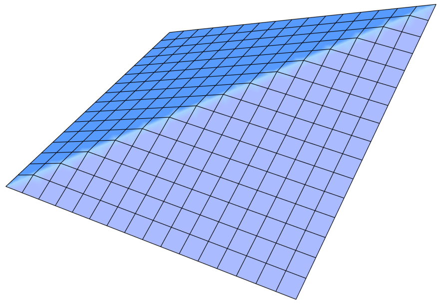


PMATH 351: Real Analysis

Prof. Michael Harz • Fall 2014 • University of Waterloo
MWF 11:30AM–12:20PM in MC 4041

Transcribed, LaTeXed, and edited by Rob Zimmerman

Note: These are personal lecture notes, not official course materials. They may contain errors (which by default you should attribute to me, Rob Zimmerman, rather than the course instructor). The permanent home of these — and all of my other lecture notes — is https://rob-zimmerman.github.io/coursenotes.

Contents

1 Set Theory and Cardinality **2**
Foundations of Set Theory . . . 3
The Axiom of Choice . . . 5
Cardinality . . . 11
Cardinal Arithmetic . . . 17

2 Metric Spaces **24**
Definitions and Examples 24
Topology of Metric Spaces 31
Sequences and Topology in Metric Spaces 39
Continuity in Metric Spaces 43

3 Compactness **45**
Topological Spaces 46
Compactness 49
Compactness in Metric Spaces 56
Applications of Compactness 61

4 Completeness **63**
Complete Metric Spaces 63
Banach and Hilbert Spaces 67
The Relationship Between Compactness and Completeness 71
Completions of Metric Spaces 72
The Baire Category Theorem 76

5 Spaces of Continuous Functions **81**
The Weierstrass Approximation Theorem 81
The Stone–Weierstrass Theorem 84
The Arzelà-Ascoli Theorem 92
Urysohn's Lemma and the Tietze Extension Theorem 98

Appendix: Lecture and Tutorial Index **105**

1. Set Theory and Cardinality

Lecture 1 — Monday, September 08, 2014

Foundations of Set Theory

We will discuss concepts used for this course, but not develop an axiomatic approach. In particular, we will not define what a set is; we take that for granted, as well as elements and belonging. The empty set, which has no elements, is denoted by $\emptyset$. Given sets A and B, we write $A \subseteq B$ to mean that $a \in A \implies a \in B$. We have $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$; in axiomatic set theory, this is called the **Axiom of Extension**. Given a set A and a sentence S, we can define the subset B of A which contains the elements of A that satisfy S: that is, $B = \{a \in A \mid S(a)\}$. We can do this by the **Axiom of Specification**.

Example 1.1 Let $\mathbb{N} = \{0, 1, 2, \ldots\}$, the set of all natural numbers. Then $B = \{n \in \mathbb{N} \mid n \text{ is even}\} = \{0, 2, 4, \ldots\}$.

Definition 1.2 Let X be a set. Then $P(X) = \{A \mid A \subseteq X\}$ is called the **power set** of X.

Definition 1.3 Let $\Lambda \neq \emptyset$ and suppose that $\{X_\lambda\}_{\lambda \in \Lambda}$ is a collection of sets. The **union** of the X_λ is

$$\bigcup_{\lambda \in \Lambda} X_\lambda = \{x \mid x \in X_\lambda \text{ for some } \lambda \in \Lambda\}$$

and the **intersection** of the X_λ is

$$\bigcap_{\lambda \in \Lambda} X_\lambda = \{x \mid x \in X_\lambda \text{ for all } \lambda \in \Lambda\}.$$

We define $\bigcup_{\lambda \in \emptyset} X_\lambda = \emptyset$. Assume that we are given a set X such that $X_\lambda \subseteq X$ for all $\lambda \in \Lambda$. We define $\bigcap_{\lambda \in \emptyset} X_\lambda = X$. The condition "$x \in \bigcap_{\lambda \in \emptyset} X_\lambda$" requires $x \in X_\lambda$ for every $\lambda \in \emptyset$. There are no indices to check, so the condition is vacuously true and the empty intersection contains every point of the ambient set X.

Definition 1.4 Let $A, B \subseteq X$ be sets. Let

$$B \setminus A = \{x \in B \mid x \notin A\}.$$

If $B = X$, then we call $X \setminus A$ the **complement** of A in X and denote this set by A^c.

Naive set theory leads to issues. Consider the set $A = \{B \mid B \notin B\}$. Is $A \in A$? If so, then $A \notin A$. If not, then $A \in A$. The problem with the definition of A is that its elements are not taken from a fixed set. Given a set X (a "universe"), we may define $C = \{B \in X \mid B \notin B\}$ by the Axiom of Specification. By the argument above, $C \notin X$. Since X was arbitrary, there is no universe that contains everything.

Let $X_1, \dots, X_n$ be sets. Recall that the Cartesian product of the X_i is defined by

$$X_1 \times \cdots \times X_n := \prod_{i=1}^{n} X_i := \{(x_1, \dots, x_n) \mid x_i \in X_i \text{ for } i = 1, \dots, n\}.$$

Every $(x_1, \dots, x_n) \in \prod_{i=1}^{n} X_i$ determines a function

$$f : \{1, \dots, n\} \to \bigcup_{i=1}^{n} X_i$$

by $i \mapsto x_i$ satisfying $f(i) \in X_i$ for $i = 1, \dots, n$. Conversely, given

$$f : \{1, \dots, n\} \to \bigcup_{i=1}^{n} X_i$$

satisfying $f(i) \in X_i$ for $i = 1, \dots, n$, define

$$(x_1, \dots, x_n) \in \prod_{i=1}^{n} X_i$$

by $x_i = f(i)$ for $i = 1, \dots, n$. These constructions are inverse to each other.

Definition 1.5 Let $\Lambda \neq \emptyset$ and let $\{X_\lambda\}_{\lambda \in \Lambda}$ be a collection of sets. We define the **Cartesian product** of the X_λ as:

$$\prod_{\lambda \in \Lambda} X_\lambda = \{f : \Lambda \to \bigcup_{\lambda \in \Lambda} X_\lambda \mid f(\lambda) \in X_\lambda \text{ for all } \lambda \in \Lambda\}.$$

Lecture 2 — Wednesday, September 10, 2014

Definition 2.1 Let $\Lambda \neq \emptyset$ and let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a collection of sets. We define

$$\prod_{\lambda\in\Lambda} X_\lambda = \{f : \Lambda \to \bigcup_{\lambda\in\Lambda} X_\lambda \mid f(\lambda) \in X_\lambda \text{ for all } \lambda \in \Lambda\}.$$

Every such function is called a **choice function**. Given sets X, Y, we define

$$X^Y = \{f : Y \to X \mid f \text{ is a function}\} = \prod_{y\in Y} X_y$$

where $X_y = X$ for all $y \in Y$.

Given nonempty sets $\{X_\lambda\}_{\lambda\in\Lambda}$, is $\prod_{\lambda\in\Lambda} X_\lambda \neq \emptyset$? That is, do choice functions always exist? Here are some examples:

1. If $\Lambda \neq \emptyset$ is finite, then choice functions always exist.
2. Let $\Lambda \neq \emptyset$ and suppose $\emptyset \neq X_\lambda \subseteq \mathbb{N}$ for $\lambda \in \Lambda$. Let $f(\lambda) = \min\{X_\lambda\}$. Then $f \in \prod_{\lambda\in\Lambda} X_\lambda$ is a choice function.
3. Suppose $\Lambda \neq \emptyset$ and for $\lambda \in \Lambda$, let $X_\lambda = \{L_\lambda, R_\lambda\}$ be a pair of shoes. Let $f(\lambda) = R_\lambda$. Then $f \in \prod_{\lambda\in\Lambda} X_\lambda$ is a choice function.
4. For $n \in \mathbb{N}$, let X_n consist of a pair of identical-looking socks. How do we find $f \in \prod_{n\in\mathbb{N}} X_n$? This is a problem!

The Axiom of Choice

The question of whether choice functions always exist cannot be answered from the basic axioms of set theory.

Definition 2.2 (Axiom of choice) If $\Lambda \neq \emptyset$ and for each $\lambda \in \Lambda$, X_λ is a nonempty set, then

$$\prod_{\lambda\in\Lambda} X_\lambda \neq \emptyset.$$

The choice function in Figure 1 selects one element from every member of the family, even when no uniform selection rule is available.

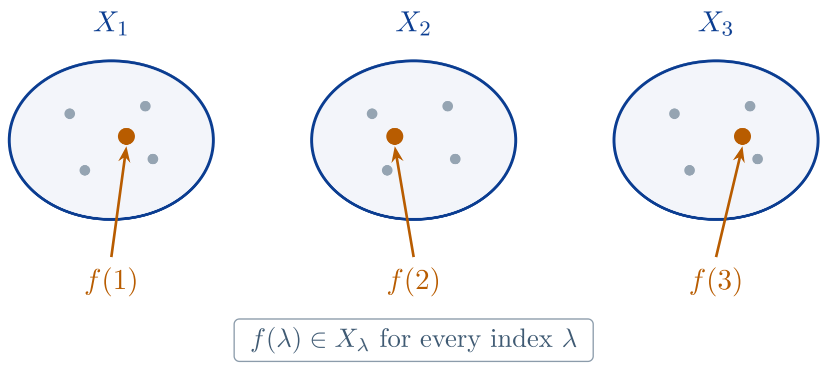


Figure 1

Proposition 2.3 The following statement is equivalent to the axiom of choice: given a nonempty set X, there is a function $f : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ such that $f(A) \in A$ for all $A \in \mathcal{P}(X) \setminus \{\emptyset\}$.

Proof. *Forward implication.* Assume the axiom of choice. Let $X \neq \emptyset$, and set $\Lambda = \mathcal{P}(X) \setminus \{\emptyset\}$. For each $A \in \Lambda$, let $X_A := A$. By the axiom of choice, there exists $(x_A)_{A\in\Lambda} \in \prod_{A\in\Lambda} X_A$, that is, $x_A \in A$ for each A. Define $f : \Lambda \to X$ by $f(A) = x_A$. Then $f(A) \in A$ for all $A \in \Lambda$.

Reverse implication. Conversely, assume the stated property of choice functions. Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of nonempty sets. Put $Y = \bigcup_{\lambda\in\Lambda} X_\lambda$. By assumption, there is $c : \mathcal{P}(Y) \setminus \{\emptyset\} \to Y$ with $c(S) \in S$ for all nonempty $S \subseteq Y$. Define $g : \Lambda \to Y$ by $g(\lambda) = c(X_\lambda)$. Then $g(\lambda) \in X_\lambda$ for all λ, so $(g(\lambda))_{\lambda\in\Lambda} \in \prod_{\lambda\in\Lambda} X_\lambda$. Hence the axiom of choice holds. □

The axiom of choice has some counterintuitive consequences, such as the Banach–Tarski paradox: it is possible to carve up a unit ball in $\mathbb{R}^3$ into finitely many pieces and, using only translations and rotations, reassemble those pieces into two balls, each of which has the same volume as the original. The proof is nonconstructive (world hunger is still a problem). On the other hand, the negation of the axiom of choice also has dreadful consequences. For example, it implies that there are sets A and B such that neither can be mapped injectively into the other (this is bad). We generally assume the axiom of choice, but mention when we use it in a proof.

Definition 2.4 A **relation** R on a set X is a subset of $X \times X$. We write xRy if $(x, y) \in R \subseteq X \times X$. A relation "$\leq$" is called a **partial order** on X if it satisfies

1. $x \leq x$ (**reflexivity**),
2. $x \leq y$ and $y \leq z \implies x \leq z$ (**transitivity**),
3. $x \leq y$ and $y \leq x \implies x = y$ (**antisymmetry**).

The pair $(X, \leq)$ is called a **partially ordered set** (or a **poset**). A **chain** C in X is a subset of X such that, for $x, y \in C$, either $x \leq y$ or $y \leq x$ (these are also called **totally ordered** or **linearly ordered** sets).

Example 2.5

1. $\mathbb{R}$ with the usual order is a totally ordered set (and hence a partially ordered set).
2. Let $X \neq \emptyset$ and define a relation "$\leq$" on $\mathcal{P}(X)$ by $A \leq B$ if and only if $A \subseteq B$. Then $(\mathcal{P}(X), \leq)$ is a poset. We say that $\mathcal{P}(X)$ is **ordered by inclusion**. If X has more than one element, then $(\mathcal{P}(X), \leq)$ is not totally ordered. Why? If $a, b \in X$ with $a \neq b$, then $\{a\} \not\subseteq \{b\}$ and $\{b\} \not\subseteq \{a\}$. For example, let $X = \{1, 2, 3\}$. Then $\{\emptyset, \{1\}, \{1, 3\}\}$ is a chain, but $\{\{1\}, \{2\}, \{2, 3\}\}$ is not a chain because $\{1\} \not\subseteq \{2, 3\}$.
3. Let $X \neq \emptyset$ and define a relation "$\leq$" on $\mathcal{P}(X)$ by $A \leq B$ if and only if $B \subseteq A$. Then $(\mathcal{P}(X), \leq)$ is a poset. We say that $\mathcal{P}(X)$ is **ordered by containment**.
4. Let $C([0, 1], \mathbb{R}) = \{f : [0, 1] \to \mathbb{R} \mid f \text{ is continuous}\}$. Say $f \leq g$ if and only if $f(x) \leq g(x)$ for all $x \in [0, 1]$. This is a partial order but not a total order.

Definition 2.6 Let $(X, \leq)$ be a poset. We say that $x \in X$ is **maximal** if $y \in X$ and $x \leq y$ imply $y = x$. We say $m \in X$ is a **maximum element** if $m \geq y$ for all $y \in X$. We say that $z \in X$ is **minimal** in X if $y \in X$ and $y \leq z$ imply $y = z$. An element $n \in X$ is called a **minimum element** if $n \leq y$ for all $y \in X$.

Every maximum element is maximal, but not vice versa. A maximal element need not be comparable to every element.

Lecture 3 — Friday, September 12, 2014

Example 3.1

1. $(\mathbb{R}, \leq)$ where $\leq$ is the usual order has no maximal elements (and hence no maximum element).
2. Let $X = \{1,2,3\}$ and consider $\mathcal{P}(X) \setminus \{X\}$ ordered by inclusion. Then $\{1,2\}, \{1,3\}, \{2,3\}$ are maximal in $\mathcal{P}(X) \setminus \{X\}$, but none are maximum elements. Indeed, $\{1,2\} \not\subseteq \{1,3\}$. In fact, $\mathcal{P}(X) \setminus \{X\}$ has no maximum element.
3. Consider
$$C([0,1], \mathbb{C}) = \{f : [0,1] \to \mathbb{C} \mid f \text{ continuous}\}$$
and let $(X, \leq)$ be the set of proper ideals of $(C([0,1], \mathbb{C}))$ ordered by inclusion. For each $y \in [0,1]$, let
$$P_y : C([0,1], \mathbb{C}) \to \mathbb{C}, \quad f \mapsto f(y).$$
Then P_y is linear and multiplicative, hence $\mathfrak{m}_y = \ker(P_y)$ is an ideal in $C([0,1], \mathbb{C})$. Note that $\ker(P_y) = \{f \in C([0,1], \mathbb{C}) \mid f(y) = 0\}$. By linear algebra,
$$C([0,1], \mathbb{C}) / \ker(P_y) \cong \text{Range}(P_y) = \mathbb{C}$$
, as P_y is surjective. That is, $\ker(P_y)$ has codimension 1 in $C([0,1], \mathbb{C})$, hence $\ker(P_y)$ is a maximal ideal. It can be shown that these are all maximal ideals.
4. Every finite poset has a maximal element.

Definition 3.2 Let X be a poset and let $A \subseteq X$. We say that $x \in X$ is an **upper bound** for A if $y \leq x$ for all $y \in A$. If A has an upper bound, we say that A is **bounded above**. An element $x \in X$ is called a **least upper bound** for A if x is an upper bound for A, and if z is an upper bound for A, then $x \leq z$. We also say that x is the **supremum** for A. **Lower bounds** and **greatest lower bounds** (**infima**) are defined similarly.

Example 3.3

1. $\mathbb{R}$ with the usual order has the **least upper bound property**: if $\emptyset \neq A \subset \mathbb{R}$ is bounded above, then A has a least upper bound.
2. Let X be a set and consider $\mathcal{P}(X)$ ordered by inclusion. Given $\{X_\lambda\}_{\lambda \in \Lambda} \subseteq \mathcal{P}(X)$, the union $\bigcup_\lambda X_\lambda$ is the least upper bound and $\bigcap_\lambda X_\lambda$ is the greatest lower bound.
3. Consider $\mathbb{Q}$ with the usual order $\leq$. Then $\{x \in \mathbb{Q} \mid x^2 \leq 2\}$ is bounded above, but it does not have a least upper bound (because $\sqrt{2}$ is irrational).

Theorem 3.4 (Zorn's lemma) Suppose $\emptyset \neq X$ and $(X, \leq)$ is a poset. If every chain in X has an upper bound, then $(X, \leq)$ has a maximal element.

We omit the proof of Zorn's lemma, as well as the proof that this statement is equivalent to the axiom of choice.

Definition 3.5 Let V be a k-vector space, where k is a field (for example, $k = \mathbb{R}$). A subset $B \subset V$ is called **linearly independent** if every finite subset of B is linearly independent. It is called a **spanning set** if every $v \in V$ can be expressed as a finite linear combination

$$v = \sum \alpha_i b_i,$$

where $\alpha_i \in k$ and $b_i \in B$. A **basis** is a linearly independent spanning set.

Theorem 3.6 Every vector space has a basis.

Proof. If $V = \{0\}$, then $\emptyset$ is a basis. Assume $V \neq \{0\}$. Let

$$\mathcal{L} = \{A \subset V \mid A \text{ is linearly independent}\},$$

ordered by inclusion. Then $\mathcal{L} \neq \emptyset$, since $\{v\} \in \mathcal{L}$ for every $v \in V \setminus \{0\}$. Let $\{A_\lambda\}_{\lambda \in \Lambda}$ be a chain in $\mathcal{L}$ and define $A = \bigcup_{\lambda \in \Lambda} A_\lambda$. We claim that $A \in \mathcal{L}$.

Let $\{x_1, \ldots, x_n\} \subseteq A$ be a finite subset. For $i = 1, \ldots, n$, there is $\lambda_i \in \Lambda$ such that $x_i \in A_{\lambda_i}$. Since $\{A_\lambda\}_{\lambda \in \Lambda}$ is a chain, one of the sets A_{λ_i} contains all the others, so $\{x_1, \ldots, x_n\}$ is contained in a single member of the chain. Since that member is linearly independent, $\{x_1, \ldots, x_n\}$ is linearly independent. Thus $A \in \mathcal{L}$, and every chain in $\mathcal{L}$ has an upper bound.

By Zorn's lemma, $\mathcal{L}$ has a maximal element B. We claim that B is a basis. We know $B \in \mathcal{L}$, so it is enough to show B spans V. To this end, let $v \in V$. If $v \in B$, then v is a linear combination of elements in B. If $v \notin B$, then $B \cup \{v\} \supsetneq B$. Maximality of B means that $B \cup \{v\}$ is not linearly independent. Thus, there exist $\alpha, \alpha_1, \ldots, \alpha_n \in k$ and $b_1, \ldots, b_n \in B$ such that $\alpha v + \alpha_1 b_1 + \cdots + \alpha_n b_n = 0$. Note $\alpha \neq 0$ as $\{b_1, \ldots, b_n\}$ is linearly independent. Thus,

$$v = -\frac{1}{\alpha} \sum_{i=1}^{n} \alpha_i b_i.$$

Hence B spans V. □

Lecture 4 — Monday, September 15, 2014

Definition 4.1 A nonempty poset $(X, \leq)$ is said to be **well-ordered** if every nonempty subset $A \subseteq X$ has a minimum element.

