it is compact. Since $j$ is continuous, $j[F]$ is a compact subset of $I^S$. However a compact subset of Hausdorff space is closed. So $j[F]$ is closed. This says that the inverse image of each closed set under the inverse of $j$ is a closed set. It follows that the inverse of $j$ is continuous.

This proves that $j$ is an embedding of $B^*$ into $I^S$. However since $I^S$ is a countable product of metric spaces, the product topology on this space is given by a metric. Such a metric induces a metric on $B^*$. This can be taken to have the explicit form $d(f, f') = \sum_{n=1}^{\infty} |f(s_n) - f'(s_n)|/2^n$. □

Examples:

1. Let $E = L^1$, so $E^* = L^\infty$. The unit ball consists of all functions with absolute value essentially bounded by one. It is possible that a sequence of positive functions with essential bound one converges weak* to zero. For example, on the line the sequence of functions $f_n$ that are the indicator functions of intervals $[n, n+1]$ converge to zero. This is because for each fixed $u$ in $L^1$ we have $\mu(f_n u) \to 0$, by the dominated convergence theorem. Such an example is even possible when the measure space is finite. Here the example would be given by the indicator functions of the sets $[0, 1/n]$. Yet another example is convergence to zero by oscillation. Consider the functions $\cos(nx)$ on the interval $[0, 2\pi]$. These converge weakly to zero in the weak* topology of $L^\infty$, by the Riemann-Lebesgue lemma.

2. Let $E = L^p$ with $1 < p < \infty$, so $E^* = L^q$ with $1 < q < \infty$. Here $1/p + 1/q = 1$. It is possible that a sequence of positive functions with $L^q$ norm equal to one converges weak* to zero. The sequence of indicator functions of the sets $[n, n+1]$ provide the most obvious example. In the case when the measure space is finite, an example is where $f_n$ is the $n^{\frac{1}{q}}$ times the indicator function of the set $[0, 1/n]$. This example is less obvious. It is clear that for $u$ bounded we have $|\mu(f_n u)| \leq n^{\frac{1}{q}} \|u\|_\infty/n \to 0$. Since bounded functions are dense in $L^p$ and we have a bound on the $L^q$ norm of the $f_n$, it follows that we have $\mu(f_n u) \to 0$ for each $u$ in $L^p$. There are yet more examples, such as convergence by oscillation.

3. Consider the space $L^1$ with the weak topology. The closed unit ball is not compact. In fact, let $g_n$ be $n$ times the indicator function of $1/n$. If, for instance, $w$ is a bounded continuous function, then $\mu(wg_n) \to w(0)$. This indicates that $g_n$ is converging to something that acts like a point measure at the origin. This is no longer in the space $L^1$. A sequence of densities with bounded total mass can converge to something that is not a density. In physical terms: conservation of mass is not enough to make something to remain a function. (This should be contrasted with the $L^p$ with $p > 1$ case above. For $L^2$ this says that conservation of energy is enough to maintain the constraint of being a function.)

4. Consider a compact metric space $X$. Let $E = C(X)$. Then the signed measures in the unit ball of the dual space form a weak* compact set. Such a measure $\mu$ is an ordinary positive measure provided that for each positive continuous function $u \geq 0$ the value $\mu(u) \geq 0$. From this it is clear that the positive measures of total mass at most one form a weak* closed subset. (This is because $\mu \mapsto \mu(f)$ is continuous, so the inverse image of the closed set $[0, +\infty)$ is closed.) Therefore they are a compact subsets. Furthermore, the probability measures form a closed subset of these, since the requirement for a positive measure to be a probability measure is that $\mu(1) = 1$. (This is because $\mu \mapsto \mu(1)$ is continuous, so the image of the closed set $\{1\}$ is closed.) The conclusion is that the space of probability measures on a compact metric space is weak* compact. There is no way to lose probability from a compact space! Notice that this example explains what is going on in the preceding example. Consider the sequence $\mu_n$ of probability measures that have density with respect to Lebesgue measure that is $n$ times the indicator function of $[0, 1/n]$. This sequence converges to the point measure $\delta_0$ at the origin, which is still a probability measure.

5. Consider a separable locally compact metric space $X$. Let $E = C_0(X)$. Again the signed measures in the unit ball of $E^*$ form a compact set. The positive measures of total mass at most one again form a compact subset. However the function 1 does not belong to the space $C_0(X)$. So we cannot conclude that the set of probability measures is closed or compact. In fact, we can take the measures with density given by the indicator function of $[n, n+1]$. These probability measures converge weak* to zero.

This seems mysterious until we choose to look instead at the one-point compactification of $X$. Then it is seen that the probability has all gone to the point at infinity.

## Problems

1. Recall that $f : X \to (-\infty, +\infty]$ is lower semicontinuous (LSC) if and only if the inverse image of each interval $(a, +\infty]$, where $-\infty < a < +\infty$, is open in $X$. Show that if $f : X \to \mathbb{R}$ is LSC, and $X$ is compact, then there is a point in $X$ at which $f$ has a minimum value. Show by example that if $f : X \to \mathbb{R}$ is LSC, and $X$ is compact, then there need not be a point in $X$ at which $f$ has a maximum value.

2. Let $H$ be a real Hilbert space. Let $L : H \to \mathbb{R}$ be continuous and linear. Thus in particular $L$ is Lipschitz, that, is, there is an $M$ with $|L(u)| \leq M\|u\|$ for all $u$ in $H$. Consider the problem of proving that there is a point in $H$ at which the function $F : H \to \mathbb{R}$ defined by $F(u) = (1/2)\|u\|^2 - L(u)$ has a minimum value. This can be done using complete metric space ideas, but can it be done using compact topological space ideas? One approach would be to look at a sufficiently large closed ball centered at the origin and argue that if there were a minimum, it would be in that closed ball. If $H$ were finite dimensional, that ball would be compact, and the result is obvious. Show how to carry out the compactness proof for infinite dimensional $H$. Hint: Switch to the weak topology. Be explicit about which functions are continuous or lower semicontinuous.

# Chapter 31

# Radon measures

## 31.1 Topology and measure

The interaction of topology and measure is complicated. A topological structure on a space $X$ may somehow determine a measurable structure on $X$. The simplest example of this is that the topology itself generates the Borel $\sigma$-algebra. It turns out, however, that various technicalities arise. In particular, there are two directions that may be taken.

The first possible direction is to take $X$ to be a locally compact Hausdorff space. This seems quite general, since such a space need not be a metrizable space. (A typical example where this is so is when $X$ is an uncountable product of compact Hausdorff spaces, so that $X$ is itself a compact Hausdorff space, not metrizable.) However in this generality there are technicalities due to the fact that a topology involves uncountable operations, and these interact uneasily with measure theory, which is primarily based on countable operations. In particular, while it is true that for a compact space (or for a $\sigma$-compact locally compact space) that $C_c(X)$ and $BC(X)$ generate the same $\sigma$-algebra, it is quite possible even for compact spaces that this is much smaller than the Borel $\sigma$-algebra. This is something of a nightmare, and so this first direction is not emphasized in the present treatment.

The other direction is to take $X$ to be a Polish space, that is, a separable completely metrizable space. This is general in a different way, since such a space need not be locally compact. (A typical example is when $X$ is an infinite-dimensional Banach space; such a space is never locally compact.) In this setting compactness issues can be something of a struggle. However certainly $C(X)$ generates the Borel $\sigma$-algebra, so it is easy to fix on the concept of measurability. We just talk of Borel subsets.