Remark 4.2 Well-ordered sets are totally ordered. (Apply the definition to two-element subsets).

Example 4.3 $\mathbb{N}$ with the usual order is well-ordered, but $\mathbb{R}$ is not.

Theorem 4.4 The following are equivalent:

1. The axiom of choice
2. Zorn's lemma
3. Every nonempty set admits a well-ordering (this is the **well-ordering principle**).

Most of the proof is postponed to PMATH 433, but we do record one implication.

Proof of (3) $\implies$ (1). Assume (3), the well-ordering principle. Let $\{X_\lambda\}_{\lambda \in \Lambda}$ be a collection of nonempty sets and put $X = \bigcup_{\lambda \in \Lambda} X_\lambda$. By assumption, there is a well-ordering on X. For each $\lambda \in \Lambda$, let $f(\lambda)$ be the least element of $X_\lambda \subseteq X$. Then $f : \Lambda \to \bigcup_{\lambda \in \Lambda} X_\lambda$ is a choice function, so the axiom of choice holds. □

Cardinality

Definition 4.5 A relation $\sim$ on a set X is an **equivalence relation** if it satisfies:

1. $x \sim x$ for all $x \in X$ (**reflexivity**),
2. $x \sim y \implies y \sim x$ (**symmetry**), and
3. $x \sim y, y \sim z \implies x \sim z$ (**transitivity**).

Given $x \in X$, the set $[x] = \{y \in X \mid x \sim y\}$ is called the **equivalence class** of x. Any two equivalence classes are either equal or disjoint.

Example 4.6

1. If X is a set, define a relation $\sim$ on X by $x \sim y$ if $x = y$. Then $\sim$ is an equivalence relation.
2. Define a relation $\sim$ on $\mathbb{Z}$ by $x \sim y$ if $y - x$ is even. Then $\sim$ is an equivalence relation. Our equivalence classes are the evens and the odds.
3. Let X be a set. Define a relation $\sim$ on $\mathcal{P}(X)$ by $A \sim B$ if there exists a bijection from A to B. Then $\sim$ is an equivalence relation: the identity gives reflexivity, inverses give symmetry, and composition gives transitivity.

Definition 4.7 Two sets A and B are said to be **equinumerous** if there exists a bijection from A to B, and we write $A \sim B$. We also say that A and B have the same **cardinality**, and we also write $|A| = |B|$.

Definition 4.8 A set A is said to be **finite** if $A = \emptyset$ or $A \sim \{1, 2, \ldots, n\}$ for some $n \in \mathbb{N}$, in which case we write $|A| = n$. Otherwise, we say that A is **infinite**. We say that A is **countably infinite** if there is a bijection $A \sim \mathbb{N}$, and we say that A is **countable** if it is finite or countably infinite.

If A is finite, then it is not equinumerous to a proper subset of itself (that is, there is no B such that $B \subsetneq A$ and $A \sim B$). This follows from the pigeonhole principle: if $|A| = n$ and $|B| = m < n$, then $A \not\sim B$.

Example 4.9

1. The set $\mathbb{Z}$ is countably infinite. Indeed,
$$f : \mathbb{Z} \to \mathbb{N}, n \mapsto \begin{cases} 2n & n \geq 0 \\ -2n-1 & n < 0 \end{cases},$$
so $\mathbb{Z} \sim \mathbb{N}$ even though $\mathbb{N} \subsetneq \mathbb{Z}$.

2. $\mathbb{N} \times \mathbb{N} \sim \mathbb{N}$. The diagonal traversal in Figure 2 constructs a bijection $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$.

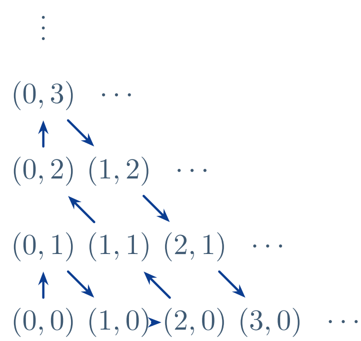


Figure 2

Proposition 4.10 Every infinite set contains a countably infinite subset.

Proof. Let X be an infinite set. By the axiom of choice, or rather by Proposition 2.3, there exists a function $f : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ such that $f(A) \in A$ for all $A \subseteq X, A \neq \emptyset$. Define $g : \mathbb{N} \to X$ recursively by $g(0) = f(X)$ and $g(n) = f(X \setminus \{g(0), g(1), \ldots, g(n-1)\})$ for $n \geq 1$. (Note: $X \setminus \{g(0), \ldots, g(n-1)\} \neq \emptyset$ since X is infinite.) Then g is injective, so $g(\mathbb{N}) \sim \mathbb{N}$ is a countably infinite subset of X. $\square$

Definition 4.11 Let A, B be sets. We write $|A| \leq |B|$ if there exists an injective function f from A to B.

Note that if $|A| = |B|$, then $|A| \leq |B|$. Also, if $|A| \leq |B|$ and $|B| \leq |C|$, then $|A| \leq |C|$.

Example 4.12 Define a function $f : \mathbb{Q} \to \mathbb{N}$ by

$$f\left(\frac{m}{n}\right) = \begin{cases} 2^m 3^n & m > 0 \\ 0 & m = 0 \\ 5^{-m} 7^n & m < 0 \end{cases}$$

where $m \in \mathbb{Z}, n \in \mathbb{N}, n \geq 1$, and $\gcd(m, n) = 1$. Then f is injective, so $|\mathbb{Q}| \leq |\mathbb{N}|$.

Lecture 5 — Wednesday, September 17, 2014

Recall that $|\mathbb{Q}| = |\mathbb{N}|$. Note that $|\mathbb{N}| \leq |\mathbb{Q}|$. Does it follow that $|\mathbb{N}| = |\mathbb{Q}|$? More generally, if $|A| \leq |B|$ and $|B| \leq |A|$, does it follow that $|A| = |B|$?

Proposition 5.1 Given nonempty sets A, B, the following are equivalent:

1. $|A| \leq |B|$ (that is, there exists an injection $f : A \to B$)
2. There exists a surjection $g : B \to A$

Proof. (1) *implies* (2). Let $f : A \to B$ be an injection. Then $f : A \to f(A)$ is a bijection. Let $a_0 \in A$ be arbitrary and define $g : B \to A$ by

$$b \mapsto \begin{cases} f^{-1}(b) & b \in f(A) \\ a_0 & b \notin f(A) \end{cases}.$$

Then g is surjective.

(2) *implies* (1). Let $g : B \to A$ be surjective (that is, $g^{-1}(\{a\}) \neq \emptyset$ for all $a \in A$).

By the axiom of choice, there is a function $h : \mathcal{P}(B) \setminus \{\emptyset\} \to B$ such that $h(L) \in L$ for every $L \in \mathcal{P}(B) \setminus \{\emptyset\}$. Define $f : A \to B$ by $a \mapsto h(g^{-1}(\{a\}))$. Then f is injective, as $g(f(a)) = a$ for all $a \in A$. $\square$

Lemma 5.2 Let X be a set and let $\phi : \mathcal{P}(X) \to \mathcal{P}(X)$ be increasing in the following sense: $A \subseteq B \subseteq X$ implies $\phi(A) \subseteq \phi(B)$. Then there exists $T \in \mathcal{P}(X)$ such that $\phi(T) = T$.

Proof. Define
$$T = \bigcup\{Y \subseteq X \mid Y \subseteq \phi(Y)\}.$$
If $Y \subseteq X$ and $Y \subseteq \phi(Y)$, then $Y \subseteq T$, hence $\phi(Y) \subseteq \phi(T)$. Since $Y \subseteq \phi(Y)$, it follows that $Y \subseteq \phi(T)$. Therefore
$$T = \bigcup\{Y \mid Y \subseteq \phi(Y)\} \subseteq \phi(T).$$
Now, $\phi(T) \subseteq \phi(\phi(T))$, so $\phi(T)$ is one of the sets Y in the definition of T. Thus $\phi(T) \subseteq T$, and $\phi(T) = T$. □

Theorem 5.3 (Schroeder–Bernstein theorem) Let A, B be sets. If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.

Proof. By assumption there are injections $f : A \to B$ and $g : B \to A$. Define $\phi : \mathcal{P}(A) \to \mathcal{P}(A)$ by $X \mapsto A \setminus g(B \setminus f(X))$. Then ϕ is increasing. By Lemma 5.2, there is a subset $T \subseteq A$ such that $\phi(T) = T$; that is,
$$T = A \setminus g(B \setminus f(T)).$$
Then $A \setminus T = g(B \setminus f(T))$. Observe that g is a bijection between $B \setminus f(T)$ and $A \setminus T$. Define $h : A \to B$ by
$$h(a) = \begin{cases} f(a), & a \in T, \\ g^{-1}(a), & a \notin T. \end{cases}$$
Then h is a bijection. □

That proof did not require the axiom of choice!

Corollary 5.4 A set X is countable if and only if $|X| \leq |\mathbb{N}|$. In particular, subsets of countable sets are countable.

Proof. *Forward implication.* This follows from our definition of countability.

Reverse implication. Assume $|X| \leq |\mathbb{N}|$. If X is finite, we are done. If X is infinite, then there exists an infinite subset $Y \subseteq X$ such that $|Y| = |\mathbb{N}|$. Hence $|\mathbb{N}| \leq |X|$. By the Schroeder–Bernstein theorem, $|X| = |\mathbb{N}|$. That is, X is countably infinite. If $Y \subseteq X$ is a subset of a countable set, then $|Y| \leq |X| \leq |\mathbb{N}|$, so Y is countable by the first part. $\square$

Corollary 5.5 $\mathbb{Q}$ is countable.

Proof. We know that $|\mathbb{Q}| \leq |\mathbb{N}|$, so apply Corollary 5.4. $\square$

Proposition 5.6 Let $\{X_n\}_{n\in\mathbb{N}}$ be a collection of countable sets. Then $X = \bigcup_{n\in\mathbb{N}} X_n$ is countable.

Proof. If $X = \emptyset$, then there is nothing to prove. Otherwise, fix $x_0 \in X$ and put

$$Y_n = \begin{cases} X_n, & X_n \neq \emptyset, \\ \{x_0\}, & X_n = \emptyset. \end{cases}$$

Then every Y_n is a nonempty countable set and $X = \bigcup_{n\in\mathbb{N}} Y_n$. For every $n \in \mathbb{N}$, choose a surjection $f_n : \mathbb{N} \to Y_n$. Define $f : \mathbb{N} \times \mathbb{N} \to X$ by $(m, n) \mapsto f_n(m)$. Then f is surjective, so $|X| \leq |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$ by Example 4.9(2). Hence X is countable. $\square$

Lecture 6 — Friday, September 19, 2014

Theorem 6.1 $\mathbb{R}$ is uncountable.

Proof. We use **Cantor's diagonal argument** to show that $(0, 1)$ is uncountable. Assume that $(0, 1)$ is countable. Then we can write $(0, 1) = \{x_1, x_2, x_3, \ldots\}$. Write x_n in a decimal expansion $x_n = 0.d_{n1}d_{n2}d_{n3}\ldots$. Define $y = 0.y_1y_2y_3\ldots$ where

$$y_n = \begin{cases} 7, & \text{if } d_{nn} \in \{0, 1, \ldots, 8\} \\ 4, & \text{if } d_{nn} = 9 \end{cases}.$$

Then $y \neq x_n$ for all n, so y does not appear in our list; but certainly $y \in (0,1)$, a contradiction. Hence $(0,1)$ is uncountable. □

Definition 6.2 If X, Y are sets, we write $|X| < |Y|$ if $|X| \leq |Y|$ but $|X| \neq |Y|$ (that is, there is an injection from X to Y, but no bijection).

Thus, $|\mathbb{N}| < |\mathbb{R}|$ by what we just proved.

Theorem 6.3 For any set X, we have $|X| < |\mathcal{P}(X)|$.

Proof. Note that $X \to \mathcal{P}(X)$, $x \mapsto \{x\}$ is an injection. Now, assume that $f : X \to \mathcal{P}(X)$ is surjective. Let $A = \{x \in X \mid x \notin f(x)\}$. Since f is surjective, there is a $y \in X$ such that $A = f(y)$. If $y \in A$, then $y \notin f(y) = A$, a contradiction. If $y \notin A$, then $y \in f(y) = A$, a contradiction as well. Thus the whole assumption was false; f is not onto. □

It follows that there are infinitely many infinite sets of different cardinality:

$$|\mathbb{N}| < |\mathcal{P}(\mathbb{N})| < |\mathcal{P}(\mathcal{P}(\mathbb{N}))| < \dots.$$

Theorem 6.4 $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$.

Proof. Define $f : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by $A \mapsto 0.a_1a_2a_3\dots$ where $a_n = 1$ if $n \in A$, $a_n = 0$ if $n \notin A$. Then f is injective, so $|\mathcal{P}(\mathbb{N})| \leq |\mathbb{R}|$.

To see that $|\mathbb{R}| \leq |\mathcal{P}(\mathbb{N})|$, note that $|\mathbb{N}| = |\mathbb{Q}|$. Choose a bijection $\phi : \mathbb{N} \to \mathbb{Q}$. Then

$$\Phi : \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{Q}), \qquad A \mapsto \phi(A)$$

is a bijection with inverse $B \mapsto \phi^{-1}(B)$, so $|\mathcal{P}(\mathbb{N})| = |\mathcal{P}(\mathbb{Q})|$. Define $g : \mathbb{R} \to \mathcal{P}(\mathbb{Q})$ by $x \mapsto \{q \in \mathbb{Q} \mid q < x\}$. If $x_1 < x_2$ in $\mathbb{R}$, there is a $q \in \mathbb{Q}$ such that $x_1 < q < x_2$, so $q \in g(x_2)$ but $q \notin g(x_1)$, so $g(x_1) \neq g(x_2)$. Thus g is injective, so $|\mathbb{R}| \leq |\mathcal{P}(\mathbb{Q})|$. By the Schroeder–Bernstein theorem, $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$. □

Theorem 6.5 If A and B are sets, then $|A| \leq |B|$ or $|B| \leq |A|$, no matter what.

Proof. Consider $\mathcal{S} = \{(X, Y, f) \mid X \subseteq A, Y \subseteq B, f \text{ is a bijection from } X \text{ to } Y\}$. Note that $\mathcal{S} \neq \emptyset$ (for example, $(\emptyset, \emptyset, e) \in \mathcal{S}$, where e is the empty function). We partially order $\mathcal{S}$ by $(X_1, Y_1, f_1) \leq (X_2, Y_2, f_2)$ if $X_1 \subseteq X_2, Y_1 \subseteq Y_2$, and $f_2\Big|_{X_1} = f_1$. Let $\{(X_\lambda, Y_\lambda, f_\lambda)\}_{\lambda \in \Lambda}$ be a chain in $\mathcal{S}$. Let $X = \bigcup_{\lambda \in \Lambda} X_\lambda$ and $Y = \bigcup_{\lambda \in \Lambda} Y_\lambda$ and define $f : X \to Y$ in the following way: if $x \in X_\lambda$, then $f(x) := f_\lambda(x)$. This function is well-defined: if $x \in X_{\lambda_1} \cap X_{\lambda_2}$, we may assume $X_{\lambda_1} \subseteq X_{\lambda_2}$, and then $f_{\lambda_2}(x) = f_{\lambda_1}(x)$, as $f_{\lambda_2}\Big|_{X_{\lambda_1}} = f_{\lambda_1}$. We now check that f is a bijection. For injectivity, let $f(x_1) = f(x_2)$ and pick λ_1, λ_2 with $x_i \in X_{\lambda_i}$. Assume $X_{\lambda_1} \subseteq X_{\lambda_2}$. Then $x_1, x_2 \in X_{\lambda_2}$ and

$$f_{\lambda_2}(x_1) = f(x_1) = f(x_2) = f_{\lambda_2}(x_2),$$

so $x_1 = x_2$ because f_{λ_2} is injective. For surjectivity, let $y \in Y$. Then $y \in Y_\lambda$ for some λ. Since $f_\lambda : X_\lambda \to Y_\lambda$ is onto, there exists $x \in X_\lambda$ with $f_\lambda(x) = y$, hence $f(x) = y$. Therefore, (X, Y, f) is an upper bound for $\{(X_\lambda, Y_\lambda, f_\lambda)\}_{\lambda \in \Lambda}$.

By Zorn's lemma, there is a maximal element $(X_0, Y_0, f_0) \in \mathcal{S}$. If $X_0 = A$, then $f_0 : A \to Y_0 \subseteq B$ is injective, so $|A| \leq |B|$. If $Y_0 = B$, then $f_0^{-1} : B \to X_0 \subseteq A$ is injective, so $|B| \leq |A|$. Assume $X_0 \neq A$ and $Y_0 \neq B$, and let $a \in A \setminus X_0$ and $b \in B \setminus Y_0$. Define $g : X_0 \cup \{a\} \to Y_0 \cup \{b\}$ by $g(x) = f_0(x)$ for $x \in X_0$ and $g(a) = b$. Then g is a bijection and $(X_0 \cup \{a\}, Y_0 \cup \{b\}, g) > (X_0, Y_0, f_0)$, contradicting the maximality of (X_0, Y_0, f_0). $\square$

We know $|\mathbb{N}| < |\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$. The conjecture that there is no X such that $|\mathbb{N}| < |X| < |\mathbb{R}|$ is called the **continuum hypothesis**. Paul Cohen (1963) famously proved that the continuum hypothesis is logically independent of ZFC.

Lecture 7 — Monday, September 22, 2014

Cardinal Arithmetic

Cardinal numbers are a generalization of natural numbers. For every cardinal κ, there exists a set B such that $|B| = \kappa$. Similarly, for any set A, there exists a cardinal number α so that $|A| = \alpha$. In particular, $0 = |\emptyset|$, $n = |\{1, 2, \ldots, n\}|$ for all $n \geq 1$, $\aleph_0 = |\mathbb{N}|$, and $\mathfrak{c} = |\mathbb{R}|$.

Definition 7.1 Let α, β be cardinal numbers.

1. The sum $\alpha + \beta$ is defined as $|A \cup B|$ where A and B are disjoint with $|A| = \alpha, |B| = \beta$.
2. The product $\alpha \cdot \beta$ is defined as $|A \times B|$ where A, B are sets with $|A| = \alpha, |B| = \beta$.
3. The power β^α is defined as $|B^A|$ where A, B are sets $|A| = \alpha, |B| = \beta$.

Remark 7.2 Let X be a set. Then there is a bijection $\Theta : \mathcal{P}(X) \to \{0,1\}^X$ by $A \mapsto \chi_A$ where

$$\chi_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

The inverse sends a function $\chi : X \to \{0,1\}$ to the set $\{x \in X \mid \chi(x) = 1\}$, so Θ is a bijection. It follows that

$$|\mathcal{P}(X)| = |\{0,1\}^X| = 2^{|X|}.$$

Recall that if $\alpha = |A|$ and $\beta = |B|$ are cardinal numbers, we write $\alpha \le \beta$ if there exists an injection $f : A \to B$.

Lemma 7.3 Let α, β, γ be cardinal numbers.

1. The sum $\alpha + \beta$ is well-defined; that is, if A, B, C, D are sets with $|A| = |C| = \alpha$ and $|B| = |D| = \beta$ and $A \cap B = \emptyset = C \cap D$, then $|A \cup B| = |C \cup D|$.
2. The product is well-defined; that is, if A, B, C, D are sets with $|A| = |C| = \alpha$ and $|B| = |D| = \beta$ then $|A \times B| = |C \times D|$.
3. $\alpha + \beta = \beta + \alpha$.
4. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
5. $\alpha \cdot \beta = \beta \cdot \alpha$ and $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$.
6. $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$.
7. $0 + \alpha = \alpha$.
8. If $\alpha \le \beta$, then $\alpha + \gamma \le \beta + \gamma$ and $\alpha \cdot \gamma \le \beta \cdot \gamma$.

Proof. Choose sets A, B, C with $|A| = \alpha$, $|B| = \beta$, and $|C| = \gamma$.

For (1), suppose A, B, C, D are as in the statement, with bijections $f : A \to C$ and $g : B \to D$. Define

$$h : A \cup B \to C \cup D, \qquad h(x) = \begin{cases} f(x), & x \in A, \\ g(x), & x \in B. \end{cases}$$

Because $A \cap B = \emptyset$ and $C \cap D = \emptyset$, this map is well-defined and bijective. Hence $|A \cup B| = |C \cup D|$.

For (2), with bijections $f : A \to C$ and $g : B \to D$, define

$$\Phi : A \times B \to C \times D, \qquad \Phi(a, b) = (f(a), g(b)).$$

Its inverse is $(c, d) \mapsto (f^{-1}(c), g^{-1}(d))$, so Φ is a bijection.

For (3), we have $A \cup B = B \cup A$, so $\alpha + \beta = \beta + \alpha$.

For (4), choose representatives A, B, C pairwise disjoint. Then

$$(\alpha + \beta) + \gamma = |(A \cup B) \cup C| = |A \cup (B \cup C)| = \alpha + (\beta + \gamma).$$

For (5), commutativity follows from

$$\sigma : A \times B \to B \times A, \qquad \sigma(a, b) = (b, a),$$

which is a bijection with inverse itself. Associativity follows from

$$\tau : (A \times B) \times C \to A \times (B \times C), \qquad \tau((a, b), c) = (a, (b, c)),$$

which is a bijection with inverse $(a, (b, c)) \mapsto ((a, b), c)$.

For (6), choose $B \cap C = \emptyset$. Define

$$\Delta : A \times (B \cup C) \to (A \times B) \cup (A \times C), \qquad \Delta(a, x) = \begin{cases} (a, x), & x \in B, \\ (a, x), & x \in C. \end{cases}$$

This is a bijection because B and C are disjoint, so each $x \in B \cup C$ lies in exactly one branch.

For (7), if $|A| = \alpha$, then

$$0 + \alpha = |\emptyset \cup A| = |A| = \alpha.$$

For (8), assume $\alpha \leq \beta$, and fix an injection $i : A \to B$. Choose G with $|G| = \gamma$ and $A \cap G =$

$B \cap G = \emptyset$. Define

$$j : A \cup G \to B \cup G, \qquad j(x) = \begin{cases} i(x), & x \in A, \\ x, & x \in G. \end{cases}$$

Then j is injective, so $\alpha + \gamma \leq \beta + \gamma$. Also define

$$k : A \times G \to B \times G, \qquad k(a, g) = (i(a), g).$$

If $k(a, g) = k(a', g')$, then $i(a) = i(a')$ and $g = g'$, hence $a = a'$ and $g = g'$. So k is injective, and $\alpha \cdot \gamma \leq \beta \cdot \gamma$. □

Lemma 7.4 Let α, β, γ be cardinal numbers.

1. If $\alpha \leq \beta$, then $\alpha^\gamma \leq \beta^\gamma$.
2. $\alpha^{(\beta+\gamma)} = \alpha^\beta \cdot \alpha^\gamma$.
3. $(\alpha^\beta)^\gamma = \alpha^{\beta \cdot \gamma}$.

Proof. Choose sets A, B, C with $|A| = \alpha$, $|B| = \beta$, and $|C| = \gamma$.

For (1), if $\alpha \leq \beta$, choose an injection $i : A \to B$. Then

$$A^C \to B^C, \qquad f \mapsto i \circ f$$

is injective, so $\alpha^\gamma \leq \beta^\gamma$.

For (2), choose representatives for $\beta + \gamma$ as a disjoint union $B \sqcup C$. Restriction gives a bijection

$$A^{B \sqcup C} \to A^B \times A^C, \qquad f \mapsto (f\big|_B, f\big|_C),$$

with inverse obtained by combining a function on B and a function on C.

For (3), define $\Theta : A^{B \times C} \to (A^B)^C$ by sending f to the function $\Theta(f) : C \to A^B$ given by

$$[\Theta(f)](c)(b) = f(b, c).$$

The inverse sends $\Phi : C \to A^B$ to the function $B \times C \to A$ given by $(b, c) \mapsto \Phi(c)(b)$. Hence $(\alpha^\beta)^\gamma = \alpha^{\beta \cdot \gamma}$. □

Lemma 7.5 Let β be an infinite cardinal number and n be a finite cardinal number. Then $n + \beta = \beta$.

Proof. First suppose $\beta = \aleph_0 = |\mathbb{N}|$. Let $A = \{a_1, \ldots, a_n\}$ be disjoint from $\mathbb{N}$ and put $B = \mathbb{N} \cup A$. The map $\theta : \mathbb{N} \to B$ defined by

$$m \mapsto \begin{cases} a_{m+1} & \text{if } 0 \leq m \leq n-1 \\ m - n & \text{if } m \geq n \end{cases}$$

is a bijection. Thus $|A \cup \mathbb{N}| = |\mathbb{N}|$. That is, $n + \aleph_0 = \aleph_0$.

More generally, let B be a set with $|B| = \beta$. Since B is infinite, it contains a denumerable subset D. Let $A = \{a_1, \ldots, a_n\}$ be a set with $|A| = n$ and $A \cap B = \emptyset$. Then

$$n + \beta = |A \cup B| = |A \cup D \cup (B \setminus D)| = |A \cup D| + |B \setminus D| = |D| + |B \setminus D|$$

by the first case, and hence $n + \beta = |B| = \beta$. $\square$

Theorem 7.6 Let α, β be cardinal numbers, and assume that $\alpha \leq \beta$ and β is infinite. Then $\alpha + \beta = \beta$.

Proof. Suppose we can prove that $\beta + \beta = \beta$. Then

$$\beta \leq \alpha + \beta \leq \beta + \beta = \beta.$$

Then by the Schroeder–Bernstein theorem, $\alpha + \beta = \beta$.

Thus, it remains to show that $\beta + \beta = \beta$. Consider a set B with $|B| = \beta$. Then

$$B \times \{0, 1\} = \{(b, 0) \mid b \in B\} \cup \{(b, 1) \mid b \in B\}$$

so $|B \times \{0, 1\}| = |B| + |B|$, and it suffices to show that $|B \times \{0, 1\}| = |B|$.

Lecture 8 — Wednesday, September 24, 2014

Proof continued. Last time we reduced the problem to showing that $\beta + \beta = \beta$; that is, if B is a set with $|B| = \beta$, then $|B| = |B \times \{0,1\}|$. To prove this, consider

$$\mathcal{F} = \{(X, f) \mid X \subseteq B \text{ and } f : X \to X \times \{0,1\} \text{ is a bijection}\}$$

partially ordered by $(X_1, f_1) \leq (X_2, f_2)$ if $X_1 \subseteq X_2$ and $f_2\Big|_{X_1} = f_1$. We want to apply Zorn's lemma, so we need to ensure that $\mathcal{F} \neq \emptyset$. But B is infinite, and so it contains a denumerable set D. By Proposition 5.6, $D \times \{0,1\}$ is also denumerable and so there exists a bijection $f : D \to D \times \{0,1\}$. Thus $(D, f) \in \mathcal{F}$, whence $\mathcal{F} \neq \emptyset$.

Let $\mathcal{C} = \{(X_\lambda, f_\lambda) \mid \lambda \in \Lambda\}$ be a chain in $\mathcal{F}$. Consider $X = \bigcup_{\lambda \in \Lambda} X_\lambda$ and define $f : X \to X \times \{0,1\}$ by $x \mapsto f_\lambda(x)$ if $x \in X_\lambda$. Note that f is well-defined, because if $x \in X_{\lambda_1} \cap X_{\lambda_2}$, say, with $(X_{\lambda_1}, f_{\lambda_1}) \leq (X_{\lambda_2}, f_{\lambda_2})$, then $f_{\lambda_2}(x) = f_{\lambda_1}(x)$ as $f_{\lambda_2}\Big|_{X_{\lambda_1}} = f_{\lambda_1}$. We claim f is a bijection. For injectivity, let $f(x_1) = f(x_2)$ and choose λ_1, λ_2 with $x_i \in X_{\lambda_i}$. Assume $X_{\lambda_1} \subseteq X_{\lambda_2}$. Then $x_1, x_2 \in X_{\lambda_2}$ and

$$f_{\lambda_2}(x_1) = f(x_1) = f(x_2) = f_{\lambda_2}(x_2),$$

so $x_1 = x_2$ because f_{λ_2} is injective. For surjectivity, let $(u, i) \in X \times \{0,1\}$. Choose λ with $u \in X_\lambda$. Since f_λ is onto $X_\lambda \times \{0,1\}$, there exists $x \in X_\lambda$ with $f_\lambda(x) = (u, i)$, and hence $f(x) = (u, i)$. Therefore f is bijective. $(X, f) \geq (X_\lambda, f_\lambda)\ \forall \lambda \in \Lambda$, so that (X, f) is an upper bound for $\mathcal{C}$.

By Zorn's lemma, $\mathcal{F}$ admits a maximal element, say (Y, g). We claim that $B \setminus Y$ is finite. To verify the claim, suppose otherwise; then $B \setminus Y$ is infinite, and so contains some denumerable set E by Proposition 4.10. As above, there exists a bijection $h : E \to E \times \{0,1\}$. Then $\hat{g} : Y \cup E \to (Y \cup E) \times \{0,1\}$ defined by

$$\hat{g}(x) = \begin{cases} g(x), & x \in Y \\ h(x), & x \in E \end{cases}$$

is well-defined because $E \subseteq B \setminus Y$, so $Y \cap E = \emptyset$. Also,

$$\hat{g}(Y) = g(Y) = Y \times \{0,1\} \qquad \text{and} \qquad \hat{g}(E) = h(E) = E \times \{0,1\},$$

and these two sets are disjoint. Since g and h are bijections, $\hat{g}$ is injective and

$$\hat{g}(Y \cup E) = (Y \times \{0,1\}) \cup (E \times \{0,1\}) = (Y \cup E) \times \{0,1\},$$

so $\hat{g}$ is surjective. Therefore $\hat{g}$ is a bijection, and so $(Y \cup E, \hat{g}) > (Y, g)$ as $\hat{g}\Big|_Y = g$ and $Y \subsetneq Y \cup E$. This contradicts the maximality of (Y, g), and thus $|B \backslash Y| < \infty$. Since Y contains the denumerable set used to start the construction, Y is infinite, and Lemma 7.5 gives

$$|B| = |Y| + |B \setminus Y| = |Y|.$$

The bijection $g : Y \to Y \times \{0, 1\}$ now yields

$$|B| = |Y| = |Y \times \{0, 1\}| = |Y| + |Y| = |B| + |B|,$$

as required. □

Theorem 8.1 Let α and β be cardinal numbers with $1 \leq \alpha \leq \beta$. If β is infinite, then $\alpha \cdot \beta = \beta$.

Proof. Note that

$$\beta = 1 \cdot \beta \leq \alpha \cdot \beta \leq \beta \cdot \beta,$$

so if we can show $\beta \cdot \beta = \beta$, then by the Schroeder–Bernstein theorem, it follows that $\alpha \cdot \beta = \beta$. Let B be a set with $|B| = \beta$. We must prove that $|B| = |B \times B|$. Let

$$\mathcal{F} = \{(X, f) \mid X \subseteq B, f : X \to X \times X \text{ is a bijection}\}$$

partially ordered by $(X_1, f_1) \leq (X_2, f_2)$ if $X_1 \subseteq X_2$ and $f_2\Big|_{X_1} = f_1$. Since B is infinite, it contains a denumerable subset $D = \{d_n\}_{n=1}^{\infty}$, and since $D \times D = \bigcup_{n=1}^{\infty} D \times \{d_n\}$ is a countable union of denumerable sets, $D \times D$ is denumerable as well, and so a bijection $f : D \to D \times D$ exists. Thus $(D, f) \in \mathcal{F} \neq \emptyset$.

As before, we can apply Zorn's lemma to show that $\mathcal{F}$ admits a maximal element, say (Y, g). We claim that $|Y| = \beta$. Suppose instead that $|Y| < \beta$, and put $\gamma = |B \setminus Y|$. Then $\beta = |Y| + \gamma$. Cardinal comparability gives either $\gamma \leq |Y|$ or $|Y| \leq \gamma$. The first possibility would give $\beta = |Y|$ by Theorem 7.6, contrary to our assumption. Hence $|Y| \leq \gamma$, and another application of Theorem 7.6 gives $\beta = \gamma$. Thus there exists a subset $Z \subseteq B \setminus Y$ with $|Z| = |Y|$. Because g already bijects Y with $Y \times Y$, the sets $Y \times Y$, $Y \times Z$, $Z \times Y$, and $Z \times Z$ all have cardinality $|Y|$. Therefore

$$|(Y \cup Z) \times (Y \cup Z)| = |Y \times Y| + |Y \times Z| + |Z \times Y| + |Z \times Z| = |Y| + |Y| + |Y| + |Y| = |Y| + |Y| = |Y \cup Z|.$$

Write

$$(Y \cup Z) \times (Y \cup Z) = (Y \times Y) \sqcup (Y \times Z) \sqcup (Z \times Y) \sqcup (Z \times Z).$$

The map g already bijects Y with $Y \times Y$. By the cardinal identity above, the remaining part

$$(Y \times Z) \sqcup (Z \times Y) \sqcup (Z \times Z)$$

has cardinality $|Y| = |Z|$, so there exists a bijection

$$h : Z \to (Y \times Z) \sqcup (Z \times Y) \sqcup (Z \times Z).$$

Define

$$\hat{g} := g \cup h : Y \cup Z \to (Y \cup Z) \times (Y \cup Z).$$

Then $\hat{g}$ is a bijection extending g. Hence $(Y \cup Z, \hat{g}) \in \mathcal{F}$ and $(Y, g) < (Y \cup Z, \hat{g})$, contradicting maximality. Therefore $|Y| = \beta$, and then $|B \times B| = |B|$, so $\beta \cdot \beta = \beta$, and finally $\alpha \cdot \beta = \beta$ by the inequalities at the start and the Schroeder–Bernstein theorem. □

Example 8.2 Consider

$$\mathbb{R}^{\mathbb{N}} = \{(x_n)_{n=0}^{\infty} \mid x_n \in \mathbb{R}, n \geq 0\}.$$

Then

$$|\mathbb{R}^{\mathbb{N}}| = |\mathbb{R}|^{|\mathbb{N}|} = \mathfrak{c}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0} = \mathfrak{c}.$$

2. Metric Spaces

Definitions and Examples

Definition 8.3 A **metric space** is a nonempty set X together with a function $d : X \times X \to \mathbb{R}$ satisfying the following properties:

1. $d(x, y) \geq 0$ (**nonnegativity**) and $d(x, y) = 0$ if and only if $x = y$.
2. $d(x, y) = d(y, x)$ (**symmetry**).
3. $d(x, y) \leq d(x, z) + d(z, y)$ (the **triangle inequality**).

Such a function is called a **metric** on X.

Example 8.4

1. Let $X = \mathbb{R}$ and $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto |x - y|$, the **usual** (**standard**) **metric** on $\mathbb{R}$. This is the motivating example for the definition of a metric space.
2. Let $X \neq \emptyset$ be any set. Define $d : X \times X \to \mathbb{R}$ by

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

   This is the **discrete metric** on X. The triangle inequality is immediate: if $x = y$, then $d(x, y) = 0$, while if $x \neq y$, then at least one of $x \neq z$ and $z \neq y$ holds, so $d(x, z) + d(z, y) \geq 1 = d(x, y)$.
3. Let p be a prime number. Given $q \in \mathbb{Q}$, we can write $q = p^n \cdot \frac{a}{b}$ where $\gcd(a, p) = \gcd(b, p) = 1$ and $n \in \mathbb{Z}$. We define $|q|_p = p^{-n}$ and $|0|_p = 0$. The standard p-adic valuation identities show that $d_p(r, q) = |r - q|_p$ is a metric on $\mathbb{Q}$, called the ***p*-adic metric**.

Lecture 9 — Friday, September 26, 2014

In the sequel, we will write $\mathbb{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$. Many important examples of metric spaces come from the following construction.

Definition 9.1 Let X be a $\mathbb{K}$-vector space. A **norm** on X is a function $\|\cdot\| : X \to \mathbb{R}$ satisfying:

1. $\|x\| \geq 0$ for all $x \in X$ and $\|x\| = 0$ if and only if $x = 0$.
2. $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{K}$ and $x \in X$.
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

The pair $(X, \|\cdot\|)$ is called a **normed linear space**.

A norm $\|\cdot\|$ on a vector space X induces a metric on X defined by

$$d : X \times X \to \mathbb{R}, \qquad (x, y) \mapsto \|x - y\|.$$

Unless otherwise specified, we will always consider this metric when dealing with a normed linear space. Note that the Euclidean metric on $\mathbb{R}^n$ is an example of this construction.

Remark 9.2 The p-adic metric on $\mathbb{Q}$ does not come from a norm in the sense of the last remark, even though the definition of d_p looks somewhat similar to the above construction. Firstly, $\mathbb{Q}$ is not a vector space over $\mathbb{R}$ or $\mathbb{C}$. Moreover, $|\cdot|_p$ is not homogeneous; for example, $|pq|_p = p^{-1} = |p|_p$ for any prime $q \neq p$.

Example 9.3 On $\mathbb{K}^n$ we define three norms as follows: given $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$, let

1. $$\|x\|_1 = \sum_{k=1}^{n} |x_k|.$$

2. $$\|x\|_\infty = \max_{k=1,\ldots,n} |x_k|.$$

3. $$\|x\|_2 = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{1/2}.$$

These are referred to as the **1-norm**, the **infinity norm**, and the **2-norm** (or the **Euclidean norm**), respectively. The norm axioms for $\|\cdot\|_1$ and $\|\cdot\|_\infty$ follow directly from their definitions. Positive definiteness and homogeneity for $\|\cdot\|_2$ do as well; we will prove its triangle inequality later.

Although these norms induce the same topology on $\mathbb{R}^2$, their unit balls have the visibly different shapes in Figure 3.

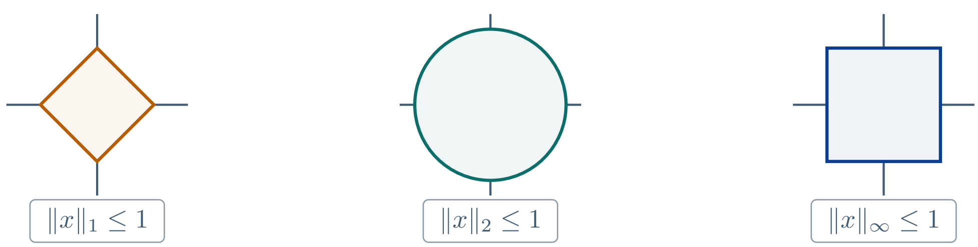


FIGURE 3

Example 9.4 Let
$$\ell_1 = \left\{ x = (x_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \mid \sum_{n=0}^{\infty} |x_n| < \infty \right\}.$$
Then ℓ_1 is a vector space over $\mathbb{K}$ and
$$\|x\|_1 = \sum_{n=0}^{\infty} |x_n|$$
defines a norm on ℓ_1, once again called the 1-norm on ℓ_1.