The best possible world is when the space $X$ is a second countable locally compact Hausdorff space. This is the same as being a locally compact Polish space. This is general enough for many applications, and most of the technicalities are gone. In particular, $C_c(X)$ generates the Borel $\sigma$-algebra.

## 31.2 Locally compact metrizable spaces

Next we need to explore the consequences of having a metrizable space that is locally compact. The following lemma is crucial.

**Lemma 31.1 (Urysohn's lemma (locally compact metric case))** *Let $X$ be a locally compact metric space. Let $K \subset X$ be a compact subset. Then there is a continuous function with values in $[0, 1]$ with compact support that has the value 1 on $K$.*

Proof: Since $X$ is locally compact, it is not hard to prove that there is an open set $U$ and a compact set $L$ with $K \subset U \subset L$. Take $f$ to be 1 on $K$ and zero on the complement of $U$. Then $f$ has support in $L$. □

Notice that this result fails without the hypothesis of local compactness. For example, consider a point in infinite dimensional Hilbert space. Say that there is a real continuous function on the Hilbert space that is one at the point. Then it is non-zero on some non-empty open set. However a non-empty open set is never a subset of a compact set.

A topological space is said to be *$\sigma$-compact* if it is a countable union of compact subsets.

**Proposition 31.2** *A second countable locally compact Hausdorff space is $\sigma$-compact.*

Proof: Since the space is locally compact, each point belongs to an open subset that is included in a compact subset. The collection of these open subsets is a cover of the space. By Lindelöf's theorem there is a countable subcover. The compact sets that correspond to this subcover are a countable collection whose union is the entire space. □

A $\sigma$-compact space need not be locally compact. In fact the space $\mathbb{Q} \subset \mathbb{R}$ is $\sigma$-compact. Also, a $\sigma$-compact space need not be second countable; in fact even a compact space need not be second countable. On the other hand, a $\sigma$-compact metrizable space is second countable. This is because each compact metrizable space is separable, and a countable union of separable metric spaces is a separable metric space, hence second countable.

**Theorem 31.3** *Let $X$ be a locally compact metric space that is $\sigma$-compact. Then $C_c(X)$ generates the Borel $\sigma$-algebra.*

Proof: Since $X$ is $\sigma$-compact, it may be written as an increasing union of compact subsets $K_n$. Let $f_n$ be in $C_c(X)$ with $0 \leq f_n \leq 1$ and $f_n = 1$ on $K_n$. Then the $f_n$ converge pointwise to 1. If $g$ is in $C(X)$, we can take a sequence $gf_n$ in $C_c(X)$ that converges pointwise to $g$. So $C_c(X)$ generates the same $\sigma$-algebra as $C(X)$. However for a metric space this is the Borel $\sigma$-algebra. □

## 31.3 Riesz representation

The following lemma will be useful in the following.

**Lemma 31.4** *Let $X$ be a locally compact metric space. Suppose in addition that $X$ is $\sigma$-compact. Then the $\sigma$-ring generated by $C_c(X)$ is a $\sigma$-algebra.*

Proof: Since $X$ is $\sigma$-compact, it may be written as an increasing union of compact subsets $K_n$. Let $f_n$ be in $C_c(X)$ with $0 \leq f_n \leq 1$ and $f_n = 1$ on $K_n$. Then the $f_n$ converge pointwise to 1. Therefore 1 is the $\sigma$-ring of functions generated by $C_c(X)$, and so this $\sigma$-ring is a $\sigma$-algebra. □

**Theorem 31.5 (Dini)** *Suppose that $X$ is a compact space. Let $f_n \downarrow 0$ be a decreasing sequence of continuous functions that converges pointwise to zero. Then $f_n$ converges uniformly to zero.*

Proof: Consider $\epsilon > 0$. The set where $f_n \geq \epsilon$ is a closed subset of a compact space and hence is compact. The pointwise convergence implies that the intersection of the collection of all these sets is zero. Hence there is a finite subcollection whose intersection is zero. However this is a decreasing sequence of sets. Therefore these sets are empty from some index on. That is, from some index on the set $f_n < \epsilon$ is the whole space. This implies uniform convergence. □

A *Radon measure* on a space of real continuous functions is a linear function $\mu$ from the space to the reals that is order preserving. Thus in particular $f \geq 0$ implies $\mu(f) \geq 0$. A Radon measure might better be called a Radon integral, since it acts on functions rather than on sets, but both terms are used. The following theorem is sometimes called a Riesz representation theorem, even though the space $C_c(X)$ is not a Banach space, and the theorem as stated here is only for linear functionals that preserve order.

**Theorem 31.6 (Riesz representation (compact support case))** *Let $X$ be a locally compact Polish space. Then there is a natural bijective correspondence between Radon measures on $C_c(X)$ and Borel measures that are finite on compact sets.*

Remark: A locally compact Polish space is the same as a second countable locally compact Hausdorff space.

Proof: Suppose each $f_n$ is in $C_c(X)$ and that $f_n \downarrow 0$ pointwise. There is a fixed compact set $K$ such that each $f_n$ has support in $K$. According to Dini's theorem, $f_n \to 0$ uniformly. Let $g$ be in $C_c(X)$ and have the value 1 on $K$. Then $0 \leq f_n \leq \|f_n\|_{\sup} g$, so $0 \leq \mu(f_n) \leq \|f_n\|_{\sup}\mu(g)$. Thus $\mu(f_n) \to 0$. That is, $\mu$ satisfies the monotone convergence theorem on $C_c(X)$. Thus $\mu$ is an elementary integral.

The elementary integral extends uniquely to the $\sigma$-ring generated by $C_c(X)$. Since $X$ is $\sigma$-compact, this is the same as the $\sigma$-algebra generated by $C_c(X)$. In the present context this is the Borel $\sigma$-algebra. □

It is worth observing that the natural class of maps for such Radon measures is the class of $\phi : X \to Y$ such that $\phi$ is measurable and such that $L \subset Y$ compact implies there is a $K \subset X$ compact with $\phi^{-1}[L] \subset K$. Thus if $\mu$ is finite on compact subsets of $X$, then for each compact subset $L$ of $Y$ the image measure $\phi[\mu]$ satisfies the property that $\phi[\mu](L) = \mu(\phi^{-1}[L]) \leq \mu(K)$ is finite. When $\phi$ is continuous this is often called a *proper map*.

Now we look at the situation for the larger space $C_0(X)$ of continuous functions that vanish at infinity.