Let
$$\ell_\infty = \left\{ x = (x_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} \mid \sup_{n\in\mathbb{N}} |x_n| < \infty \right\}.$$
Then ℓ_∞ is a vector space over $\mathbb{K}$ and
$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$$
defines a norm on ℓ_∞, also called the **infinity-norm** or **supremum norm** on ℓ_∞.

Example 9.5 Let $C([0,1],\mathbb{K}) = \{f : [0,1] \to \mathbb{K} \mid f \text{ is continuous}\}$. We define a norm on $C([0,1],\mathbb{K})$ by
$$\|f\|_\infty = \sup_{x\in[0,1]} |f(x)| = \max_{x\in[0,1]} |f(x)|.$$
This norm is also called the **supremum norm**. Moreover,
$$\|f\|_1 = \int_0^1 |f(x)|\,\mathrm{d}x$$
also defines a norm on $C([0,1],\mathbb{K})$, which is also called the **1-norm**. Indeed, nonnegativity is clear. If $\|f\|_1 = 0$, then $|f|$ is continuous, nonnegative, and has integral 0, so $|f(x)| = 0$ for all $x \in [0,1]$, hence $f = 0$. Homogeneity follows from
$$\|\alpha f\|_1 = \int_0^1 |\alpha f(x)|\,\mathrm{d}x = |\alpha| \int_0^1 |f(x)|\,\mathrm{d}x = |\alpha|\,\|f\|_1,$$
and the triangle inequality follows from $|f(x) + g(x)| \le |f(x)| + |g(x)|$:
$$\|f + g\|_1 \le \int_0^1 (|f(x)| + |g(x)|)\,\mathrm{d}x = \|f\|_1 + \|g\|_1.$$

Example 9.6 Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces over the same field $\mathbb{K}$. Recall that

$$\mathrm{Hom}(X, Y) = \{T : X \to Y \mid T \text{ is linear}\}$$

is again a vector space over $\mathbb{K}$. For $T \in \mathrm{Hom}(X, Y)$, we define

$$\|T\| = \sup\{\|Tx\|_Y \mid x \in X \text{ with } \|x\|_X \leq 1\},$$

and say that T is bounded if $\|T\| < \infty$. Then

$$B(X, Y) = \{T \in \mathrm{Hom}(X, Y) \mid \|T\| < \infty\}$$

is a subspace of $\mathrm{Hom}(X, Y)$, and $\|\cdot\|$ is a norm on $B(X, Y)$. To see this, let $S, T \in B(X, Y)$ and $\alpha, \beta \in \mathbb{K}$. For $\|x\|_X \leq 1$,

$$\|(\alpha S + \beta T)x\|_Y \leq |\alpha|\,\|Sx\|_Y + |\beta|\,\|Tx\|_Y \leq |\alpha|\,\|S\| + |\beta|\,\|T\|,$$

so $\alpha S + \beta T \in B(X, Y)$. Thus $B(X, Y)$ is a subspace. Also, for $T \in B(X, Y)$, $\|T\| \geq 0$ and

$$\|\alpha T\| = |\alpha|\,\|T\| \qquad \text{and} \qquad \|S + T\| \leq \|S\| + \|T\|.$$

If $\|T\| = 0$, then $\|Tx\|_Y = 0$ whenever $\|x\|_X \leq 1$. For $x \neq 0$, apply this to $x/\|x\|_X$ and use linearity to get $Tx = 0$. Hence $T = 0$.

Lecture 10 — Monday, September 29, 2014

Definition 10.1 For $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ and $1 \leq p < \infty$, we define ***p*-norm**

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

This generalizes $\|\cdot\|_1$ and $\|\cdot\|_2$. We will show that $\|\cdot\|_p$ is a norm on $\mathbb{K}^n$. For $p = 1$, the claim follows directly from the definition. For general p, only the triangle inequality requires work.

Lemma 10.2 (Young's inequality) Let $p, q > 1$ satisfy

$$\frac{1}{p} + \frac{1}{q} = 1.$$

If $a, b \geq 0$, then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof. If $a = 0$ or $b = 0$, the inequality is immediate. Assume $a, b > 0$. Let $t = 1/p \in (0, 1)$ so $1 - t = 1/q$. Since log is concave,

$$\log(ta^p + (1-t)b^q) \geq t\log(a^p) + (1-t)\log(b^q) = \log(a) + \log(b) = \log(ab).$$

Applying the exponential function to both sides gives

$$ta^p + (1-t)b^q \geq ab,$$

and therefore

$$\frac{a^p}{p} + \frac{b^q}{q} \geq ab.$$

□

Theorem 10.3 (Hölder's inequality) Let $1 < p, q < \infty$ be reals such that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then

$$\|xy\|_1 \leq \|x\|_p \|y\|_q$$

for all $x, y \in \mathbb{K}^n$.

Proof. If $x = 0$ or $y = 0$, the result is immediate, so assume $x, y \neq 0$. By homogeneity, it suffices to prove the result under the normalization $\|x\|_p = 1$ and $\|y\|_q = 1$. Write $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Then by Young's inequality,

$$\|xy\|_1 = \sum_{i=1}^n |x_i y_i| \leq \sum_{i=1}^n \left(\frac{|x_i|^p}{p} + \frac{|y_i|^q}{q}\right) = \frac{1}{p}\|x\|_p^p + \frac{1}{q}\|y\|_q^q$$

$$= \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q.$$

□

Theorem 10.4 (Minkowski's inequality) Let $1 \leq p < \infty$. Then

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

for all $x, y \in \mathbb{K}^n$.

Proof. If $p = 1$, the result follows immediately from the scalar triangle inequality:

$$\|x + y\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1.$$

Assume henceforth that $p > 1$. Let q be such that $1/p + 1/q = 1$, and write $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Then

$$\begin{aligned}
\|x + y\|_p^p &= \sum_{i=1}^{n} |x_i + y_i|^p \\
&= \sum_{i=1}^{n} |x_i + y_i| \, |x_i + y_i|^{p-1} \\
&\leq \sum_{i=1}^{n} |x_i| \, |x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i| \, |x_i + y_i|^{p-1}.
\end{aligned}$$

By Hölder's inequality,

$$\sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} \leq \|x\|_p \left(\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q} \right)^{1/q}.$$

Note that

$$\frac{1}{p} + \frac{1}{q} = 1 \implies q = \frac{p}{p-1} \implies q(p-1) = p.$$

Also $1/q = (p-1)/p$, so

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p \right)^{1/q} = \|x + y\|_p^{p/q} = \|x + y\|_p^{p-1}.$$

Similarly,

$$\sum_{i=1}^{n} |y_i||x_i + y_i|^{p-1} \le \|y\|_p \|x + y\|_p^{p-1}.$$

Combining these estimates gives

$$\|x + y\|_p^p \le (\|x\|_p + \|y\|_p)\|x + y\|_p^{p-1}.$$

If $x + y = 0$, the result is immediate; otherwise, divide by $\|x + y\|_p^{p-1}$ to obtain the result. $\square$

Topology of Metric Spaces

Definition 10.5 Let (X, d) be a metric space, let $x \in X$ and $r > 0$. The **open ball of radius r around x** is $B_r(x) = \{y \in X \mid d(x, y) < r\}$.

1. A set $N \subseteq X$ is said to be a **neighborhood** of x if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq N$.
2. A set $U \subseteq X$ is said to be **open** if for every $x \in U$, there exists $\varepsilon > 0$ (depending on x) such that $B_\varepsilon(x) \subseteq U$.
3. A set $F \subseteq X$ is said to be **closed** if $X \setminus F$ is open.

By our convention, neighborhoods are not assumed to be open. Also observe that a subset $U \subseteq X$ is open if and only if it is a neighborhood of each of its points.

Proposition 10.6 Let (X, d) be a metric space, $x \in X, r > 0$. The open ball $B_r(x)$ is open.

Proof. Let $y \in B_r(x)$. We want to show that there exists $\varepsilon > 0$ such that $B_\varepsilon(y) \subseteq B_r(x)$. Since $y \in B_r(x)$, $d(y, x) < r$. Let $\varepsilon = r - d(x, y) > 0$. If $z \in B_\varepsilon(y)$, then

$$d(z, x) \le d(z, y) + d(y, x) < \varepsilon + d(x, y) = r.$$

Thus $B_\varepsilon(y) \subseteq B_r(x)$. Hence $B_r(x)$ is open. $\square$

Lecture 11 — Wednesday, October 01, 2014

Proposition 11.1 Let (X, d) be a metric space, $x \in X, r \geq 0$. Then $\overline{B}_r(x) = \{y \in X \mid d(x, y) \leq r\}$ is closed. In particular, $\{x\}$ is closed.

Proof. Let $y \notin \overline{B}_r(x)$, so $d(x, y) > r$. Set $\varepsilon := (d(x, y) - r)/2 > 0$. If $z \in B_\varepsilon(y)$, then

$$d(x, z) \geq d(x, y) - d(y, z) > d(x, y) - \varepsilon = \frac{d(x, y) + r}{2} > r,$$

so $z \notin \overline{B}_r(x)$. Hence $B_\varepsilon(y) \subseteq X \setminus \overline{B}_r(x)$, showing the complement is open. Therefore $\overline{B}_r(x)$ is closed. Taking $r = 0$ shows that $\{x\}$ is closed. □

Example 11.2 Consider $\mathbb{R}$ equipped with the Euclidean metric, and let $a, b \in \mathbb{R}$ with $a < b$.

1. (a, b) is open.
2. $[a, b]$ is closed.
3. The intervals $(-\infty, a), (b, \infty)$ are open.
4. $(-\infty, a]$ and $[b, \infty)$ are closed.
5. $[a, b), (a, b]$ are neither open nor closed.

Warning Sets are not doors: they need not be open or closed.

Theorem 11.3 Let (X, d) be a metric space.

1. $\emptyset$ and X are open.
2. If $\{U_\lambda\}_{\lambda\in\Lambda}$ is any collection of open sets, then $\bigcup_{\lambda\in\Lambda} U_\lambda$ is open.
3. If $\{V_1, \ldots, V_m\}$ is a finite collection of open sets in X, then $V = \bigcap_{i=1}^m V_i$ is open.

Proof. For (1), $\emptyset$ is open vacuously, and X is open because for every $x \in X$, $B_1(x) \subseteq X$.

For (2), let
$$U = \bigcup_{\lambda \in \Lambda} U_\lambda$$
where each U_λ is open. If $x \in U$, then $x \in U_{\lambda_0}$ for some $\lambda_0 \in \Lambda$. Since U_{λ_0} is open, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U_{\lambda_0} \subseteq U$. Hence U is open.

For (3), let
$$V = \bigcap_{i=1}^{m} V_i$$
where $V_1, \ldots, V_m$ are open. If $x \in V$, then $x \in V_i$ for each $i = 1, \ldots, m$. For each i, choose $\varepsilon_i > 0$ such that $B_{\varepsilon_i}(x) \subseteq V_i$. Set
$$\varepsilon = \min\{\varepsilon_1, \ldots, \varepsilon_m\} > 0.$$
Then $B_\varepsilon(x) \subseteq V_i$ for every i, so
$$B_\varepsilon(x) \subseteq \bigcap_{i=1}^{m} V_i = V.$$
Therefore V is open. $\square$

Example 11.4 Infinite intersections of open sets need not be open. For $n \geq 1$, let $V_n = (-1/n, 1) \subset \mathbb{R}$. Then each V_n is open, but $\bigcap_{n=1}^{\infty} V_n = [0, 1)$ which is neither open nor closed.

Example 11.5 Let $X \neq \emptyset$ and let d be the discrete metric. For $x \in X$, the singleton $\{x\} = B_{\frac{1}{2}}(x)$ is open. If $A \subset X$, then $A = \bigcup_{x \in A}\{x\}$ is therefore open. Since every subset of (X, d) is open, every subset is also closed.

Theorem 11.6 Let (X, d) be a metric space.

1. $\emptyset$ and X are closed.
2. If $\{F_\lambda\}_{\lambda \in \Lambda}$ is a collection of closed sets, then $F = \bigcap_{\lambda \in \Lambda} F_\lambda$ is closed.
3. If $\{G_1, \ldots, G_n\}$ is a finite collection of closed sets, then $G = \bigcup_{i=1}^{n} G_i$ is closed.

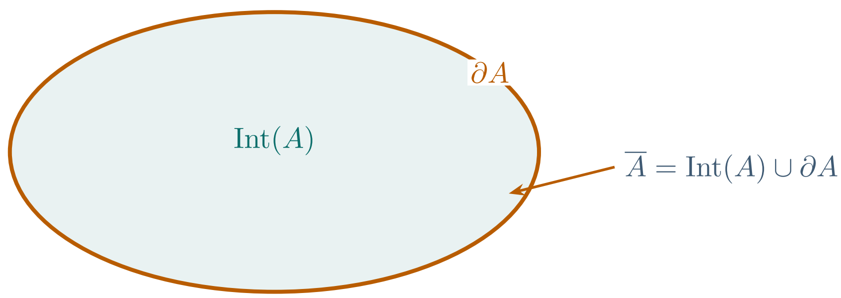
Proof. Take complements in Theorem 11.3. $\square$

Definition 11.7 Let (X, d) be a metric space and let $A \subset X$.

1. The **closure** of A is defined to be the set $\overline{A} = \bigcap\{F \subset X \mid A \subset F, F \text{ is closed}\}$.
2. The **interior** of A is defined to be $\mathrm{Int}(A) = \bigcup\{U \subset X \mid U \subset A, U \text{ is open}\}$.
3. The **boundary** of A is $\partial A := \overline{A} \setminus \mathrm{Int}(A)$.
4. An element $x \in X$ is said to be an **accumulation point** of A if $(B_\varepsilon(x) \setminus \{x\}) \cap A \neq \emptyset$. We write A' for the set of all accumulation points of A.

$\overline{A}$ is closed by Theorem 11.6, and is the smallest closed set that contains A (in the sense that if $F \subset X$ is closed and $A \subset F$, then $\overline{A} \subset F$). Similarly, $\mathrm{Int}(A)$ is the largest open set contained in A.

The three regions are separated in Figure 4: the interior excludes the edge, the boundary is the edge, and the closure contains both.



closure adds every boundary point; interior keeps only points with room around them

FIGURE 4

Proposition 11.8 Let (X, d) be a metric space and let $A \subset X$.

1. A is closed if and only if $A = \overline{A}$.
2. A is open if and only if $A = \mathrm{Int}(A)$.
3. $X \setminus \overline{A} = \mathrm{Int}(X \setminus A)$ and $X \setminus \mathrm{Int}(A) = \overline{X \setminus A}$.
4. ∂A is closed and $\overline{A} = A \cup \partial A$.

Proof. For (1), if $A = \overline{A}$, then A is closed because $\overline{A}$ is closed. Conversely, if A is closed, then

$A = \overline{A}$ because $\overline{A}$ is the smallest closed set containing A.

For (2), the argument is analogous: A is open if and only if $A = \operatorname{Int}(A)$, since $\operatorname{Int}(A)$ is the largest open subset of A.

For (3), De Morgan's laws give

$$X \setminus \overline{A} = \bigcup \{X \setminus F \mid F \subset X \text{ is closed}, A \subset F\}.$$

Observe that F is closed and $A \subset F$ if and only if $X \setminus F$ is open and $X \setminus F \subset X \setminus A$. Thus $X \setminus \overline{A} = \bigcup \{U \mid U \subset X \text{ is open}, U \subset X \setminus A\} = \operatorname{Int}(X \setminus A)$. Take complements and get $X \setminus \operatorname{Int}(A) = \overline{X \setminus A}$.

For (4),

$$\partial A = \overline{A} \cap (X \setminus \operatorname{Int}(A))$$

is an intersection of closed sets, hence closed. Also, since $\operatorname{Int}(A) \subset A$, we have

$$A \cup \partial A = A \cup (\overline{A} \setminus \operatorname{Int}(A)) = \overline{A}.$$

□

Lecture 12 — Friday, October 03, 2014

Proposition 12.1 Let (X, d) be a metric space, $A \subset X$, and $x \in X$.

1. $x \in \overline{A}$ if and only if $B_\varepsilon(x) \cap A \neq \emptyset$ for all $\varepsilon > 0$.
2. $x \in \operatorname{Int}(A)$ if and only if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset A$.
3. $x \in \partial A$ if and only if $B_\varepsilon(x) \cap A \neq \emptyset$ and $B_\varepsilon(x) \cap (X \setminus A) \neq \emptyset$ for all $\varepsilon > 0$.

Proof. For (1), define

$$B := \{y \in X \mid B_\varepsilon(y) \cap A \neq \emptyset \text{ for all } \varepsilon > 0\}.$$

We claim that $\overline{A} = B$. First, let $x \in X \setminus \overline{A}$. Since $X \setminus \overline{A}$ is open, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset X \setminus \overline{A}$. In particular, $B_\varepsilon(x) \cap A = \emptyset$, so $x \notin B$. Hence $B \subset \overline{A}$.

Now we prove the reverse inclusion. We first show that B is closed. Let $x \in X \setminus B$. Then there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \cap A = \emptyset$. We claim that $B_\varepsilon(x) \cap B = \emptyset$. Suppose not; pick $y \in B_\varepsilon(x) \cap B$. Set $\delta := \varepsilon - d(x,y) > 0$. Since $y \in B$, there exists $z \in A$ such that $d(y,z) < \delta$. Then

$$d(x,z) \leq d(x,y) + d(y,z) < d(x,y) + \varepsilon - d(x,y) = \varepsilon,$$

so $z \in B_\varepsilon(x) \cap A$, contradicting $B_\varepsilon(x) \cap A = \emptyset$. Therefore $B_\varepsilon(x) \subset X \setminus B$, so $X \setminus B$ is open and B is closed.

Also, $A \subset B$, because if $a \in A$ and $\varepsilon > 0$, then $a \in B_\varepsilon(a) \cap A \neq \emptyset$. Since B is a closed set containing A, the minimality of $\overline{A}$ yields $\overline{A} \subset B$. Combining with $B \subset \overline{A}$ gives $\overline{A} = B$, as required.

For (2),

$$x \in \mathrm{Int}(A) \iff x \in X \setminus \overline{X \setminus A}$$

by Proposition 11.8(3). By (1) applied to $X \setminus A$, this is equivalent to the existence of $\varepsilon > 0$ such that

$$B_\varepsilon(x) \cap (X \setminus A) = \emptyset,$$

which is equivalent to $B_\varepsilon(x) \subset A$.

For (3),

$$x \in \partial A \iff x \in \overline{A} \setminus \mathrm{Int}(A) \iff x \in \overline{A} \text{ and } x \in X \setminus \mathrm{Int}(A) \iff x \in \overline{A} \text{ and } x \in \overline{X \setminus A},$$

where the last equivalence is by Proposition 11.8(3). By (1), the condition $x \in \overline{A}$ is equivalent to

$$B_\varepsilon(x) \cap A \neq \emptyset \quad \text{for all } \varepsilon > 0.$$

Again by (1), now applied to $X \setminus A$, the condition $x \in \overline{X \setminus A}$ is equivalent to

$$B_\varepsilon(x) \cap (X \setminus A) \neq \emptyset \quad \text{for all } \varepsilon > 0.$$

Combining these two statements gives the desired characterization. □

Example 12.2 Consider $\mathbb{R}$ equipped with the Euclidean metric, and let $A = (0,1)$. Since A is open, $\mathrm{Int}(A) = A = (0,1)$. The points 0 and 1 are in $\overline{A}$; indeed, given $\varepsilon > 0$, $B_\varepsilon(0) \cap A \neq \emptyset$ and $B_\varepsilon(1) \cap A \neq \emptyset$. So $\{0,1\} \subset \overline{A}$ by Proposition 12.1(1). Hence $A = (0,1) \subset [0,1] \subset \overline{A}$. Since $[0,1]$ is closed, we have that $\overline{A} \subset [0,1]$, so $\overline{A} = [0,1]$. Finally,

$$\partial A = \overline{A} \setminus \mathrm{Int}(A) = [0,1] \setminus (0,1) = \{0,1\}.$$

For $A = \{0\}$,

$$\partial A = \overline{A} \setminus \mathrm{Int}(A) = \{0\} \setminus \emptyset = \{0\},$$

$A' = \emptyset$.

Corollary 12.3 Let (X, d) be a metric space, $A \subset X$. Then $\overline{A} = A \cup A'$.

Proof. If $x \in A'$, then $(B_\varepsilon(x) \setminus \{x\}) \cap A \neq \emptyset$ for all $\varepsilon > 0$, hence $B_\varepsilon(x) \cap A \neq \emptyset$ for all $\varepsilon > 0$, so $x \in \overline{A}$ by Proposition 12.1(1). Thus $A' \subset \overline{A}$. Since $A \subset \overline{A}$, $A \cup A' \subset \overline{A}$. Conversely, let $x \in \overline{A}$. If $x \in A$, we are done. If $x \notin A$, $B_\varepsilon(x) \cap A \neq \emptyset$ by Proposition 12.1(1); since $x \notin A$,

$$B_\varepsilon(x) \cap A = (B_\varepsilon(x) \setminus \{x\}) \cap A \neq \emptyset$$

for all $\varepsilon > 0$; that is, $x \in A'$. Therefore, $\overline{A} \subset A \cup A'$. □

Example 12.4

1. Let (X, d) be a discrete metric space, and $A \subset X$. If $\varepsilon = 1/2$, then

$$B_\varepsilon(x) \cap A = \begin{cases} \{x\}, & x \in A \\ \emptyset, & x \notin A \end{cases}.$$

   In both cases, $(B_\varepsilon(x) \setminus \{x\}) \cap A = \emptyset$. Thus $A' = \emptyset$. Moreover, $\overline{A} = A \cup A' = A$, which we already knew.

2. Consider $\mathbb{R}$ with the Euclidean metric. Let $A = \{1/n \mid n \in \mathbb{N}, n \geq 1\}$. Then 0 is an accumulation point of A. Indeed, if $\varepsilon > 0$, let $n \in \mathbb{N}$ be such that $1/n < \varepsilon$. Then $1/n \in (B_\varepsilon(0) \setminus \{0\}) \cap A$. In fact, $A' = \{0\}$, so that

$$\overline{A} = A \cup A' = \{0\} \cup \{1/n \mid n \in \mathbb{N}, n \geq 1\}.$$

Corollary 12.5 Let (X, d) be a metric space and let $A \subset X$. Then the following are equivalent:

1. A is closed.

2. $A' \subset A$.
3. $\partial A \subset A$.
4. $A = \overline{A}$.

Proof. We know that A is closed if and only if $A = \overline{A}$. Since $\overline{A} = A \cup A'$, it follows that A is closed if and only if $A' \subset A$. Similarly, $\overline{A} = A \cup \partial A$, so A is closed if and only if $\partial A \subset A$. □

Note that $\overline{A} = A \cup \partial A = A \cup A'$, but $A' \neq \partial A$ in general (for example, consider a discrete metric space).

Definition 12.6 Let (X, d) be a metric space. A subset $A \subset X$ is said to be **dense** in X if $\overline{A} = X$. We say that X is **separable** if it admits a countable dense subset.

Example 12.7

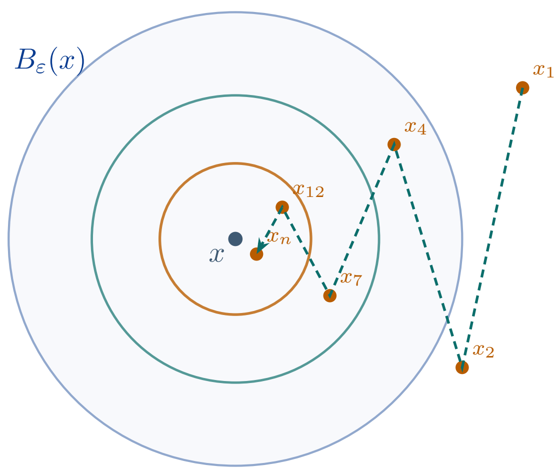
1. Every point in $\mathbb{R}$ is an accumulation point of $\mathbb{Q}$, so $\overline{\mathbb{Q}} = \mathbb{R}$. Since $\mathbb{Q}$ is countable, $\mathbb{R}$ is separable.
2. Let (X, d) be a discrete space. Since $\overline{A} = A$ for all $A \subset X$, the only dense subset of X is X itself. Then X is separable if and only if X itself is countable.
3. $(\ell_1, \|\cdot\|_1)$ is separable.
4. $(\ell_\infty, \|\cdot\|_\infty)$ is not separable.

Lecture 13 — Monday, October 06, 2014

Sequences and Topology in Metric Spaces

Definition 13.1 Let $\{x_n\}_n$ be a sequence in a metric space (X, d) and let $x \in X$. We say that $\{x_n\}_n$ **converges** to x and write $\lim_{n\to\infty} x_n = x$ if for every $\varepsilon > 0$, there exists $N > 0$ such that $d(x_n, x) < \varepsilon$ whenever $n \geq N$. The sequence $\{x_n\}_n$ is said to be a **Cauchy sequence** if for every $\varepsilon > 0$, there exists $N > 0$ such that $d(x_n, x_m) < \varepsilon$ whenever $n, m \geq N$.

Convergence means that every fixed neighborhood eventually captures the whole tail. In Figure 5, later terms lie in progressively smaller balls around x; the finitely many early terms are irrelevant.



for every $\varepsilon > 0$, all sufficiently late terms lie in $B_\varepsilon(x)$

Figure 5

Example 13.2 Let (X, d) be a discrete metric space and let $\{x_n\}_n$ be a sequence in X. If $\lim_{n\to\infty} x_n = x$, then there exists $N > 0$ such that $d(x_n, x) < 1/2$ whenever $n \geq N$. Hence $x_n = x$ for all $n \geq N$.

Conversely, if there exists $N > 0$ such that $x_n = x$ for all $n \geq N$, then $\lim_{n\to\infty} x_n = x$. Indeed, if $\varepsilon > 0$, then $d(x_n, x) = 0 < \varepsilon$ whenever $n \geq N$, so that $\lim_{n\to\infty} x_n = x$. Therefore, a sequence in (X, d) converges if and only if it is eventually constant.

Example 13.3 For $n \geq 3$, define $f_n \in C([0,1], \mathbb{R})$ by

$$f_n(x) = \begin{cases} nx, & 0 \leq x \leq 1/n \\ -nx + 2, & 1/n < x \leq 2/n \\ 0, & x > 2/n \end{cases}$$

The triangular spike in Figure 6 makes the shrinking support and fixed height explicit.

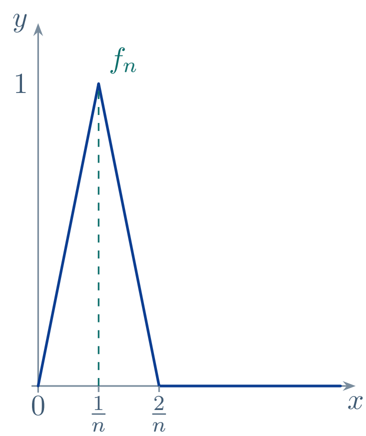


Figure 6

Then $\{f_n\}_n$ converges pointwise to 0; that is, $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in [0,1]$. Indeed, if $x = 0$, then $f_n(x) = 0$ for all n, so $\lim_{n\to\infty} f_n(x) = 0$. On the other hand, if $x > 0$, choose $N > 0$ such that $2/N < x$. Then $f_n(x) = 0$ for all $n \geq N$, hence $\lim_{n\to\infty} f_n(x) = 0$.

However, $\{f_n\}_n$ does not converge to 0 in $(C([0,1], \mathbb{R}), \|\cdot\|_\infty)$. Indeed,

$$\|f_n\|_\infty = \sup_{x\in[0,1]} |f_n(x)| = 1$$

for all $n \in \mathbb{N}$. In fact, $\{f_n\}_n$ is not even Cauchy in $(C([0,1], \mathbb{R}), \|\cdot\|_\infty)$. Indeed,

$$\|f_n - f_{2n}\|_\infty \geq \left| f_n\left(\frac{1}{n}\right) - f_{2n}\left(\frac{1}{n}\right) \right| = |1 - 0| = 1.$$

Lemma 13.4 Let (X, d) be a metric space and let $\{x_n\}_n$ be a sequence in X.

1. If $\{x_n\}_n$ converges, then the limit is unique.
2. If $\{x_n\}_n$ converges, then $\{x_n\}_n$ is a Cauchy sequence.
3. If $\{x_n\}_n$ is Cauchy and admits a subsequence $\{x_{n_k}\}_k$ which converges, with $\lim_{k\to\infty} x_{n_k} = x$, then $\{x_n\}_n$ converges to x.

Proof. For (1), suppose that x and y are different limits of $\{x_n\}_n$. Then $d(x, y) > 0$. Let $\varepsilon = \frac{d(x,y)}{2} > 0$. Since $\lim_{n\to\infty} x_n = x$, there exists $N_1 > 0$ such that $d(x_n, x) < \varepsilon$ whenever $n \geq N_1$. Since $\lim_{n\to\infty} x_n = y$, there exists $N_2 > 0$ such that $d(x_n, y) < \varepsilon$ whenever $n \geq N_2$. Let $N = \max\{N_1, N_2\}$. Then

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \varepsilon + \varepsilon = d(x, y),$$

a contradiction; hence $x = y$.

For (2), suppose $x_n \to x$. Given $\varepsilon > 0$, choose N such that $d(x_n, x) < \varepsilon/2$ for all $n \geq N$. Then for $m, n \geq N$,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon.$$

So $\{x_n\}$ is Cauchy.

For (3), suppose $\{x_n\}$ is Cauchy and $x_{n_k} \to x$. Given $\varepsilon > 0$, choose N_1 such that $d(x_n, x_m) < \varepsilon/2$ for all $m, n \geq N_1$. Also choose K such that $k \geq K$ implies $d(x_{n_k}, x) < \varepsilon/2$, and pick $k_0 \geq K$ with $n_{k_0} \geq N_1$. Then for $n \geq N_1$,

$$d(x_n, x) \leq d(x_n, x_{n_{k_0}}) + d(x_{n_{k_0}}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $x_n \to x$. □

Proposition 13.5 Let (X, d) be a metric space, $A \subset X$, $x \in X$.

1. $x \in \overline{A}$ if and only if there exists a sequence $\{x_n\}_n \subset A$ with $\lim_{n\to\infty} x_n = x$.
2. $x \in A'$ if and only if there exists a sequence $\{x_n\}_n$ in A with $\lim_{n\to\infty} x_n = x$ and $x_n \neq x$ for all $n \in \mathbb{N}$.

Proof. For (1):

Forward implication. Suppose that $\{x_n\}_n \subset A$ and $\lim_{n\to\infty} x_n = x$. Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon$ for all $n \geq N$. Hence $x_n \in B_\varepsilon(x) \cap A$ for all $n \geq N$, so $B_\varepsilon(x) \cap A \neq \emptyset$. Since $\varepsilon > 0$ was arbitrary, $x \in \overline{A}$.

Reverse implication. Suppose that $x \in \overline{A}$. Then $B_\varepsilon(x) \cap A \neq \emptyset$ for all $\varepsilon > 0$. For each $n \geq 1$, choose

$$x_n \in B_{1/n}(x) \cap A.$$

Then $\{x_n\}_n \subset A$, and $d(x_n, x) < 1/n$ for all n, so $x_n \to x$.

For (2):

Forward implication. Suppose that $x \in A'$. By definition, for every $\varepsilon > 0$,

$$(B_\varepsilon(x) \setminus \{x\}) \cap A \neq \emptyset.$$

For each $n \geq 1$, choose

$$x_n \in (B_{1/n}(x) \setminus \{x\}) \cap A.$$

Then $x_n \in A$, $x_n \neq x$ for all n, and $d(x_n, x) < 1/n$, hence $x_n \to x$.

Reverse implication. Suppose there exists a sequence $\{x_n\}_n$ in A such that $x_n \to x$ and $x_n \neq x$ for all $n \in \mathbb{N}$. Let $\varepsilon > 0$. Since $x_n \to x$, there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon$ for all $n \geq N$. In particular, $x_N \in B_\varepsilon(x) \cap A$, and because $x_N \neq x$, we get

$$x_N \in (B_\varepsilon(x) \setminus \{x\}) \cap A.$$

Thus $(B_\varepsilon(x) \setminus \{x\}) \cap A \neq \emptyset$ for every $\varepsilon > 0$, so $x \in A'$. $\square$

Remark 13.6 Proposition 13.5 *really* takes advantage of the fact that we are working in a metric space.

Corollary 13.7 Let (X, d) be a metric space and $A \subset X$. Then A is closed if and only if the limit of every convergent sequence in A is also in A.

Proof. This is immediate from Proposition 13.5. $\square$

Example 13.8 $(0,1] \subset \mathbb{R}$ is not closed, as $(1/n)_{n>0}$ is a sequence in $(0,1]$ but $\lim_{n\to\infty} 1/n = 0 \notin (0,1]$.

Definition 13.9 Let $\{x_n\}_n$ be a sequence in a metric space (X,d). An element $x \in X$ is said to be a **cluster point** of the sequence $\{x_n\}_n$ if there exists a subsequence $\{x_{n_k}\}_k$ such that $\lim_{k\to\infty} x_{n_k} = x$.

It is important to distinguish between cluster points of a sequence and accumulation points of the set $A = \{x_n \mid n \in \mathbb{N}\}$.

Proposition 13.10 Let $\{x_n\}_n$ be a sequence in a metric space (X,d) and let $A = \{x_n \mid n \in \mathbb{N}\}$. Then every accumulation point of A is a cluster point of $\{x_n\}_n$.

Proof. Let x be an accumulation point of A. Then for every $k \geq 1$ and every $N \in \mathbb{N}$, there exists $n \geq N$ such that $x_n \in B_{1/k}(x) \setminus \{x\}$. Otherwise, for some k there exists N such that $x_n \notin B_{1/k}(x) \setminus \{x\}$ for all $n \geq N$, so $(B_{1/k}(x) \setminus \{x\}) \cap A$ is finite. If its elements have positive distances $r_1, \ldots, r_j$ from x, then a ball whose radius is smaller than both $1/k$ and every r_i misses them all; this contradicts that x is an accumulation point. Thus we may choose a strictly increasing sequence $(n_k)_{k\geq 1}$ with $x_{n_k} \in B_{1/k}(x) \setminus \{x\}$, and then $x_{n_k} \to x$. So x is a cluster point. $\square$

This implication is not reversible. For example, let $x_n = 1$ for all $n \in \mathbb{N}$. Then 1 is a cluster point of $\{x_n\}_n$, but $A = \{x_n \mid n \in \mathbb{N}\} = \{1\}$, which has no accumulation points.

Lecture 14 — Wednesday, October 08, 2014

Continuity in Metric Spaces

Definition 14.1 Let (X, d_X) and (Y, d_Y) be metric spaces and let $f : X \to Y$. We say that f is **continuous at $x \in X$** if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), f(y)) < \varepsilon$ whenever $d_X(x,y) < \delta$. We say that f is **continuous** if it is continuous at each $x \in X$.

Recall that if $f : X \to Y$ is a function and $B \subset Y$, then

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$$

is called the **preimage** of B under f.

Theorem 14.2 Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. Then the following are equivalent:

1. f is continuous at x.
2. If W is a neighborhood of $f(x)$ in Y, then $f^{-1}(W)$ is a neighborhood of $x \in X$.
3. If $\{x_n\}_n$ is a sequence in X which converges to x, then $\lim_{n\to\infty} f(x_n) = f(x)$.

Proof. (1) *implies* (2). Suppose that f is continuous at x and let W be a neighborhood of $f(x)$. Then there exists $\varepsilon > 0$ such that $B_\varepsilon(f(x)) \subseteq W$. Since f is continuous at x, there exists $\delta > 0$ such that $f(y) \in B_\varepsilon(f(x))$ whenever $y \in B_\delta(x)$. Hence

$$f(B_\delta(x)) \subseteq B_\varepsilon(f(x)) \subseteq W,$$

so $B_\delta(x) \subseteq f^{-1}(W)$. Therefore $f^{-1}(W)$ is a neighborhood of x.

(2) *implies* (3). Assume (2) and let $\{x_n\}_n$ be a sequence in X with $x_n \to x$. Let W be a neighborhood of $f(x)$. By (2), $f^{-1}(W)$ is a neighborhood of x, so there exists $\delta > 0$ such that $B_\delta(x) \subseteq f^{-1}(W)$. Since $x_n \to x$, there exists $N \in \mathbb{N}$ such that $x_n \in B_\delta(x)$ for all $n \geq N$. Thus $f(x_n) \in W$ for all $n \geq N$, which proves $f(x_n) \to f(x)$.

(3) *implies* (1). Suppose f is not continuous at x. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$, there exists $y \in X$ with

$$d_X(x, y) < \delta \qquad \text{but} \qquad d_Y(f(x), f(y)) \geq \varepsilon_0.$$

For each $n \geq 1$, taking $\delta = 1/n$, choose $x_n \in X$ with

$$d_X(x, x_n) < \frac{1}{n} \qquad \text{and} \qquad d_Y(f(x), f(x_n)) \geq \varepsilon_0.$$

Then $x_n \to x$, but $f(x_n)$ does not converge to $f(x)$, contradicting (3). Therefore f is continuous at x. $\square$

Corollary 14.3 Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : X \to Y$. Then the following are equivalent:

1. f is continuous.
2. Whenever V is open in Y, then $f^{-1}(V)$ is open in X.
3. Whenever $\{x_n\}_n$ is a sequence in X which converges to some $x \in X$, then $\lim_{n\to\infty} f(x_n) = f(x)$.

Proof. (1) *implies* (2). Let $V \subseteq Y$ be open and let $x \in f^{-1}(V)$. Since V is a neighborhood of $f(x)$ and f is continuous at x, $f^{-1}(V)$ is a neighborhood of x. Thus $f^{-1}(V)$ is open.

(2) *implies* (3). Let $x_n \to x$ in X and let W be a neighborhood of $f(x)$. Choose an open set $V \subseteq W$ with $f(x) \in V$. By (2), $f^{-1}(V)$ is open and contains x, so eventually $x_n \in f^{-1}(V)$. Hence eventually $f(x_n) \in V \subseteq W$, that is, $f(x_n) \to f(x)$.