**Theorem 31.7** *If $f$ is a locally compact metric space, then the closure of the space $C_c(X)$ of continuous functions with compact support in the space $BC(X)$ of bounded continuous functions is $C_0(X)$, the space of continuous functions that vanish at infinity.*

Proof: Since a function with compact support vanishes at infinity, it follows that the closure of $C_c(X)$ is a subset of $C_0(X)$. The converse is slightly more complicated. Suppose that $f$ is in $C_0(X)$. Then there is a compact set $K_n$ such that $|f| < 1/n$ on the complement of $K_n$ Let $g_n$ be in $C_c(X)$ with $0 \leq g_n \leq 1$ and with $g_n = 1$ on $K_n$. Then $g_n f$ is in $C_c(X)$. Furthermore, $(1 - g_n)|f|$ is bounded by $1/n$. Hence $g_n f \to f$ uniformly. So $f$ is in the closure of $C_c(X)$. □

**Theorem 31.8** *Let $X$ be a locally compact metric space. Let $\mu$ be as Radon measure on $C_0(X)$. Then $\mu$ is a Lipschitz function on $C_0(X)$ with the uniform norm.*

Proof: First we argue that there is a $M$ so that for $f$ in $C_0(X)$ with $0 \leq f \leq 1$ we have $\mu(f) \leq M$. Otherwise, there is sequence of $f_k$ with $0 \leq f_k \leq 1$ and $\mu(f_k) \geq 2^k$. Let $g = \sum_{k=1}^{\infty} \frac{1}{2^k} f_k$. Then $g$ is also in $C_0(X)$. Furthermore, for each $n$ we have $g \geq g_n = \sum_{k=1}^{n} \frac{1}{2^k} f_k$. It follows that $\mu(g) \geq \mu(g_n) \geq n$. If $n > \mu(g)$ this is a contradiction.

Since $\mu$ is linear, if $-1 \leq f \leq 1$, then we have $\mu(f) = \mu(f_+) - \mu(f_-)$ and so $|\mu(f)| \leq \mu(f_+) + \mu(f_-) \leq 2M$. It then follows for arbitrary $f$ in $C_0(X)$ that $|\mu(f)| \leq 2M\|f\|$. □

**Corollary 31.9 (Riesz representation theorem (locally compact case))** *Let $X$ be locally compact Polish space. Then there is a natural bijective correspondence between Radon measures on $C_0(X)$ and finite Borel measures.*

Remark: A locally compact Polish space is the same as a second countable locally compact Hausdorff space.

Proof: A Radon measure $\mu$ on $C_0(X)$ restricts to a measure $\mu'$ on $C_c(X)$. By the previous theorem $\mu'$ uniquely corresponds to a Borel measure finite on compact subsets. If this Borel measure $\mu'$ is not finite, then one can construct $f_n$ in $C_c(X)$ with $0 \leq f_n \leq 1$ and $f_n \uparrow 1$, and so $\mu(f_n) \uparrow +\infty$. This is a contradiction, so $\mu'$ is a finite measure.

Since $\mu = \mu'$ are both Lipschitz on $C_0(X)$ and agree on the dense subset $C_c(X)$, we see that $\mu = \mu'$ on $C_0(X)$. □

There is no corresponding theorem for $BC(X)$. One can always start with a linear order-preserving functional $\mu$ on $BC(X)$. Again it will be Lipschitz. One can restrict it to a Radon measure $\mu'$ on $C_c(X)$ that gives rise to a finite Borel measure. However it is no longer the case that we can conclude that $\mu'$ agrees with the original $\mu$ on $BC(X)$. The correspondence from $\mu$ to $\mu'$ can be many to one. The problem is that $C_c(X)$ is not dense in $BC(X)$. The functional $\mu$ can depend on asymptotic properties of the functions on $BC(X)$ that are not captured by a measure on $X$.

**Corollary 31.10 (Riesz representation theorem (compact case))** *Let $X$ be a a compact metrizable space. Then there is a natural bijective correspondence between Radon measures on $C(X)$ and finite Borel measures.*

Remark: A compact metrizable space is the same as a second countable compact Hausdorff space.

Again there is a result where the Radon measure $\mu$ is not required to be order preserving but only to be continuous on $C(X)$ with compact $X$. The conclusion is that $\mu$ is given by a finite signed Borel measure.

## 31.4 Lower semicontinuous functions

Let us look more closely at the extension process in the case of a Radon measure. We begin with the positive linear functional on the space $L = C_c(X)$ of continuous functions with compact support. The construction of the integral associated with the Radon measure proceeds in the standard two stage process. The first stage is to consider the integral on the spaces $L\uparrow$ and $L\downarrow$. The second stage is to use this extended integral to define the integral of an arbitrary summable function.

A function $f$ from $X$ to $(-\infty, +\infty]$ is said to be *lower semicontinuous* (LSC) if for each real $a$ the set $\{x \mid f(x) > a\}$ is an open set. A function $f$ from $X$ to $[-\infty, +\infty)$ is said to be *upper semicontinuous* (USC) if for each real $a$ the set $\{x \mid f(x) < a\}$ is an open set. Clearly a continuous real function is both LSC and USC.

**Theorem 31.11** *If each $f_n$ is LSC and if $f_n \uparrow f$, then $f$ is LSC. If each $f_n$ is USC and if $f_n \downarrow f$, then $f$ is USC.*

It follows from this theorem that space $L\uparrow$ consists of functions that are LSC. Similarly, the space $L\downarrow$ consists of functions that are USC. These functions can already be very complicated. The first stage of the construction of the integral is to use the monotone convergence theorem to define the integral on the spaces $L\uparrow$ and $L\downarrow$.

In order to define the integral for a measurable functions, we approximate such a function from above by a function in $L\uparrow$ and from below by a function in $L\downarrow$. This is the second stage of the construction. The details were presented in an earlier chapter.

The following is a useful result that we state without proof.

**Theorem 31.12** *If $\mu$ is a Radon measure and if $1 \le p < \infty$, then $C_c(X)$ is dense in $L^p(X, \mathcal{B}, \mu)$.*

The corresponding result for $p = \infty$ is false. The uniform closure of $C_c(X)$ is $C_0(X)$, which in general is much smaller than $L^\infty(X, \mathcal{B}, \mu)$. A bounded function does not have to be continuous, nor does it have to vanish at infinity.

## 31.5 Weak* convergence

In order to emphasize the duality between the space of measures and the space of continuous functions, we sometimes write the value of the Radon measure $\mu$ on the continuous function $f$ as

$$\mu(f) = \langle \mu, f \rangle. \tag{31.1}$$

As before, we consider only positive Radon measures, though there is a generalization to signed Radon measures. We consider finite Radon measures, that is, Radon measures for which $\langle \mu, 1 \rangle < \infty$. Such a measure extends by continuity to $C_0(X)$, the space of real continuous functions that vanish at infinity. In the case when $\langle \mu, 1 \rangle = 1$ we are in the realm of probability. Throughout we take $X$ to be a separable locally compact metric space, though a more general setting is possible.

In this section we describe *weak* convergenceweak* convergence* for finite Radon measures. In probability this is often called *vague convergence*. A sequence $\mu_n$ of finite Radon measures is said to weak* converge to a finite Radon measure $\mu$ if for each $f$ in $C_0(X)$ the numbers $\langle \mu_n, f \rangle \to \langle \mu, f \rangle$.

The importance of weak* convergence is that it gives a sense in which two probability measures with very different qualitative properties can be close. For instance, consider the measure

$$\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{\frac{k}{n}}. \tag{31.2}$$

This is a Riemann sum measure. Also, consider the measure

$$\langle \lambda, f \rangle = \int_0^1 f(x)\, dx. \tag{31.3}$$

This is Lebesgue measure on the unit interval. Then $\mu_n \to \lambda$ in the weak* sense, even though each $\mu_n$ is discrete and $\lambda$ is continuous.

A weak* convergent sequence can lose mass. For instance, a sequence of probability measures $\mu_n$ can converge in the weak* sense to zero. A simple example is the sequence $\delta_n$. The following theory shows that a weak* convergent sequence cannot gain mass.