(3) *implies* (1). Suppose f is not continuous at x. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ there exists y with $d_X(y, x) < \delta$ but $d_Y(f(y), f(x)) \geq \varepsilon_0$. For $n \geq 1$, take $\delta = 1/n$ and choose x_n with

$$d_X(x_n, x) < \frac{1}{n} \qquad \text{and} \qquad d_Y(f(x_n), f(x)) \geq \varepsilon_0.$$

Then $x_n \to x$, but $f(x_n) \not\to f(x)$, contradicting (3). Thus f is continuous. □

Example 14.4 Let (X, d_X) be a discrete metric space and let (Y, d_Y) be an arbitrary metric space. Since every subset of X is open, every function $f : X \to Y$ is continuous.

Definition 14.5 Let (X, d_X) and (Y, d_Y) be metric spaces. A bijection $f : X \to Y$ is a **homeomorphism** if both f and f^{-1} are continuous.

If f is a homeomorphism from $X \to Y$, then a subset $U \subset X$ is open if and only if $f(U)$ is open in Y. Thus, X and Y can be identified topologically.

3. Compactness

Topological Spaces

Definition 14.6 Let $X \neq \emptyset$. A subset $\tau \subset \mathcal{P}(X)$ is a **topology** and (X, τ) is a **topological space** if the following are satisfied:

1. $\emptyset, X \in \tau$.
2. If $\{U_\alpha\}_{\alpha \in A} \subset \tau$, then $\bigcup_{\alpha \in A} U_\alpha \in \tau$.
3. If $U_1, \ldots, U_n \in \tau$, then $\bigcap_{i=1}^{n} U_i \in \tau$.

The elements of τ are said to be **$\boldsymbol{\tau}$-open**. A subset $F \subset X$ is said to be **closed** if $X \setminus F \in \tau$.

Example 14.7 Let (X, d) be a metric space and let

$$\tau = \{U \subset X \mid U \text{ is open in the metric space sense}\}.$$

Equivalently, $U \in \tau$ if and only if for every $x \in U$, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset U$. By Theorem 11.3, τ is a topology on X, called the **metric topology induced by $\boldsymbol{d}$**.

Example 14.8

1. Let $X \neq \emptyset$ and let $\tau = \{\emptyset, X\}$. Then τ is a topology on X, called the **trivial topology**.
2. Let $X \neq \emptyset$ and let $\tau = \mathcal{P}(X)$. This topology is called the **discrete topology**. This is the metric topology induced by the discrete metric on X.

Lecture 15 — Friday, October 10, 2014

Example 15.1

1. Let $X = \{a, b\}$ and $\tau = \{\emptyset, X, \{a\}\}$. Then τ is a topology on X.
2. Let $X \neq \emptyset$ and let $\tau_{cf} = \{\emptyset\} \cup \{U \subset X \mid X \setminus U \text{ is finite}\}$. Then τ_{cf} is a topology on

X, called the **cofinite topology**.

Definition 15.2 Let (X, τ) be a topological space, $x \in N$ and $N \subset X$. N is said to be a **neighborhood of x** if there exists $U \in \tau$ with $x \in U \subset N$.

Definition 15.2 is consistent with our definition of neighborhoods for metric spaces: if (X, d) is a metric space, then $N \subset X$ is a neighborhood of x if and only if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset N$.

Lemma 15.3 Let (X, τ) be a topological space and let $U \subset X$. Then $U \in \tau$ if and only if U is a neighborhood of each of its points.

Proof. Suppose that U is open and let $x \in U$. Then $x \in U \subset U$ and $U \in \tau$, so U is a neighborhood of x. Conversely, suppose U is a neighborhood of each of its points. Then for $x \in U$, there exists $U_x \in \tau$ such that $x \in U_x \subset U$. Then $U \subset \bigcup_{x \in U} U_x$, that is, $U = \bigcup_{x \in U} U_x$, and hence $U \in \tau$. $\square$

Definition 15.4 Let (X, τ_X) and (Y, τ_Y) be topological spaces and let $f : X \to Y$. Then f is said to be **continuous at $x \in X$** if for every neighborhood W of $f(x)$, $f^{-1}(W)$ is a neighborhood of x. We say that f is continuous if it is continuous at each $x \in X$.

By Theorem 14.2, this definition is consistent with the corresponding definition for metric spaces.

Lemma 15.5 Let (X, τ_X) and (Y, τ_Y) be topological spaces, and $f : X \to Y$. Then f is continuous if and only if $f^{-1}(V) \in \tau_X$ for every $V \in \tau_Y$.

Proof. *Forward implication.* Assume f is continuous (in the neighborhood sense). Let $V \in \tau_Y$ and $x \in f^{-1}(V)$. Since V is a neighborhood of $f(x)$, continuity gives that $f^{-1}(V)$ is a neighborhood of x. Thus $f^{-1}(V)$ is a neighborhood of each of its points; by Lemma 15.3, $f^{-1}(V) \in \tau_X$.

Reverse implication. Conversely, assume $f^{-1}(V) \in \tau_X$ for every $V \in \tau_Y$. Let $x \in X$ and let W be a neighborhood of $f(x)$. Choose $V \in \tau_Y$ with $f(x) \in V \subseteq W$. Then $f^{-1}(V) \in \tau_X$, $x \in f^{-1}(V) \subseteq f^{-1}(W)$, so $f^{-1}(W)$ is a neighborhood of x. Hence f is continuous at x, and therefore continuous. $\square$

Definition 15.6 Let (X, τ) be a topological space and $Y \subset X$. The **relative topology** on Y inherited from X is $\tau_Y = \{U \cap Y \mid U \in \tau\}$. We say that $U \in \tau_Y$ is **relatively open** in Y. Also, $F \subset Y$ is **relatively closed** in Y if $Y \setminus F \in \tau_Y$.

Example 15.7 Let τ be the metric topology on $\mathbb{R}$, induced by the Euclidean metric. Consider $\mathbb{Z} \subset \mathbb{R}$. Let $n \in \mathbb{Z}$. Then

$$\{n\} = (n-1, n+1) \cap \mathbb{Z} \in \tau_{\mathbb{Z}}.$$

If $A \subset \mathbb{Z}$, then $A = \bigcup_{n \in A} \{n\} \in \tau_{\mathbb{Z}}$. So $\tau_{\mathbb{Z}}$ is the discrete topology!

Many notions from metric spaces (e.g., closure, interior, and boundary) extend to topological spaces. For example, if $A \subset X$ and $x \in X$, then $x \in \overline{A}$ if and only if $A \cap N \neq \emptyset$ for every neighborhood N of x.

Definition 15.8 Let (X, τ) be a topological space, and let $\{x_n\}_n$ be a sequence in X and $x \in X$. We say that $\{x_n\}_n$ **converges to x** if for every neighborhood N of x, there exists N_0 such that $x_n \in N$ for all $n \geq N_0$.

Example 15.9 Let $X \neq \emptyset$ and $\tau = \{\emptyset, X\}$. Let $\{x_n\}_n$ be a sequence in X. Every $x \in X$ has only one neighborhood, namely X itself, so every sequence converges to every point of X. Thus limits need not be unique in topological spaces.

Definition 15.10 A topological space (X, τ) is said to be **Hausdorff** if for every $x, y \in X$ with $x \neq y$, there exist open sets U, V with $U \cap V = \emptyset$ and $x \in U, y \in V$.

Example 15.11

1. If $|X| \geq 2$ and $\tau = \{\emptyset, X\}$, then (X, τ) is not Hausdorff.
2. If (X, d) is a metric space and τ is the metric topology induced by d, then (X, τ) is Hausdorff. If $x \neq y \in X$, then $r = d(x, y) > 0$. Also $B_{r/2}(x) \cap B_{r/2}(y) = \emptyset$.

Lemma 15.12 If (X, τ) is a Hausdorff space, then limits of sequences in X are unique.

Proof. Let $\{x_n\}_n$ be a sequence in X such that $x_n \to x$ and $x_n \to y$.

Suppose, for contradiction, that $x \neq y$. Since (X, τ) is Hausdorff, there exist open sets $U, V \in \tau$ such that

$$x \in U, \qquad y \in V, \qquad \text{and} \qquad U \cap V = \emptyset.$$

Because $x_n \to x$, there exists $N_1 \in \mathbb{N}$ such that $x_n \in U$ for all $n \geq N_1$. Because $x_n \to y$, there exists $N_2 \in \mathbb{N}$ such that $x_n \in V$ for all $n \geq N_2$.

If $n \geq \max\{N_1, N_2\}$, then $x_n \in U$ and $x_n \in V$, so $x_n \in U \cap V = \emptyset$, which is impossible.

This contradiction shows that $x \neq y$ cannot occur. Therefore $x = y$, so the limit is unique. $\square$

Warning In general topological spaces (even in Hausdorff spaces), sequences are not enough to characterize closures of sets or continuity of functions.

Lecture 16 — Wednesday, October 15, 2014

Compactness

Definition 16.1 Let (X, τ) be a topological space and let $A \subset X$. An **open cover** of A is a collection $\{U_\lambda\}_{\lambda \in \Lambda}$ of open subsets of X such that $A \subset \bigcup_{\lambda \in \Lambda} U_\lambda$. Given an open cover of A as above, a **finite subcover** is a finite subcollection $\{U_{\lambda_1}, \ldots, U_{\lambda_n}\}$ such that $A \subset \bigcup_{i=1}^n U_{\lambda_i}$. A subset $K \subset X$ is called **compact** if every open cover of K admits a finite subcover.

Compactness is the ability to discard all but finitely many members of any open cover. In Figure 7, the coloured sets already cover K; any additional members of the original cover are superfluous.

Example 16.2

1. Finite subspaces of topological spaces are compact.
2. If (X, τ) is a topological space endowed with the trivial topology, then every subset of X is compact.

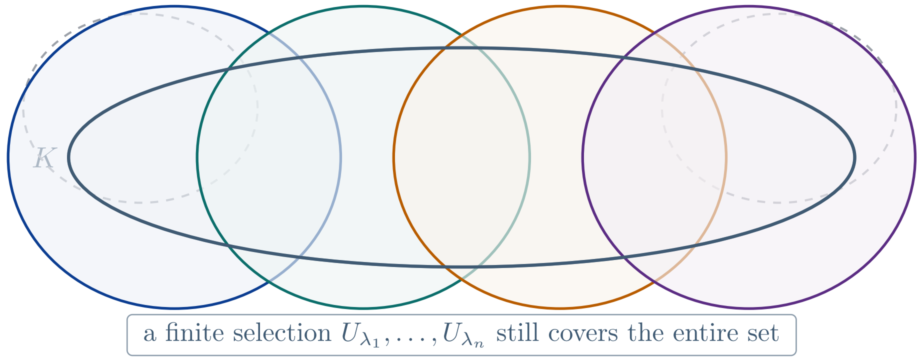


Figure 7

3. If (X, τ) is a space endowed with the discrete topology, then $K \subset X$ is compact if and only if K is finite; indeed, every finite K is compact by (1). Conversely, if K is compact, then the open cover $\{\{x\}\}_{x \in K}$ has a finite subcover, so K is finite.
4. $\mathbb{R}$ equipped with the Euclidean topology is not compact. Indeed, the open cover $\{(n, \infty)\}_{n \in \mathbb{Z}}$ has no finite subcover.
5. $(0,1) \subset \mathbb{R}$ is not compact. Indeed, $\{(x, 2)\}_{x \in (0,1)}$ has no finite subcover.

Theorem 16.3 Let $\mathbb{R}$ be endowed with the Euclidean topology, and let $a, b \in \mathbb{R}$ with $a < b$. Then $[a, b]$ is compact.

Proof. Assume for a contradiction that $\{U_\lambda\}_{\lambda \in \Lambda}$ is an open cover of $[a, b]$ which does not admit a finite subcover. Let $I_0 = [a, b]$. Split I_0 into two subintervals $[a, c]$ and $[c, b]$ where $c = (a + b)/2$. Since $[a, b]$ cannot be covered by finitely many of the U_λ, one of these subintervals — call it I_1 — cannot be covered by finitely many of the U_λ. Proceeding recursively, we obtain a sequence $\{I_n\}_{n=0}^{\infty}$ of closed subintervals of $[a, b]$ such that the following hold:

1. $I_{n+1} \subset I_n$ for all $n \in \mathbb{N}$.
2. $\operatorname{diam}(I_n) = \frac{b-a}{2^n}$ for all $n \in \mathbb{N}$.
3. No I_n can be covered by finitely many of the U_λ.

By (1) and (2) and the nested intervals theorem, $\bigcap_{n=0}^{\infty} I_n = \{z\}$ for some $z \in [a, b]$. Since $\{U_\lambda\}_{\lambda \in \Lambda}$ is an open cover of $[a, b]$, there exists $\lambda_0 \in \Lambda$ such that $z \in U_{\lambda_0}$. Since U_{λ_0} is open, there exists $\varepsilon > 0$ such that $(z - \varepsilon, z + \varepsilon) \subset U_{\lambda_0}$. But then $I_n \subset (z - \varepsilon, z + \varepsilon) \subset U_{\lambda_0}$ whenever n is so large that $\operatorname{diam}(I_n) < \varepsilon$. This contradicts (3). Therefore, $[a, b]$ is compact. □

Remark 16.4 Don't fret! We will get to the Heine–Borel theorem soon enough.

Example 16.5 Let X be an infinite set and let d be the discrete metric on X. Then X is closed in itself and bounded, since $d(x, y) \leq 1$ for all $x, y \in X$. It is not compact: the open cover $\{\{x\}\}_{x\in X}$ has no finite subcover.

Proposition 16.6 Let (X, τ) be a topological space and let $Y \subset X$ and $K \subset Y$. Then the following are equivalent:

1. K is compact as a subset of (X, τ).
2. K is compact as a subset of (Y, τ_Y).

So compactness does not rely on the ambient space.

Proof. (1) *implies* (2). Assume K is compact as a subset of (X, τ) and let $\{V_\lambda\}_{\lambda\in\Lambda}$ be a τ_Y-open cover of K.

For each $\lambda \in \Lambda$, there exists $U_\lambda \in \tau$ such that $V_\lambda = Y \cap U_\lambda$. Then $\{U_\lambda\}_{\lambda\in\Lambda}$ is a τ-open cover of K. By compactness of K in (X, τ), there are $\lambda_1, \ldots, \lambda_n$ such that $K \subset \bigcup_{i=1}^n U_{\lambda_i}$. Since $K \subset Y$,

$$K = K \cap Y = (\bigcup_{i=1}^n U_{\lambda_i}) \cap Y = \bigcup_{i=1}^n (U_{\lambda_i} \cap Y) = \bigcup_{i=1}^n V_{\lambda_i}.$$

Thus, $\{V_\lambda\}_{\lambda\in\Lambda}$ admits a finite subcover of K, and hence K is compact in (Y, τ_Y).

(2) *implies* (1). Assume K is compact as a subset of (Y, τ_Y) and let $\{U_\lambda\}_{\lambda\in\Lambda}$ be a τ-open cover of K. For each $\lambda \in \Lambda$, set $V_\lambda = U_\lambda \cap Y$. Then $V_\lambda \in \tau_Y$ for every λ, and

$$K \subseteq \bigcup_{\lambda\in\Lambda} U_\lambda \implies K = K \cap Y \subseteq \bigcup_{\lambda\in\Lambda} (U_\lambda \cap Y) = \bigcup_{\lambda\in\Lambda} V_\lambda.$$

So $\{V_\lambda\}_{\lambda\in\Lambda}$ is a τ_Y-open cover of K. By compactness of K in (Y, τ_Y), there exist $\lambda_1, \ldots, \lambda_n \in \Lambda$

such that

$$K \subseteq \bigcup_{i=1}^{n} V_{\lambda_i} \subseteq \bigcup_{i=1}^{n} U_{\lambda_i}.$$

Hence $\{U_\lambda\}_{\lambda\in\Lambda}$ admits a finite subcover of K, and therefore K is compact as a subset of (X, τ). $\square$

Proposition 16.7 Let (X, τ) be a topological space. If X is compact, and F is a closed subset of X, then F is compact.

Proof. Let $\{U_\lambda\}_{\lambda\in\Lambda}$ be an open cover of F. Since F is closed, $\{U_\lambda\}_{\lambda\in\Lambda} \cup \{X \setminus F\}$ is an open cover of X. By compactness of X, there exist $\lambda_1, \ldots, \lambda_n \in \Lambda$ such that $X \subset (\bigcup_{i=1}^n U_{\lambda_i}) \cup (X \setminus F)$. But then $F \subset \bigcup_{i=1}^n U_{\lambda_i}$, so $\{U_\lambda\}_{\lambda\in\Lambda}$ admits a finite subcover of F. Thus F is compact. $\square$

Lecture 17 — Friday, October 17, 2014

Proposition 17.1 Let (X, τ) be a topological space. If X is Hausdorff and $K \subset X$ is compact, then K is closed.

Proof. Let $K \subset X$ be a compact subset. We want to show that $X \setminus K$ is open. Let $x \in X \setminus K$. Since X is Hausdorff, for every $y \in K$, there exist disjoint open sets U_y and V_y such that $y \in U_y$ and $x \in V_y$. Then $\{U_y\}_{y\in K}$ is an open cover, and by compactness has a finite subcover: there exist $y_1, \ldots, y_n \in K$ such that $K \subset \bigcup_{i=1}^n U_{y_i}$. Let $V = \bigcap_{i=1}^n V_{y_i}$. Then V is open and $V \cap U_{y_i} = \emptyset$ for all $i = 1, \ldots, n$. Hence $V \cap K = \emptyset$, so $V \subset X \setminus K$ and $x \in V$. Therefore, $X \setminus K$ is a neighborhood of each of its points, and hence open. Thus K is closed. $\square$

Definition 17.2 Let (X, d) be a metric space and let $\emptyset \neq A \subset X$. The **diameter** of A is defined to be

$$\operatorname{diam}(A) = \sup\{d(x, y) \mid x, y \in A\}.$$

We set $\operatorname{diam}(\emptyset) = 0$. We say A is **bounded** if $\operatorname{diam}(A) < \infty$.

Example 17.3

1. If $(X, \|\cdot\|)$ is a normed space and $\emptyset \neq A \subset X$, then A is bounded if and only if
$$\sup\{\|x\| \mid x \in A\} < \infty.$$
2. If (X, d) is a metric space and $x \in X, r > 0$, then $\operatorname{diam}(B_r(x)) \leq 2r$. Equality does not hold in general; for example, if d is the discrete metric and $0 < r \leq 1$, then $B_r(x) = \{x\}$ and hence $\operatorname{diam}(B_r(x)) = 0$.

Proposition 17.4 Let (X, d) be a metric space and let $K \subset X$ be compact. Then K is bounded.

Proof. Fix $x \in X$. The collection $\{B_r(x) \mid r > 0\}$ is an open cover of K. Since K is compact, this cover admits a finite subcover. Therefore, there exists $R > 0$ such that $K \subset B_R(x)$, and hence
$$\operatorname{diam}(K) \leq 2R < \infty.$$
□

Theorem 17.5 Let (X, τ_X) and (Y, τ_Y) be topological spaces, and let $f : X \to Y$ be continuous. If $K \subset X$ is compact, then $f(K) \subset Y$ is compact.

Proof. Let $\{U_\lambda\}_{\lambda \in \Lambda}$ be an open cover of $f(K)$. Because f is continuous, $\{f^{-1}(U_\lambda)\}_{\lambda \in \Lambda}$ is an open cover of K. Because K is compact, there are $\lambda_1, \ldots, \lambda_n \in \Lambda$ such that $K \subset \bigcup_{i=1}^n f^{-1}(U_{\lambda_i})$. Therefore,
$$f(K) \subseteq f\left(\bigcup_{i=1}^n f^{-1}(U_{\lambda_i})\right) \subseteq \bigcup_{i=1}^n U_{\lambda_i}.$$
Therefore $f(K)$ is compact. □

Theorem 17.6 Let (X, τ) be a compact topological space and let $f : X \to \mathbb{R}$ be continuous. Then f attains its maximum and its minimum (that is, there exists $x_n, x_m \in X$ such that $f(x_n) \leq f(x) \leq f(x_m)$ for all $x \in X$).

Proof. By Theorem 17.5, $f(X) \subset \mathbb{R}$ is compact. By Propositions 17.1 and 17.4, $f(X)$ is closed and bounded. Thus $\sup f(X) \in f(X)$ and $\inf f(X) \in f(X)$. Choose $x_n, x_m \in X$ such that $f(x_n) = \inf f(X)$ and $f(x_m) = \sup f(X)$. $\square$

Definition 17.7 Let X be a set. A collection of subsets $\mathcal{F} \subset \mathcal{P}(X)$ is said to have the **finite intersection property** if for every finite subcollection $\{F_1, \ldots, F_n\} \subseteq \mathcal{F}$, we have $\bigcap_{i=1}^n F_i \neq \emptyset$.

Theorem 17.8 Let (X, τ) be a topological space. Then the following are equivalent:

1. X is compact.
2. If $\{F_\lambda\}_{\lambda \in \Lambda}$ is a collection of closed subsets of X with finite intersection property, then $\bigcap_{\lambda \in \Lambda} F_\lambda \neq \emptyset$.

Proof. (1) *implies* (2). Suppose that X is compact and $\{F_\lambda\}_{\lambda \in \Lambda}$ is a collection of closed subsets of X with finite intersection property. Assume for a contradiction that $\bigcap_{\lambda \in \Lambda} F_\lambda = \emptyset$. For $\lambda \in \Lambda$, let $U_\lambda = X \setminus F_\lambda$. Then $\{U_\lambda\}_{\lambda \in \Lambda}$ is an open cover of X. Since X is compact, there are $\lambda_1, \ldots, \lambda_n \in \Lambda$ such that $X = \bigcup_{i=1}^n U_{\lambda_i}$, and hence $\bigcap_{i=1}^n F_{\lambda_i} = \emptyset$. This contradicts finite intersection property for $\{F_\lambda\}_{\lambda \in \Lambda}$.

(2) *implies* (1). Suppose (2) holds and let $\{U_\lambda\}_{\lambda \in \Lambda}$ be an open cover of X. For $\lambda \in \Lambda$, let $F_\lambda = X \setminus U_\lambda$. Then each F_λ is closed, and $\bigcap_{\lambda \in \Lambda} F_\lambda = \emptyset$ since $\bigcup_{\lambda \in \Lambda} U_\lambda = X$. By (2), $\{F_\lambda\}_{\lambda \in \Lambda}$ does not have finite intersection property. Hence there are $\lambda_1, \ldots, \lambda_n \in \Lambda$ such that $\bigcap_{i=1}^n F_{\lambda_i} = \emptyset$, so then $\bigcup_{i=1}^n U_{\lambda_i} = X$. Thus $\{U_\lambda\}_{\lambda \in \Lambda}$ admits a finite subcover, and X is compact. $\square$

Lecture 18 — Monday, October 20, 2014

Recall that in a compact topological space (X, τ), if $\{F_\alpha\}_{\alpha \in A}$ is a family of closed sets with the finite intersection property, then $\bigcap_{\alpha \in A} F_\alpha \neq \emptyset$.

Corollary 18.1 Let (X, τ) be a compact topological space. If

$$F_1 \supseteq F_2 \supseteq F_3 \supseteq \dots$$

is a sequence of nonempty closed subsets of X, then $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$. This is a generalization of the nested interval theorem.

Proof. The family $\{F_n\}_{n\in\mathbb{N}}$ has the finite intersection property, and hence, by Theorem 17.8, $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$. □

Definition 18.2 Let (X, τ) be a topological space, and let $A \subset X$. An element $x \in X$ is said to be an **accumulation point** of A if $(N \setminus \{x\}) \cap A \neq \emptyset$ for every neighborhood N of x.

Definition 18.3 A topological space (X, τ) is said to be **sequentially compact** if every sequence $\{x_n\}_n$ in X admits a subsequence $\{x_{n_k}\}$ which converges to a point in X. We say that (X, τ) has the **Bolzano–Weierstrass property** if every infinite subset of X has an accumulation point in X.

Example 18.4

1. The closed interval $[0, 1]$ with the standard topology is sequentially compact. Indeed, every sequence in $[0, 1]$ is bounded and therefore contains a convergent subsequence by the usual Bolzano–Weierstrass theorem from MATH 147. Since $[0, 1]$ is closed, the limit of that subsequence remains in $[0, 1]$.

2. Let $1 \leq p \leq \infty$ and let
$$B = \{x \in \ell_p : \|x\|_p \leq 1\}$$
be the unit ball in ℓ_p. Let $e_n = (0, \dots, 0, 1, 0, \dots)$, where the 1 appears in the nth coordinate. We have $\|e_n\|_p = 1$, so $\{e_n\}_n \subset B$. For $m \neq n$,
$$\|e_m - e_n\|_p = \begin{cases} 2^{1/p} & \text{if } 1 \leq p < \infty \\ 1 & \text{if } p = \infty \end{cases}.$$
So the distance between any two distinct terms in the sequence $\{e_n\}_n$ is bounded

below by a positive constant. Therefore, no subsequence of $\{e_n\}_n$ can be Cauchy, and in particular $\{e_n\}_n$ has no convergent subsequence. Thus B contains a sequence with no convergent subsequence, so B is not sequentially compact.

Theorem 18.5 If (X, τ) is compact or sequentially compact, then X has the Bolzano–Weierstrass property.

Proof. First assume that (X, τ) is compact. For a contradiction, assume X has an infinite subset A with no accumulation point. Then for every $x \in X$, there exists a neighborhood (assume to be open) U_x of x such that $(U_x \setminus \{x\}) \cap A = \emptyset$, and hence $U_x \cap A \subseteq \{x\}$. Then $\{U_x\}_{x \in X}$ is an open cover of X, so by compactness of X, there are $x_1, \ldots, x_n \in X$ such that $X = \bigcup_{i=1}^n U_{x_i}$. Then

$$A = A \cap X = \bigcup_{i=1}^n (A \cap U_{x_i}) \subseteq \{x_1, \ldots, x_n\},$$

a contradiction, because A is infinite. Hence X has the Bolzano–Weierstrass property.

Now assume that (X, τ) is sequentially compact, and let A be an infinite subset of X. Let $\{a_n\}_{n \in \mathbb{N}}$ be a countable infinite subset of A. By sequential compactness, $\{a_n\}_{n \in \mathbb{N}}$ has a convergent subsequence $\{a_{n_k}\}$ such that $\lim_{k \to \infty} a_{n_k} = x \in X$. Let N be a neighborhood of x. Then there exists $K \in \mathbb{N}$ such that $a_{n_k} \in N$ for all $k \geq K$. Therefore, $|N \cap A| = \infty$. In particular, $(N \setminus \{x\}) \cap A \neq \emptyset$, so x is an accumulation point of A. Thus X has the Bolzano–Weierstrass property. $\square$

Warning In general topological spaces, none of the remaining implications between compactness, sequential compactness, and the Bolzano–Weierstrass property hold true. Counterexamples are a bit too hard for us to construct in this course, but they do exist.

Compactness in Metric Spaces

Proposition 18.6 If (X, d) is a metric space with the Bolzano–Weierstrass property, then (X, d) is sequentially compact.

Proof. Let $\{x_n\}_n$ be a sequence in X. If $\{x_n\}_n$ admits a constant subsequence, then that subsequence converges to that constant value. So we may assume that $\{x_n\}_n$ has no constant subsequence, and hence that $A = \{x_n \mid n \in \mathbb{N}\}$ is infinite. By the Bolzano–Weierstrass property, A has an accumulation point $x \in X$. We claim that $A \cap B_\epsilon(x)$ is infinite for every $\epsilon > 0$. Otherwise, for some $\epsilon_0 > 0$, the set

$$F := (A \cap B_{\epsilon_0}(x)) \setminus \{x\}$$

is finite. If $F = \emptyset$, then $(B_{\epsilon_0}(x) \setminus \{x\}) \cap A = \emptyset$, contradicting that x is an accumulation point. If $F \neq \emptyset$, let

$$\eta := \min\{d(x,y) : y \in F\} > 0.$$

Then $(B_r(x) \setminus \{x\}) \cap A = \emptyset$ for $r = \min\{\epsilon_0, \eta/2\}$, again a contradiction. So $A \cap B_\epsilon(x)$ is infinite for all $\epsilon > 0$.

We will recursively construct a subsequence $\{x_{n_k}\}$ of $\{x_n\}_n$ which converges to x. Let n_1 be such that $x_{n_1} \in B_1(x)$. If $x_{n_1}, \ldots, x_{n_k}$ have been constructed, let $n_{k+1} > n_k$ be such that $x_{n_{k+1}} \in B_{\frac{1}{k+1}}(x)$. This is possible because $B_{\frac{1}{k+1}}(x) \cap A$ is infinite. Then $d(x, x_{n_k}) < 1/k$ for all $k \in \mathbb{N}$, so $\lim_{k\to\infty} x_{n_k} = x$. Therefore $\{x_n\}_n$ has a convergent subsequence, and so X is sequentially compact. $\square$

For metric spaces, we have shown that compactness implies the Bolzano–Weierstrass property, and that the Bolzano–Weierstrass property implies sequential compactness. In any topological space, sequential compactness implies the Bolzano–Weierstrass property. The remaining direction for metric spaces will soon follow from total boundedness.

Definition 18.7 Let (X,d) be a metric space. A collection of elements $\{x_\lambda\}_{\lambda\in\Lambda} \subset X$ is called an **ε-net** if $\bigcup_{\lambda\in\Lambda} B_\varepsilon(x_\lambda) = X$. We say that X is **totally bounded** if for every $\varepsilon > 0$, X admits a finite ε-net. A subset $A \subset X$ is called totally bounded if $(A, d\big|_{A\times A})$ is totally bounded.

Lecture 19 — Friday, October 24, 2014

Proposition 19.1 Let (X,d) be a metric space. If X is totally bounded, then X is bounded.

Proof. Since X is totally bounded, X admits a finite 1-net $\{x_1, \ldots, x_n\}$. Let

$$R = \max\{d(x_i, x_j) : 1 \leq i, j \leq n\} < \infty.$$

If $x, y \in X$, then there exist $i, j \in \{1, \ldots, n\}$ such that $x \in B_1(x_i)$ and $y \in B_1(x_j)$. Thus,

$$d(x, y) \leq d(x, x_i) + d(x_i, x_j) + d(x_j, y) \leq R + 2,$$

so $\operatorname{diam}(X) \leq R + 2 < \infty$. □

Example 19.2

1. The open interval $(0,1)$ is totally bounded. Indeed, given $\varepsilon > 0$, let $N \geq 2$ with $1/N < \varepsilon$. Then
$$\left\{\frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}\right\}$$
is a finite ε-net for $(0,1)$.
2. If X is an infinite set and d is the discrete metric on X, then X is bounded, but not totally bounded. Indeed, X does not admit a finite $(1/2)$-net.
3. Let $1 \leq p \leq \infty$ and let
$$B = \{x \in \ell_p : \|x\|_p \leq 1\}$$
be the unit ball in ℓ_p. B is not sequentially compact by Example 18.4(2), and hence not compact. For $n \in \mathbb{N}$, define $e^{(n)} = (e_k^{(n)})_k \in B$ by
$$e_k^{(n)} = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k \end{cases}.$$
Then $\left\|e^{(n)} - e^{(m)}\right\|_p \geq 1$ if $n \neq m$. Thus, if $x \in B$, then $B_{1/2}(x)$ contains at most one of the $e^{(n)}$, so B does not admit a finite $(1/2)$-net. Hence, B is not totally bounded.

Proposition 19.3 If (X, d) is a sequentially compact metric space, then X is totally bounded.

Proof. Suppose that X is not totally bounded. Then there exists $\varepsilon > 0$ such that X does not admit a finite ε-net. Recursively, we may construct a sequence (x_n) in X such that

$$x_n \notin \bigcup_{k=0}^{n-1} B_\varepsilon(x_k)$$

for all $n \geq 1$. It follows that $d(x_n, x_m) \geq \varepsilon$ if $n \neq m$, so that (x_n) does not admit a Cauchy subsequence. In particular, (x_n) has no convergent subsequence, hence X is not sequentially compact. The assertion now follows. $\square$

That proof actually shows the following somewhat stronger statement: if every sequence in X admits a subsequence which is Cauchy, then X is totally bounded. We will see that the converse of this statement is also true.

Definition 19.4 Let $\{U_\lambda\}_{\lambda\in\Lambda}$ be an open cover of a metric space (X, d). We say that $\delta > 0$ is a **Lebesgue number** for $\{U_\lambda\}_{\lambda\in\Lambda}$ if for every $A \subseteq X$ with $\operatorname{diam}(A) < \delta$, there exists $\lambda_0 \in \Lambda$ such that $A \subseteq U_{\lambda_0}$.

Lemma 19.5 (Lebesgue's covering lemma) Let (X, d) be a sequentially compact metric space and let $\{U_\lambda\}_{\lambda\in\Lambda}$ be an open cover of X. Then $\{U_\lambda\}_{\lambda\in\Lambda}$ admits a (positive) Lebesgue number.

Proof. Suppose otherwise. Then for each $n \geq 0$, there exists $A_n \subseteq X$ such that $\operatorname{diam}(A_n) < \frac{1}{n+1}$, and such that $A_n \not\subseteq U_\lambda$ for all $\lambda \in \Lambda$. Choose $x_n \in A_n$ for each $n \in \mathbb{N}$. Since X is sequentially compact, (x_n) admits a convergent subsequence (x_{n_k}), say $x = \lim_{k\to\infty} x_{n_k}$. Since $\{U_\lambda\}_{\lambda\in\Lambda}$ is an open cover of X, there exists $\lambda_0 \in \Lambda$ such that $x \in U_{\lambda_0}$. Since U_{λ_0} is open, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U_{\lambda_0}$. Because $\lim_{k\to\infty} x_{n_k} = x$, there exists $K \in \mathbb{N}$ such that

$$d(x_{n_k}, x) < \frac{\varepsilon}{2}$$

whenever $k \geq K$, and we may assume that $\frac{1}{n_K+1} \leq \frac{\varepsilon}{2}$. If $k \geq K$ and $y \in A_{n_k}$, then

$$d(y, x) \leq d(y, x_{n_k}) + d(x_{n_k}, x) \leq \operatorname{diam}(A_{n_k}) + d(x_{n_k}, x) < \frac{1}{n_K + 1} + \frac{\varepsilon}{2} \leq \varepsilon,$$

hence $y \in B_\varepsilon(x)$. Therefore,

$$A_{n_k} \subseteq B_\varepsilon(x) \subseteq U_{\lambda_0}$$

for $k \geq K$, a contradiction. Therefore, $\{U_\lambda\}_{\lambda\in\Lambda}$ admits a Lebesgue number. $\square$

Lecture 20 — Monday, October 27, 2014

Recall that if $\{U_\lambda\}_\lambda$ is an open cover of (X, d), then $\delta > 0$ is a Lebesgue number if, for every $A \subset X$ with $\text{diam}(A) < \delta$, there is a $\lambda_0 \in \Lambda$ such that $A \subset U_{\lambda_0}$.

Theorem 20.1 Let (X, d) be a metric space. The following are equivalent:

1. X is compact.
2. X is sequentially compact.
3. X has the Bolzano–Weierstrass property.

Proof. We only need to show that (2) $\implies$ (1). Suppose X is sequentially compact, and let $\{U_\lambda\}_\lambda$ be an open cover of X.

By Lemma 19.5, $\{U_\lambda\}_\lambda$ admits a Lebesgue number δ. Let $0 < \varepsilon < \delta/2$. By Proposition 19.3, X admits a finite ε-net $\{x_1, \ldots, x_n\}$. Since $\text{diam}(B_\varepsilon(x_i)) \leq 2\varepsilon < \delta$, for $k = 1, \ldots, n$, there is a $\lambda_k \in \Lambda$ such that $B_\varepsilon(x_k) \subset U_{\lambda_k}$. Then

$$X = \bigcup_{k=1}^{n} B_\varepsilon(x_k) \subset \bigcup_{k=1}^{n} U_{\lambda_k},$$

so $\{U_\lambda\}_\lambda$ admits a finite subcover. □

Corollary 20.2 (Heine–Borel theorem) Suppose that $\mathbb{R}^n$ is equipped with the Euclidean topology, and let $K \subset \mathbb{R}^n$. The following are equivalent:

1. K is compact.
2. K is sequentially compact.
3. K is closed and bounded.

Proof. (1) $\Leftrightarrow$ (2) by Theorem 20.1, and (1) $\implies$ (3) by Propositions 17.1 and 17.4.

(3) *implies* (2). Let K be closed and bounded. Let $\{x_n\}_n$ be a sequence in K. By the Bolzano–

Weierstrass theorem, $\{x_n\}_n$ has a convergent subsequence $\{x_{n_k}\}_k$. Since K is closed, $\lim_{k\to\infty} x_{n_k} \in K$. Thus K is sequentially compact. $\square$

Applications of Compactness

Definition 20.3 Let $(X, d_X), (Y, d_Y)$ be metric spaces. A function $f : X \to Y$ is said to be **uniformly continuous** if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), f(y)) < \varepsilon$ whenever $d_X(x, y) < \delta$.

Example 20.4 The continuous function $f : \mathbb{R} \to \mathbb{R}$ defined by $x \mapsto x^2$ is not uniformly continuous. Indeed, let $\varepsilon = 1$. Given any $\delta > 0$, choose

$$x = \frac{1}{\delta} \qquad \text{and} \qquad y = x + \frac{\delta}{2}.$$

Then $|x - y| = \frac{\delta}{2} < \delta$, but

$$|f(x) - f(y)| = |x^2 - y^2| = |x - y|\,|x + y| = \frac{\delta}{2}\left(\frac{2}{\delta} + \frac{\delta}{2}\right) = 1 + \frac{\delta^2}{4} > 1 = \varepsilon.$$

So no single δ works for $\varepsilon = 1$, and therefore f is not uniformly continuous on $\mathbb{R}$.

Theorem 20.5 Let (X, d_X) and (Y, d_Y) be metric spaces, and assume that X is compact. If $f : X \to Y$ is continuous, then f is uniformly continuous.

Proof. Let $\varepsilon > 0$. Since f is continuous, $\{f^{-1}(B_{\varepsilon/2}(z))\}_{z\in Y}$ is an open cover of X. By Lebesgue's covering lemma, this open cover admits a Lebesgue number $\delta > 0$. If $d_X(x_1, x_2) < \delta$, then $\operatorname{diam}(\{x_1, x_2\}) < \delta$. Hence there exists $z \in Y$ such that $\{x_1, x_2\} \subset f^{-1}(B_{\varepsilon/2}(z))$. Thus $f(x_1), f(x_2) \in B_{\varepsilon/2}(z)$, and

$$d_Y(f(x_1), f(x_2)) \le d_Y(f(x_1), z) + d_Y(z, f(x_2)) < \varepsilon.$$

$\square$

Let X be a $\mathbb{K}$-vector space. Recall that two norms $\|\cdot\|, \|\cdot\|'$ on X are **equivalent** if there exist $C_1, C_2 > 0$ such that $C_1\|x\| \le \|x\|' \le C_2\|x\|$ for all $x \in X$.