**Theorem 31.13** *If $\mu_n \to \mu$ in the weak* sense, then $\langle \mu, f\rangle \leq \liminf_{n\to\infty}\langle \mu_n, f\rangle$ for all $f \geq 0$ in $BC(X)$.*

Proof: It is sufficient to show this for $f$ in $BC$ with $0 \leq f \leq 1$. Choose $\epsilon > 0$. Let $0 \leq g \leq 1$ be in $C_0$ so that $\langle \mu, (1-g)\rangle < \epsilon$. Notice that $gf$ is in $C_0$. Furthermore, $(1-g)f \leq (1-g)$ and $gf \leq f$. It follows that

$$\langle \mu, f\rangle \leq \langle \mu, gf\rangle + \langle \mu, (1-g)\rangle \leq \langle \mu, gf\rangle + \epsilon \leq \langle \mu_k, gf\rangle + 2\epsilon \leq \langle \mu_k, f\rangle + 2\epsilon \quad (31.4)$$

for $k$ sufficiently large. □

The following theorem shows that if a weak* convergent sequence does not lose mass, then the convergence extends to all bounded continuous functions.

**Theorem 31.14** *If $\mu_n \to \mu$ in the weak* sense, and if $\langle \mu_n, 1\rangle \to \langle \mu, 1\rangle$, then $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for all $f$ in $BC(X)$.*

Proof: It is sufficient to prove the result for $f$ in $BC$ with $0 \leq f \leq 1$. The preceding result gives an inequality in one direction, so it is sufficient to prove the inequality in the other direction. Choose $\epsilon > 0$. Let $0 \leq g \leq 1$ be in $C_0$ so that $\langle \mu, (1-g)\rangle < \epsilon$. Notice that $gf$ is in $C_0$. Furthermore, $(1-g)f \leq (1-g)$ and $gf \leq f$. For this direction we note that the extra assumption implies that $\langle \mu_n, (1-g)\rangle \to \langle \mu, (1-g)\rangle$. We obtain

$$\langle \mu_n, f\rangle \leq \langle \mu_n, gf\rangle + \langle \mu_n, (1-g)\rangle \leq \langle \mu, gf\rangle + \langle \mu, (1-g)\rangle + 2\epsilon \leq \langle \mu, gf\rangle + 3\epsilon \leq \langle \mu, f\rangle + 3\epsilon \quad (31.5)$$

for $n$ sufficiently large. □

It is not true in general that the convergence works for discontinuous functions. Take the function $f(x) = 1$ for $x \leq 0$ and $f(x) = 0$ for $x > 0$. Then the measures $\delta_{\frac{1}{n}} \to \delta_0$ in the weak* sense. However $\langle \delta_{\frac{1}{n}}, f\rangle = 0$ for each $n$, while $\langle \delta_0, f\rangle = 1$.

We now want to argue that the convergence takes place also for certain discontinuous functions. A quick way to such a result is through the following concept. For present purposes, we say that a bounded measurable function $g$ has $\mu$-negligible discontinuities if for every $\epsilon > 0$ there are bounded continuous functions $f$ and $h$ with $f \leq g \leq h$ and such that $\mu(f)$ and $\mu(h)$ differ by less than $\epsilon$.

Example: If $\lambda$ is Lebesgue measure on the line, then every bounded piecewise continuous function with jump discontinuities has $\lambda$-negligible discontinuities.

Example: If $\delta_0$ is the Dirac mass at zero, then the indicator function of the interval $(-\infty, 0]$ does not have $\delta_0$-negligible discontinuities.

**Theorem 31.15** *If $\mu_n \to \mu$ in the weak* sense, and if $\langle \mu_n, 1\rangle \to \langle \mu, 1\rangle$, then $\langle \mu_n, g\rangle \to \langle \mu, g\rangle$ for all bounded measurable $g$ with $\mu$-negligible discontinuities.*

Proof: Take $\epsilon > 0$. Take $f$ and $h$ in $BC$ such that $f \leq g \leq h$ and $\mu(f)$ and $\mu(h)$ differ by at most $\epsilon$. Then $\mu_n(f) \leq \mu_n(g) \leq \mu_n(h)$ for each $n$. It follows that $\mu(f) \leq \liminf_{n\to\infty} \mu_n(g) \leq \limsup_{n\to\infty} \mu_n(g) \leq \mu(h)$. But also $\mu(f) \leq \mu(g) \leq \mu(h)$. This says that $\liminf_{n\to\infty} \mu_n(g)$ and $\limsup_{n\to\infty} \mu_n(g)$ are each within $\epsilon$ of $\mu(g)$. Since $\epsilon > 0$ is arbitrary, this proves that $\lim_n \mu_n(g)$ is $\mu(g)$. □

## 31.6 Central limit theorem for coin tossing

This section gives a statement of the for the special case of coin-tossing. The purpose is to illustrate the of the weak* convergence concept.

For each $n$ consider the space $\{0,1\}^n$ for the outcomes of $n$ coin tosses. We shall think of 0 as heads and 1 as tails. The probability measure for fair tosses of a coin is the uniform measure that assigns measure $1/2^n$ to each of the $2^n$ one point sets.

For $j = 1, \ldots, n$ let $x_j$ be the function on $\{0,1\}^n$ that has value $x_j(\omega) = 1$ if $\omega_j = 0$ and $\omega_j = -1$ if $\omega_j = 1$. Thus $x_1 + \cdots + x_n$ is the number of heads minus the number of tails. It is a number between $-n$ and $n$.

Let $\mu_n$ be the probability measure on the line that is the image of the uniform probability measure on $\{0,1\}^n$ under the map

$$\frac{1}{\sqrt{n}}(x_1 + \cdots + x_n) : \{0,1\}^n \to \mathbb{R}. \tag{31.6}$$

This measure assigns mass

$$p_k = \binom{n}{k}\frac{1}{2^n} \tag{31.7}$$

to the point $(2k-n)/\sqrt{n}$, for $k = 0, 1, \ldots, n$. So it is very much a discrete measure.

**Theorem 31.16 (Central limit theorem for coin tossing)** *The measures $\mu_n$ converge in the weak* sense to the standard normal probability measure $\mu$ with density*

$$\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2} \tag{31.8}$$

*with respect to Lebesgue measure on the line.*

## 31.7 Weak* probability convergence and Wiener measure

Consider a Polish space $X$ that is not necessarily locally compact. There is still a useful notion of convergence, provided that we restrict the discussion to probability measures. A sequence $\mu_n$ of Borel probability measures is said to converge to a Borel probability measure $\mu$ in the *weak* probability sense* if for every $f$ in $BC(X)$ we have $\mu_n(f) \to \mu(f)$ as $n \to \infty$.

The nice feature of this definition is that it applies not only to locally compact spaces, but also to more general spaces, such as infinite dimensional separable Banach spaces. See the book by Dudley [4] for a thorough development of this theory. Notice that if $X$ does happens to be locally compact, then on the class of probability measures this agrees with the previous definition of weak* convergence of Radon measures.

Perhaps the most famous probability measure on an infinite dimensional space is *Wiener measure* (otherwise known as the Einstein model for *Brownian motion*). Fix a time interval $[0, T]$. The space on which the measure lives is the space $X = C([0, T])$ of all real functions on $[0, T]$. These are considered as function from time to one dimensional space. According to Einstein there is a parameter $\sigma > 0$ that related time to space. This is the *diffusion constant.* It is defined so that the expectation of the square of the distance travelled is $\sigma$ times the elapsed time.

Consider a natural number $N$ and define $\Delta t = T/N$. For each $j = 0, 1, 2, 3 \ldots, N$ there are corresponding time instants $0, \Delta t, 2\Delta t, 3\Delta t, T$. For a coin toss sequence $\omega$ in $\{0, 1\}^N$ and a time $t$ in $[0, T]$ with $j\Delta t \leq t \leq (j + 1)\Delta t$ define

$$W(t, \omega) = \sigma\sqrt{\Delta t}(x_1(\omega) + \ldots + x_j(\omega) + \frac{t - j\Delta t}{\Delta t} x_{j+1}(\omega)). \tag{31.9}$$

For each coin toss sequence $\omega$ the function $t \mapsto W(t, \omega)$ is a piecewise linear continuous path in the space $X = C([0, T])$. Define $\mu_N$ as the image of the coin tossing measure on $\{0, 1\}^N$ in the space $X = C([0, T])$.

**Theorem 31.17 (Existence of Wiener measure)** *Fix the total time $T$ and the diffusion constant $\sigma$. For each $N$ there is a probability measure $\mu_N$ in $X = C([0, T])$ defined as above by $N$ tosses of a fair coin. The assertion is that there is a probability measure $\mu$ in $X = C([0, T])$ such that $\mu_n \to \mu$ in the weak* probability sense as $n \to \infty$.*

The theorem is proved in probability texts, such as the one by Fristedt and Gray [6]. The Wiener measure has remarkable properties.

1. For each $t$ in $[0, T]$ the expectation $\mu(W(t)) = 0$.
2. For each $t$ in $[0, T]$ and $h \geq 0$ with $t+h$ in $[0, T]$ we have $\mu(W(t)W(t+h)) = \sigma^2 t$. In particular the variance is $\mu(W(t))^2 = \sigma^2 t$.
3. For each $t$ the random variable $W(t)$ has the normal distribution with density

$$\phi_t(x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{x^2}{2\sigma^2 t}}. \tag{31.10}$$

The second property implies that the next increment $W(t + h) - W(t)$ is uncorrelated with the present $W(t)$, that is, $\mu((W(t + h) - W(t))W(t)) = 0$.

As a consequence of these properties we see that for $h \geq 0$ we have

$$\mu((W(t + h) - W(t))^2) = \sigma^2 h \tag{31.11}$$

This is consistent with the fact that the paths are continuous. However for $h > 0$ this also says that

$$\mu(\left(\frac{W(t + h) - W(t)}{h}\right)^2) = \frac{\sigma^2}{h}. \tag{31.12}$$

This suggests that though the typical Wiener path is continuous, it is not differentiable.

## 31.8 Supplement: Measure theory on locally compact spaces

In this section we explore properties of locally compact Hausdorff spaces that are not metrizable. Much of the elementary theory about spaces of continuous functions continues to work [5], but there are new measure-theoretic technicalities.

If $X$ is a topological space, then it determines a measurable space by taking the *Borel $\sigma$-algebra of subsets*. This is the smallest $\sigma$-algebra $\mathcal{B}o$ that contains all the open sets of the topological space. Since it is closed under complements, it also contains all the closed sets.

The smallest $\sigma$-algebra of subsets for which every function in $C(X)$ is measurable will be called the *continuous Baire* $\sigma$-algebra and denoted $\mathcal{B}c$. It is clear that $\mathcal{B}c \subset \mathcal{B}o$. It is not difficult to see that this $\sigma$-algebra may also be generated by $BC(X)$.

For topological spaces that are not metrizable these may be different. An example is obtained with a compact Hausdorff space $Y$ (such as the two point set $\{0, 1\}$ or the unit interval $[0, 1]$) and an uncountable index set $I$. Then $X = Y^I$ is a compact Hausdorff space, but it is not metrizable. Each set or function in the continuous Baire $\sigma$-algebra depends only on countable many coordinate in $I$. A one point set is defined by restricting all of the coordinates in $I$ to have fixed values. So the one point sets are not in the continuous Baire $\sigma$-algebra. On the other hand, each one point set in $X$ is closed, and so it is in the Borel $\sigma$-algebra

Let $X$ be a locally compact Hausdorff space. The *Baire $\sigma$-algebra $\mathcal{B}a$* is the $\sigma$-algebra of functions generated by the space $C_c(X)$ of real continuous functions on $X$, each of which has compact support. This is the same $\sigma$-algebra as that generated by the space $C_0(X)$ of real continuous functions on $X$ that vanish at infinity.

An example where the Baire $\sigma$-algebra $\mathcal{B}a$ is strictly smaller than the extended Baire $\sigma$-algebra $\mathcal{B}c$ is when $X$ is an uncountable discrete space.

The general relation of the Baire, continuous Baire, and Borel $\sigma$-algebras is $\mathcal{B}a \subset \mathcal{B}c \subset \mathcal{B}o$. See Royden [17] for yet more $\sigma$-algebras. The following result is from this source.

**Theorem 31.18** *Let $X$ be a $\sigma$-compact locally compact Hausdorff space. Then $\mathcal{B}a = \mathcal{B}c$.*

It is interesting to consider the Riesz representation theorems for non-metric spaces. In this case the simplest results are for the Baire $\sigma$-algebra $\mathcal{B}a$. Recall that a Radon measure on a space of real continuous functions is a linear function $\mu$ from the space to the reals that is order preserving. Thus in particular $f \geq 0$ implies $\mu(f) \geq 0$.

**Theorem 31.19 (Riesz representation)** *Let $X$ be a locally compact Hausdorff space. Suppose that it is also $\sigma$-compact. Then there is a natural bijective* *correspondence between Radon measures on $C_c(X)$ and Baire measures that are finite on compact sets.*

Proof: Suppose as known the usual properties of continuous functions on locally compact Hausdorff spaces, as described in Folland[5]. Here is a skecth of the proof. Consider such a Radon measure $\mu$. The local compactness and Dini's theorem are be used in the usual way to show that $\mu$ is an elementary integral on $C_c(X)$.

The elementary integral extends uniquely to the $\sigma$-ring generated by $C_c(X)$. Since $X$ is $\sigma$-compact, this is the same as the $\sigma$-algebra generated by $C_c(X)$ [17]. This is the Baire $\sigma$-algebra $\mathcal{B}a$. □

Example: Let $X = [0,1]^{\mathbb{R}}$, an uncountable product of copies of the unit interval. This is a compact Hausdorff space. The continuous functions that depend on only finitely many coordinates are dense in the space $C(X)$. It follows that each Baire subset depends on at most countably many coordinates.

There are many interesting subsets, even open subsets, that are not Baire subsets. For example, consider a non-trivial open subset $U$ of $[0,1]$. For each real $t$, the set $S_t$ of all $\omega$ in $X$ such $\omega(t)$ is in $U$ is an open subset. Let $S$ be the set of all $\omega$ in $S$ such that for some real $t$ the value $\omega(t)$ is in $U$. Then $S$ is the union of the open subsets $S_t$ for $t$ in $\mathbb{R}$, so $S$ is open. But $S$ is not a Borel subset.

The Borel $\sigma$-algebra is large enough to include most interesting subsets, and so one might want to have a correspondence between Radon measures and Borel measures. In order to make such a correspondence well-defined, the Borel measure must be assumed to have regularity properties with respect to certain possibly uncountable supremum and infimum operations [5].

One way to do this is to use the notion of lower semicontinuous (LSC) function. If $X$ is a locally compact Hausdorff space, and $f \geq 0$ is an lower semicontinuous function from $X$ to $[0, +\infty)$, then $f$ is the supremum of the set of all $g$ in $C_c(X)$ with $0 \leq g \leq f$. Notice that this supremum may be over an uncountable set of functions. It thus seems reasonable to require that an integral $\mu$ satisfy the the first stage regularity property: If $h \geq 0$ is LSC, then

$$\mu(h) = \sup\{\mu(f) \mid f \in C_c(X), 0 \leq f \leq h\}. \tag{31.13}$$

**Theorem 31.20 (Riesz representation)** *Let $X$ be a locally compact Hausdorff space. Then there is a natural bijective correspondence between Radon measures on $C_c(X)$ and Borel measures that are finite on compact sets and whose corresponding integrals $\mu$ the following the first stage regularity property and also the second stage regularity property: For each Borel measurable function $g \geq 0$*

$$\mu(g) = \inf\{\mu(h) \mid h \in \mathrm{LSC}, g \leq h\} \tag{31.14}$$

Here is an example where such issues arise [5]. Consider the space $X = C([0,T])$ that was used in the example of Wiener measure. It is a complete separable metric space, and the notion of convergence, uniform convergence, seems quite natural. However it is not locally compact. It would be nice if it were possible to to use a compact space in the construction of Wiener measure.

One approach is to use the product space $[-\infty, \infty]^{[0,T]}$. This has the topology of pointwise convergence. This space is much larger, but it is automatically compact, because of the way that the topology is defined. However it is not a metric space. Furthermore, the continuous functions are not a Baire subset, but only a Borel subset. So one has to deal with measure theory technicalities in order to get a measure on the space of continuous functions. This approach is so elegant, however, that it apparently justifies the study of Borel measures on compact Hausdorff spaces.

This approach, while elegant, may not save so much work. The estimates that are needed to show that the measure lives on the space of continuous functions are rather similar to the estimates that are needed to establish the required compactness properties in the approach that works directly with $X = C([0, T])$.

## Problems

1. Show that the union of a finite collection of compact sets is compact.

2. Recall that $X$ is locally compact if for every point $x$ in $X$ there is an open subset $U$ and a compact subset $K$ with $x \in U \subset K$. Suppose that $X$ is locally compact. Prove that for every compact subset $M \subset X$ there is an open subset $V$ and a compact subset $N$ such that $M \subset V \subset N$.

3. The problem concerns real Borel functions on the line. Let
$$g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}. \tag{31.15}$$
The integral of $g$ with respect to Lebesgue measure is 1. For $t > 0$ let $g_t(x) = (1/t)g(x/t)$. Let $\gamma_t$ be the measure whose Radon-Nikodym derivative with respect to Lebesgue measure is $g_t$. Recall that weak$^*$ convergence of measures says that for each continuous real function that vanishes at infinity the integrals of the function converge.

   (a) Find the weak$^*$ limit of $\gamma_t$ as $t \to +\infty$.

   (b) Find the weak$^*$ limit of $\gamma_t$ as $t \to 0$.

4. We know that
$$\sum_{k=1}^{n} \frac{1}{n} \delta_{\frac{k}{n}} \to \lambda_1 \tag{31.16}$$
as $n \to \infty$ in the weak$^*$ sense, where $\lambda_1$ is Lebesgue measure on the unit interval.

   (a) Evaluate the limit as $n \to \infty$ of
$$\sum_{k=1}^{n} \frac{1}{n} \delta_{\frac{k}{n^2}}. \tag{31.17}$$

(b) Evaluate the limit as $n \to \infty$ of

$$\sum_{k=1}^{n} \frac{1}{n}\delta_k. \tag{31.18}$$

(c) Evaluate the limit as $n \to \infty$ of

$$\sum_{k=1}^{n} \frac{k^2}{n^3}\delta_{\frac{k}{n}}. \tag{31.19}$$

(d) Evaluate the limit as $n \to \infty$ of

$$\sum_{k=1}^{n^2} \frac{n}{n^2+k^2}\delta_{\frac{k}{n}}. \tag{31.20}$$

5. There is an identity

$$\frac{1}{m+1}\sum_{n=0}^{m}\sum_{k=-n}^{n} e^{ikx} = \frac{1}{m+1}\frac{\sin^2(\frac{1}{2}(m+1)x)}{\sin^2(\frac{1}{2}x)}. \tag{31.21}$$

(a) Find the integral of this function over the interval from $-\pi$ to $\pi$.

(b) For each $x$ in the interval from $-\pi$ to $\pi$ find the pointwise limit of this function as $m \to \infty$.

(c) Find the weak* limit as $m \to \infty$ of the measure with this density (with respect to Lebesgue measure) in the space $C(T)^*$, where $T$ is the circle parameterized by $[-\pi, \pi)$. Justify your calculation.

6. The context is Borel functions on the real line. Let $f$ be in $L^2$ and $g$ be in $L^1$. Let $T : L^2 \to L^2$ be defined by the convolution $Tf = g * f$. Then $T$ is a continuous linear transformation. For each polynomial $p$ define $\mu(p) = \langle f, p(T)f\rangle$. Then $\mu$ defines a Radon measure and hence is given by a Borel measure on the line. Find this measure explicitly, as the image of an absolutely continuous measure.

# Bibliography

[1] Howard Becker and Alexander S. Kechris, *The descriptive set theory of Polish group actions*, London Mathematical Society Lecture Note Series, no. 232, Cambridge University Press, Cambridge, 1996.

[2] Czesław Bessaga and Aleksander Pełczyński, *Selected topics in infinite-dimensional topology*, Monografie Matematyczne, no. 58, PWN — Polish Scientific Publishers, Warsaw, 1975.

[3] R. M. Dudley, *Uniform central limit theorems*, Cambridge University Press, Cambridge, 1999.

[4] Richard M. Dudley, *Real analysis and probability*, Cambridge University Press, Cambridge, 2002.

[5] Gerald B. Folland, *Real analysis: Modern techniques and their applications*, second ed., John Wiley & Sons, New York, 1999.

[6] Bert Fristedt and Lawrence Gray, *A modern approach to probability theory*, Probability Theory and its Applications, Birkhäuser, Boston, 1997.

[7] Paul R. Halmos, *Measure theory*, Van Nostrand, Princeton, NJ, 1950.

[8] Greg Hjorth, *Classification and orbit equivalence relations*, Mathematical Surveys and Monographs, no. 75, American Mathematical Society, Providence, RI, 1955.

[9] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, no. 156, Springer-Verlag, New York, 1995.

[10] John L. Kelley, *General topology*, D. Van Nostrand, Princeton, NJ, 1955.

[11] F. W. Lawvere, *Metric spaces, generalized logic, and closed categories*, Rendiconti del Seminario Matematico e Fisico di Milano **43** (1973), 135–166.

[12] Lynn H. Loomis, *An introduction to abstract harmonic analysis*, D. Van Nostrand Company, New York, 1953.

[13] R. Lowen, *Approach spaces: The missing link in the topology-uniformity-metric triad*, Clarendon Press, Oxford, 1997.

[14] Saunders Mac Lane, *Categories for the working mathematician*, Graduate Texts in Mathematics, no. 5, Springer-Verlag, New York, 1971.

[15] George W. Mackey, *Borel structure in groups and their duals*, Transactions of the American Mathematical Society **85** (1957), 134–165.

[16] Michael D. Potter, *Sets: An introduction.*

[17] H. L. Royden, *Real analysis*, third ed., Macmillan, New York, 1988.

[18] Bernd S. W. Schröder, *Ordered sets: An introduction*, Birkhäuser, Boston, 2003.

[19] Ya. G. Sinai, *Introduction to ergodic theory*, Mathematical Notes, Princeton University Press, Princeton, NJ, 1977.

[20] S. M. Srivastava, *A course on Borel sets*, Graduate Texts in Mathematics, no. 180, Springer, New York, 1998.

[21] Daniel W. Stroock, *A concise introduction to the theory of integration*, third ed., Birkhäuser, Boston, 1999.

[22] R. J. Wood, *Ordered sets via adjunction*, Categorical Foundations: Special Topics in Order, Topology, Algebra, and Sheaf Theory (Maria Cristina Pedicchio and Walter Tholen, eds.), Encyclopedia of Mathematics and its Applications, no. 97, Cambridge University Press, New York, 2004, pp. 5–47.

# Mathematical Notation

| | |
|---|---|
| *Logic* | |
| $\forall$ | for all |
| $\wedge$ | and |
| $\exists$ | for some, there exists |
| $\vee$ | or |
| $\Rightarrow$ | implies |
| $\neg$ | not |
| $\perp$ | false statement |
| | |
| *Sets* | |
| $x \in A$ | $x$ is in $A$ |
| $A \subset B$ | $A$ is a subset of $B$ |
| $\emptyset$ | empty set |
| $\{x, y\}$ | unordered pair with $x$ and $y$ in it |
| $(x, y)$ | ordered pair with $x$ first, $y$ second |
| $\bigcap \Gamma$ | intersection of a collection $\Gamma$ of sets |
| $A \cap B$ | $\bigcap\{A, B\}$ |
| $\bigcup \Gamma$ | union of a collection $\Gamma$ of sets |
| $A \cup B$ | $\bigcup\{A, B\}$ |
| $A \setminus B$ | relative complement of $A$ in $B$ |
| $B^c$ | $X \setminus B$ |
| $A \times B$ | Cartesian product of $A, B$ |
| $A + B$ | disjoint union of $A, B$ |
| $P(X)$ | power set of set $X$ |
| $\{x \in A \mid p(x)\}$ | subset of $A$ satisfying property $p$ |
| $X/E$ | quotient of $X$ by equivalence relation $E$ |
| | |
| *Relations* | |
| $I_A$ | identity function on $A$ |
| $S \circ R$ | composition of relations |
| $R^{-1}$ | inverse relation |
| $R[A]$ | image of $A$ |
| $R^{-1}[B]$ | inverse image of $B$ |

| | |
|---|---|
| *Functions* | |
| $f : A \to B$ | $f$ function with domain $A$ and target $B$ |
| $\{x \mapsto \phi(x) : A \to B\}$ | function from $A$ to $B$ given by formula $\phi$ |
| $\prod_{t\in I} A_t$ | Cartesian product of sets $A_t$, $t \in I$ |
| $A^I$ | Cartesian power of $A$, all functions from $I$ to $A$ |
| $\sum_{t\in I} A_t$ | disjoint union of sets $A_t$, $t \in I$ |
| $I \times A$ | disjoint multiple of $A$ |
| $\omega_0$ | countable infinite cardinality |
| $c$ | cardinality of the continuuum, $2^{\omega_0}$ |
| *Ordered sets* | |
| $\leq$ | generic order relation |
| $P, \leq$ | ordered set |
| $[a, b]$ | $\{x \in P \mid a \leq x \leq b\}$ |
| $(a, b)$ | $\{x \in P \mid a < x < b\}$ |
| $\downarrow S$ | lower bounds for $S$ |
| $\uparrow S$ | upper bounds for $S$ |
| $\bigwedge S$ | infimum $\inf S$, greatest lower bound |
| $x \wedge y$ | $\bigwedge\{x, y\}$ |
| $\bigvee S$ | $\sup S$, supremum, least upper bound |
| $x \vee y$ | $\bigvee\{x, y\}$ |
| *Number systems* | |
| $\mathbb{N}$ | natural numbers starting at 0 |
| $\mathbb{N}_+$ | natural numbers starting at 1 |
| $\mathbb{Z}$ | integers |
| $\mathbb{Q}$ | rational numbers |
| $\mathbb{R}$ | real numbers |
| $\mathbb{C}$ | complex numbers |
| $\mathbb{H}$ | quaternions |
| $\mathbf{N}$ | set ordered like $\mathbb{N}$ or $\mathbb{N}_+$ |
| $\mathbf{Z}$ | set ordered like $\mathbb{Z}$ |
| $\mathbf{Q}$ | set ordered like $\mathbb{Q}$ |
| $\mathbf{R}$ | set ordered like $\mathbb{R}$ |
| *Convergence* | |
| $\to$ | approaches, converges to |
| $\uparrow$ | increases to |
| $\downarrow$ | decreases to |
| *Metric spaces* | |
| $d$ | generic metric |
| $X, d$ | metric space |
| $B(x, r)$ | open ball about $x$ of radius $r$ |
| $\mathbb{R}^n$ | all ordered $n$-tuples of real numbers |

| | |
|---|---|
| $\ell^p$ | all $p$-summable sequences of real numbers, $1 \leq p < +\infty$ |
| $\ell^\infty$ | all bounded sequences of real numbers |
| $c_0$ | all sequences of real numbers that converge to zero |
| $\mathbb{R}^\infty$ | all infinite sequences of real numbers (product metric) |
| $[0,1]^\infty$ | Hilbert cube (product metric) |
| $2^\infty$ | Cantor set, coin-tossing space $\{0,1\}^\infty$ (product metric) |

*Topological spaces*

| | |
|---|---|
| $\mathcal{T}$ | generic topology |
| $X, \mathcal{T}$ | topological space |
| $\tau(\Gamma)$ | topology generated by subsets in $\Gamma$ |
| $\bar{A}$ | closure of subset $A$ |
| $A^\circ$ | interior of subset $A$ |
| $F_\sigma$ | countable union of closed subsets |
| $G_\delta$ | countable intersection of open subsets |

*Measurable spaces*

| | |
|---|---|
| $\mathcal{F}$ | generic $\sigma$-algebra (of subsets or of real functions) |
| $X, \mathcal{F}$ | measurable space |
| $\sigma(\Gamma)$ | $\sigma$-algebra generated by $\Gamma$ |
| $\mathcal{B}o(X) = \mathcal{B}o(X, \mathcal{T})$ | Borel $\sigma$-algebra $\sigma(\mathcal{T})$ |
| $\mathcal{B}c(X) = \mathcal{B}c(X, \mathcal{T})$ | continuous Baire $\sigma$-algebra $\sigma(C(X, \mathcal{T}))$ ($= \mathcal{B}o(X, \mathcal{T})$ for metrizable $X, \mathcal{T}$) |
| $\mathcal{B}a(X) = \mathcal{B}a(X, \mathcal{T})$ | Baire $\sigma$-algebra $\sigma(C_c(X, \mathcal{T}))$ |

*Integrals and measures*

| | |
|---|---|
| $\mu$ | generic (positive) measure |
| $X, \mathcal{F}, \mu$ | measure space |
| $f_+$ | positive part of $f$, $f_+ = f \vee 0$ |
| $f_-$ | negative part of $f$, $f_+ = -(f \wedge 0)$ |
| $\mu(f)$ | integral of $f$ with respect to $\mu$, $\mu(f) = \mu(f_+) - \mu(f_-)$ |
| $1_A$ | indicator function of $A$ |
| $\mu(A)$ | measure of $A$, same as $\mu(1_A)$ |
| $\delta_p$ | unit point mass at $p$, $\delta_p(g) = g(p)$ |
| $\sum$ | summation (counting measure) |
| $\phi[\mu]$ | image of integral $\mu$ under $\phi$, $\phi[\mu](g) = \mu(g \circ \phi)$ |
| $\nu \prec \mu$ | $\nu$ is absolutely continuous with respect to $\mu$ |
| $\nu \perp \mu$ | $\nu$ and $\mu$ are mutually singular |
| $\overline{\mathcal{F}_\mu}$ | completion of $\sigma$-algebra $\mathcal{F}$ with respect to $\mu$ |
| $\bar{\mu}$ | completion of integral $\mu$, measurable functions in $\bar{\mathcal{F}}_\mu$ |

*Product measures*

| | |
|---|---|
| $g \otimes h$ | tensor product of $g, h$, $(g \otimes h)(x,y) = g(x)h(y)$ |
| $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ | product $\sigma$-algebra |
| $\mu_1 \times \mu_2$ | product integral, $(\mu_1 \times \mu_2)(g \otimes h) = \mu_1(g)\mu_2(h)$ |
| $X_1 \times X_2, \mathcal{F}_1 \bigotimes \mathcal{F}_2, \mu_1 \times \mu_2$ | product measure space |

| | |
|---|---|
| $f^{\mid 1}$ | fix first input, $f^{\mid 1}(x) = \{y \mapsto f(x,y)\}$ |
| $f^{\mid 2}$ | fix second input, $f^{\mid 2}(y) = \{x \mapsto f(x,y)\}$ |
| $\mu_2(f \mid 1)$ | partial integral fixing first input, $\mu_2 \circ f^{\mid 1}$ |
| $\mu_1(f \mid 2)$ | partial integral fixing second input, $\mu_1 \circ f^{\mid 2}$ |

*Lebesgue integral*

| | |
|---|---|
| $\mathcal{B}o$ | Borel measurable functions $\mathcal{B}o(\mathbb{R})$ |
| $\lambda$ | Lebesgue integral with Borel measurable functions, $\lambda(g) = \int_{-\infty}^{\infty} g(x)\,dx$ |
| $\overline{\mathcal{B}o_\lambda}$ | Lebesgue measurable functions |
| $\bar{\lambda}$ | completed Lebesgue integral with Lebesgue measurable functions |
| $\sigma_F$ | Lebesgue-Stieltjes integral with Borel measurable functions, $\sigma_F(g) = \int_{-\infty}^{\infty}$ |

*Function spaces*

| | |
|---|---|
| $B(X)$ | all bounded real functions on set $X$ |
| $C(X) = C(X, \mathcal{T})$ | all continuous real functions on topological space $X, \mathcal{T}$ |
| $BC(X) = BC(X, \mathcal{T})$ | all bounded continuous real functions on topological space $X, \mathcal{T}$ |
| $C_c(X) = C_c(X, \mathcal{T})$ | all real compactly-supported functions on LCH space $X, \mathcal{T}$ |
| $C_0(X) = C_0(X, \mathcal{T})$ | all real functions on LCH space $X, \mathcal{T}$ that vanish at infinity |
| $\\|f\\|_{\sup}$ | supremum norm of function $f$ |
| $\mathcal{L}^p(X, \mu) = \mathcal{L}^p(X, \mathcal{F}, \mu)$ | $\mathcal{L}^p$ space of functions, $1 \le p \le \infty$ |
| $L^p(X, \mu) = L^p(X, \mathcal{F}, \mu)$ | quotient space by $\mu$-equivalent functions, $1 \le p \le \infty$ |
| $\\|f\\|_p$ | $\mathcal{L}^p$ or $L^p$ norm, $1 \le p \le \infty$ |
| $M(X) = M(X, \mathcal{F})$ | space of finite signed measures on $X, \mathcal{F}$ |
| $\mu_+$ | positive part of signed measure $\mu$ |
| $\mu_-$ | negative part of signed measure $\mu$ |
| $\\|\mu\\|$ | variation norm $\mu_+(X) + \mu_-(X)$, where $\mu = \mu_+ - \mu_-$ |

*Banach spaces*

| | |
|---|---|
| $\\|u\\|$ | norm of vector $u$ in Banach space $E$ |
| $T : E \to F$ | $T$ is continuous linear from $E$ to $F$ |
| $\\|T\\|$ | Lipschitz norm of $T$, $\\|Tu\\| \le \\|T\\|\\|u\\|$ |
| $E^*$ | dual space of $E$, space of $\alpha : E \to \mathbb{R}$ |
| $\langle \alpha, v \rangle$ | value $\alpha(v)$ of $\alpha$ in $E^*$ on $v$ in $E$ |
| $T^* : F^* \to E^*$ | adjoint of $T$, $\langle T^*\alpha, v \rangle = \langle \alpha, Tv \rangle$ |

*Hilbert space*

| | |
|---|---|
| $\langle u, v \rangle$ | inner product of vectors $u, v$ in Hilbert space $H$ |
| $\\|v\\|$ | norm $\\|v\\| = \sqrt{\langle v, v \rangle}$ of $v$ in Hilbert space $H$ |
| $u \perp v$ | $u$ is orthogonal (perpendicular) to $v$, $\langle u, v \rangle = 0$ |
| $M^\perp$ | closed subspace of vectors orthogonal to $M$ |
| $T : H \to K$ | $T$ is continuous linear from $H$ to $K$ |
| $T^* : K \to H$ | Hilbert space adjoint of $T$, $\langle T^*u, v \rangle = \langle u, Tv \rangle$ |
| $H^*$ | dual space of $H$, space of $\alpha : H \to \mathbb{C}$ |
| $w^*$ | adjoint $w^*$ in $H^*$ of $w$ in $H$, $w^*(u) = \langle w, u \rangle$ |

*Fourier transform*

| | |
|---|---|
| $f * g$ | convolution of $f$ and $g$ , $\|f * g\|_2 \leq \|f\|_1 \|g\|_2$ |
| $f^*$ | convolution adjoint of function $f$, $\langle f^* * h, g \rangle = \langle h, f * g \rangle$ |
| $\hat{f}$ | Fourier transform of $f$, $\widehat{f * g} = \hat{f}\hat{g}$, $\widehat{f^*} = \bar{f}$ |
| $F$ | Fourier transform on $L^2$, $Fg = \hat{g}$ |

*Geometry*

| | |
|---|---|
| $T = S_1$ | circle |
| $T^n$ | $n$-torus |
| $S_{n-1}$ | unit $n-1$ sphere of area $a_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}$ |
| $B_n$ | open unit $n$ ball of volume $v_n = a_n/n$ |

# Index