Theorem 20.6 Let X be a finite-dimensional vector space. Then all norms on X are equivalent.

Proof. We may assume that $X = \mathbb{K}^n$, where $n = \dim(X)$, since this is invariant under isomorphism. Indeed, if $\phi : X \to \mathbb{K}^n$ is an isomorphism and $\|\cdot\|$ is a norm on X, then $\|\phi^{-1}(\cdot)\|$ defines a norm on $\mathbb{K}^n$. Let $e_1, \ldots, e_n$ be the standard basis for $\mathbb{K}^n$, and let $\|\cdot\|$ be a norm on $\mathbb{K}^n$. We will show that $\|\cdot\|$ is equivalent to $\|\cdot\|_2$. For $x \in \mathbb{K}^n$, we have

$$\|x\| = \|\sum_{i=1}^n x_i e_i\| \leq \sum_{i=1}^n |x_i|\|e_i\| \leq \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} \left(\sum_{i=1}^n \|e_i\|^2\right)^{1/2} = \|x\|_2 \cdot C.$$

Thus $\|x\| \leq C\|x\|_2$ for all $x \in \mathbb{K}^n$.

Moreover, $|\|x\| - \|y\|| \leq \|x - y\| \leq C\|x - y\|_2$, so $x \mapsto \|x\|$ is continuous with respect to $\|\cdot\|_2$. Let $S = \{x \in \mathbb{K}^n : \|x\|_2 = 1\}$ be the unit sphere. Then S is closed and bounded with respect to $\|\cdot\|_2$, and hence compact by the Heine–Borel theorem. By the extreme value theorem, $x \mapsto \|x\|$ attains its minimum on S. Since $\|\cdot\|$ is positive definite, this minimum m is positive. Then $\|x\| \geq m\|x\|_2$, since $x/\|x\|_2 \in S$ for $x \neq 0$. Therefore $m\|x\|_2 \leq \|x\| \leq C\|x\|_2$. $\square$

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Recall that a linear map $T : X \to Y$ is **bounded** if

$$\|T\| = \sup\{\|Tx\|_Y : \|x\|_X \leq 1\} < \infty.$$

Corollary 20.7 Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces, with $\dim(X) < \infty$. If $T : X \to Y$ is linear, then T is bounded.

Proof. Define the **graph norm** on X, $\|\cdot\|_T$, by $\|x\|_T = \|x\|_X + \|Tx\|_Y$. We show that this is a norm. Nonnegativity is clear, and if $\|x\|_T = 0$, then $\|x\|_X = 0$ and so $x = 0$. Homogeneity follows from linearity of T, and

$$\|x + y\|_T = \|x + y\|_X + \|T(x + y)\|_Y \leq (\|x\|_X + \|y\|_X) + (\|Tx\|_Y + \|Ty\|_Y) = \|x\|_T + \|y\|_T.$$

Then $\|\cdot\|_T$ and $\|\cdot\|_X$ are equivalent since $\dim(X) < \infty$. Therefore, there exists $\mu > 0$ such that $\|x\|_T \leq \mu\|x\|_X$ for all $x \in X$. It follows that

$$\|Tx\|_Y \leq \|x\|_T \leq \mu\|x\|_X,$$

so T is bounded. $\square$

Lecture 21 — Wednesday, October 29, 2014

4. Completeness

Complete Metric Spaces

Definition 21.1 A metric space (X, d) is said to be **complete** if every Cauchy sequence in X converges to an element in X. A subset $A \subset X$ is said to be complete if $(A, d\big|_A)$ is complete.

Example 21.2

1. $(\mathbb{R}^n, d)$ where d is the metric induced by the 2-norm is complete (as seen in previous lectures and proved using the Bolzano–Weierstrass theorem).
2. Let (X, d) be a discrete metric space. Then (X, d) is complete. Indeed, every Cauchy sequence in X is eventually constant, and hence converges.
3. The open interval $(0, 1)$ is not complete. Indeed, $(1/n)_{n\geq 2}$ is Cauchy, but does not converge to an element in $(0, 1)$.

The failure in the last example is shown in Figure 8. The terms crowd together near 0, so the sequence is Cauchy, but its only possible limit is the missing endpoint.

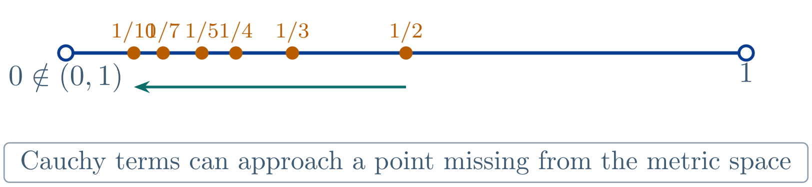


FIGURE 8

Proposition 21.3 Let (X, d) be a metric space and $A \subset X$.

1. If A is complete, then A is closed in X.
2. If X is complete and A is closed in X, then A is complete.

Proof. For (1), assume that A is complete and let $(a_n)_n$ be a sequence in A such that $a_n \to x$ for some $x \in X$. Since every convergent sequence is Cauchy, $(a_n)_n$ is Cauchy in A. By completeness of A, there exists $a \in A$ such that $a_n \to a$. Limits in metric spaces are unique, so $x = a$. Hence $x \in A$, and therefore A is closed in X.

For (2), assume that X is complete and that A is closed in X. Let $(a_n)_n$ be a Cauchy sequence in A. Since X is complete, there exists $x \in X$ such that $a_n \to x$ in X. Because A is closed and each $a_n \in A$, we must have $x \in A$. Thus $(a_n)_n$ converges in A, so A is complete. $\square$

Define $f : \mathbb{R} \to (-1, 1)$ by $x \mapsto x/(1 + |x|)$. Then f is a continuous bijection, and f^{-1} is continuous, so f is a homeomorphism. But $\mathbb{R}$ is complete, and $(-1, 1)$ is not! This shows that completeness is a property of the metric, and not a topological property.

Theorem 21.4 (Cantor's intersection theorem) Let (X, d) be a metric space. The following are equivalent:

1. (X, d) is complete.
2. If
$$F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$$
is a sequence of nonempty closed subsets of X with $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$, then $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$.

Proof. (1) *implies* (2). Suppose that (X, d) is complete and (F_n) are as above. For $n \geq 1$, choose $x_n \in F_n$. Let $\varepsilon > 0$. Since $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$, there exists $N \geq 1$ such that $\operatorname{diam}(F_N) < \varepsilon$. If $n, m \geq N$, then $x_n \in F_n \subset F_N$ and $x_m \in F_m \subset F_N$. Then

$$d(x_n, x_m) \leq \operatorname{diam}(F_N) < \varepsilon.$$

Thus, (x_n) is a Cauchy sequence. By assumption, $x = \lim_{n\to\infty} x_n$ exists. Let $n \geq 1$. If $k \geq n$, then $x_k \in F_k \subset F_n$, so $x = \lim_{k\to\infty} x_k \in F_n$ as F_n is closed. Thus, $x \in \bigcap_{n=1}^{\infty} F_n$.

(2) *implies* (1). Assume that (2) holds, and let (x_n) be a Cauchy sequence in X. For $n \geq 1$, let $F_n = \overline{\{x_k \mid k \geq n\}}$. Then $F_n \neq \emptyset$, $F_n \supseteq F_{n+1}$ for $n \geq 1$, and each F_n is closed. Let $\varepsilon > 0$. Since (x_n) is Cauchy, there exists $N \geq 1$ such that $d(x_n, x_m) < \varepsilon$ whenever $n, m \geq N$. The identity $\operatorname{diam}(\overline{A}) = \operatorname{diam}(A)$ follows by approximating any two points of $\overline{A}$ with sequences from A and passing to the limit in the distance. Hence $\operatorname{diam}(F_N) \leq \varepsilon$. Since $F_n \subset F_N$ for $n \geq N$,

$$\operatorname{diam}(F_n) \leq \operatorname{diam}(F_N) \leq \varepsilon.$$

Therefore, $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$. By assumption, there exists $x \in \bigcap_{n=1}^{\infty} F_n$. Let $\varepsilon > 0$. Since $\operatorname{diam}(F_n) \to 0$, there exists some $N \geq 1$ such that $\operatorname{diam}(F_n) < \varepsilon$ whenever $n \geq N$. If $n \geq N$, $x_n \in F_n$ and $x \in F_n$, so

$$d(x_n, x) \leq \operatorname{diam}(F_n) < \varepsilon.$$

Thus, $\lim_{n\to\infty} x_n = x$. □

Example 21.5

1. Let $x \in \mathbb{R}$ and let $F_n = [n, \infty)$. Then each F_n is closed and $F_n \supset F_{n+1}$. But $\bigcap_{n=1}^{\infty} F_n = \emptyset$. Note that $\operatorname{diam}(F_n) \not\to 0$.
2. Let $X = \mathbb{R}$ and $F_n = (0, 1/n)$. Then $\bigcap_{n=1}^{\infty} F_n = \emptyset$. The issue is that the F_n are not closed.

Definition 21.6 Let $(X, \|\cdot\|)$ be a normed linear space. A series

$$\sum_{n=0}^{\infty} x_n$$

of elements in X is said to **converge** to $x \in X$ if

$$\lim_{N\to\infty} \sum_{n=0}^{N} x_n = x$$

(that is,

$$\lim_{N\to\infty} \left\| \sum_{n=0}^{N} x_n - x \right\| = 0$$

). When this happens, we write

$$x = \sum_{n=0}^{\infty} x_n.$$

The series

$$\sum_{n=0}^{\infty} x_n$$

is said to be **absolutely convergent** if

$$\sum_{n=0}^{\infty} \|x_n\| < \infty.$$

Lecture 22 — Wednesday, October 29, 2014

Theorem 22.1 Let $(X, \|\cdot\|)$ be a normed linear space. The following are equivalent:

1. X is complete.
2. Every absolutely convergent series converges.

Proof. (1) *implies* (2). Assume X is complete, and let

$$\sum_{n=0}^{\infty} x_n$$

be an absolutely convergent series. Let

$$y_n = \sum_{k=0}^{n} x_k.$$

Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that

$$\sum_{k=N}^{\infty} \|x_k\| < \varepsilon.$$

If $m \geq n \geq N$, then

$$\|y_m - y_n\| = \left\| \sum_{k=n+1}^{m} x_k \right\| \leq \sum_{k=n+1}^{m} \|x_k\| \leq \sum_{k=N}^{\infty} \|x_k\| < \varepsilon.$$

Thus the sequence of partial sums $(y_n)_n$ is Cauchy. Since X is complete, it converges, so the series $\sum_{n=0}^{\infty} x_n$ converges in X.

(2) *implies* (1). Assume (2) and let $(y_n)_n$ be a Cauchy sequence in X. Recursively choose a subsequence $(y_{n_k})_k$ such that $\|y_{n_k} - y_{n_{k-1}}\| < 2^{-k}$ for every $k \geq 1$. Let $x_0 = y_{n_0}$, and $x_k = y_{n_k} - y_{n_{k-1}}$ for $k \geq 1$. Then

$$\sum_{k=0}^{\infty} \|x_k\| \leq \|y_{n_0}\| + \sum_{k=1}^{\infty} 2^{-k} < \infty$$

since this is a geometric series. By assumption,

$$y = \lim_{N \to \infty} \sum_{k=0}^{N} x_k$$

exists. But

$$\sum_{k=0}^{N} x_k = y_{n_N}.$$

Thus $(y_n)_n$ admits a convergent subsequence. By Lemma 13.4, $(y_n)_n$ converges. □

Banach and Hilbert Spaces

Definition 22.2 A normed linear space $(X, \|\cdot\|)$ is said to be a **Banach space** if it is complete in the metric induced by the norm.

Definition 22.3 Let $(H, \langle\cdot,\cdot\rangle)$ be an inner product space, and let $\|\cdot\|$ be the norm induced by the inner product (that is, $\|x\| = \sqrt{\langle x, x\rangle}$). H is said to be a **Hilbert space** if $(H, \|\cdot\|)$ is a Banach space.

Example 22.4

1. $\mathbb{K}^n$ with the usual inner product (the dot product) is a Hilbert space ($\langle\cdot,\cdot\rangle$ induces $\|\cdot\|_2$).
2. If $\|\cdot\|$ is a norm on $\mathbb{K}^n$, then $(\mathbb{K}^n, \|\cdot\|)$ is a Banach space. Indeed, $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent, so they have the same Cauchy/convergent sequences.
3. Let $S \neq \emptyset$ be a set, and define
$$\ell_\infty(S) = \{f : S \to \mathbb{K} \mid \|f\|_\infty < \infty\}.$$
Pointwise addition and scalar multiplication make $\ell_\infty(S)$ a vector space, and $\|f\|_\infty = \sup_{s\in S} |f(s)|$ defines a norm on it.

Remark 22.5 Let $(f_n)_n$ be a sequence in $\ell_\infty(S)$ and let $f = \lim_{n\to\infty} f_n$ in $\|\cdot\|_\infty$. For $s \in S$, $|f_n(s) - f(s)| \leq \|f_n - f\|_\infty$, so $\lim_{n\to\infty} f_n(s) = f(s)$ for all $s \in S$ (we say that $f_n \to f$ **pointwise**). The converse is not true! Let $S = \mathbb{N}$ and for $n \in \mathbb{N}$, let $e_n \in \ell_\infty(\mathbb{N})$ be defined by
$$e_n(m) = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases}.$$
Then $\lim_{n\to\infty} e_n(m) = 0$ for all $m \in \mathbb{N}$, but $\|e_n\|_\infty = 1$, so $e_n \not\to 0$ in ℓ_∞.

Theorem 22.6 Let $S \neq \emptyset$. Then $(\ell_\infty(S), \|\cdot\|_\infty)$ is a Banach space.

Proof. We have to show that the space is complete. Let $(f_n)_n$ be a Cauchy sequence in $\ell_\infty(S)$. For $s \in S$, $|f_n(s) - f_m(s)| \leq \|f_n - f_m\|_\infty$. Then $(f_n(s))_n$ is Cauchy in $\mathbb{K}$. Since $\mathbb{K}$ is complete, $(f_n(s))_n$ converges, so we may define a function $f : S \to \mathbb{K}$ by $f(s) = \lim_{n\to\infty} f_n(s)$. We claim that $f \in \ell_\infty(S)$ and that $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$. Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $\|f_n - f_m\|_\infty < \frac{\varepsilon}{2}$ for $n, m \geq N$. In particular, for any s, $|f_n(s) - f_m(s)| \leq \|f_n - f_m\|_\infty < \varepsilon/2$ for $n, m \geq N$. Taking the limit as $m \to \infty$, we have
$$|f_n(s) - f(s)| \leq \varepsilon/2 \text{ for all } n \geq N, s \in S. \tag{22.1}$$

This shows that $(f_n - f) \in \ell_\infty(S)$, so $f = (f - f_n) + f_n \in \ell_\infty(S)$. Taking the supremum over the $s \in S$ in (22.1), we deduce that $\|f_n - f\|_\infty \leq \varepsilon/2 < \varepsilon$ for all $n \geq N$. By definition, $f_n \to f$ in

$\|\cdot\|_\infty$. Hence $(\ell_\infty(S), \|\cdot\|_\infty)$ is complete. □

Lecture 23 — Friday, October 31, 2014

Theorem 23.1 Let $1 \le p < \infty$. Then $(\ell_p, \|\cdot\|_p)$ is a Banach space. In particular, $(\ell_2, \|\cdot\|_2)$ is a Hilbert space, where $\langle \cdot, \cdot \rangle$ is the usual inner product.

Proof. Let $(x^{(n)})_n$ be a Cauchy sequence in ℓ_p. Write $x^{(n)} = (x_k^{(n)})_{k\in\mathbb{N}}$. Since $|x_k^{(n)} - x_k^{(m)}| \le \|x^{(n)} - x^{(m)}\|_p$ for $k \in \mathbb{N}$, the sequence $(x_k^{(n)})_n$ is Cauchy in $\mathbb{K}$ for every $k \in \mathbb{N}$. Since $\mathbb{K}$ is complete, we may define $x_k = \lim_{n\to\infty} x_k^{(n)}$. Let $x = (x_k)_{k\in\mathbb{N}}$. We claim that $x \in \ell_p$ and that $\lim_{n\to\infty} \|x - x^{(n)}\|_p = 0$. Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $\|x^{(n)} - x^{(m)}\|_p < \varepsilon/2$ whenever $n, m \ge N$. Then for every $j \in \mathbb{N}$,

$$\sum_{k=0}^{j} |x_k^{(n)} - x_k^{(m)}|^p < (\varepsilon/2)^p$$

whenever $n, m \ge N$. Taking the limit as $m \to \infty$, we see that

$$\sum_{k=0}^{j} |x_k^{(n)} - x_k|^p \le (\varepsilon/2)^p$$

for all $n \ge N$ and all $j \in \mathbb{N}$. Taking the limit as $j \to \infty$, it follows that

$$\sum_{k=0}^{\infty} |x_k^{(n)} - x_k|^p \le (\varepsilon/2)^p$$

whenever $n \ge N$. Hence $x^{(n)} - x \in \ell_p$, and so

$$x = -(x^{(n)} - x) + x^{(n)} \in \ell_p.$$

Moreover,

$$\|x^{(n)} - x\|_p \le \frac{\varepsilon}{2} < \varepsilon$$

for all $n \ge N$. Thus $\lim_{n\to\infty} \|x - x^{(n)}\|_p = 0$, and so $(\ell_p, \|\cdot\|_p)$ is complete. □

Definition 23.2 Let S be a set and (X,d) be a metric space. Let $(f_n)_n$ be a sequence of functions from $S \to X$ and let $f : S \to X$. We say that $(f_n)_n$ converges to f **pointwise** if $\lim_{n\to\infty} f_n(s) = f(s)$ for all $s \in S$. We say that $(f_n)_n$ converges to f **uniformly** if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(f_n(s), f(s)) < \varepsilon$ whenever $n \geq N$, for all $s \in S$.

Example 23.3

1. Uniform convergence implies pointwise convergence, but not vice versa by Remark 22.5.
2. If $(f_n)_n$ is a sequence in $\ell_\infty(S)$ and if $f \in \ell_\infty(S)$, then $f_n \to f$ uniformly $\iff$ $\|f_n - f\|_\infty \to 0$.

Theorem 23.4 Let (X,τ) be a topological space and let (Y,d) be a metric space. If $(f_n)_n$ is a sequence of continuous functions $f_n : X \to Y$ which converges uniformly to $f : X \to Y$, then f is continuous.

Proof. Let $\varepsilon > 0$. Since $f_n \to f$ uniformly, there exists $N \in \mathbb{N}$ such that for all $x \in X$, $d(f_n(x), f(x)) < \varepsilon/3$ whenever $n \geq N$. Let $x_0 \in X$. Since f_N is continuous at x_0, there exists a neighborhood U of x_0 such that $d(f_N(x), f_N(x_0)) < \varepsilon/3$ for all $x \in U$. If $x \in U$, then

$$d(f(x), f(x_0)) \leq d(f(x), f_N(x)) + d(f_N(x), f_N(x_0)) + d(f_N(x_0), f(x_0)) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Thus $f^{-1}(B_\varepsilon(f(x_0)))$ is a neighborhood of x_0. Thus, f is continuous at x_0. $\square$

Example 23.5 Let $X = [0,1]$ with the Euclidean topology and $Y = \mathbb{R}$ with the Euclidean metric. Define $f_n : [0,1] \to \mathbb{R}$ by $f_n(x) = x^n$. Then

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in [0,1) \\ 1, & x = 1 \end{cases}.$$

Thus the pointwise limit of $(f_n)_n$ is not continuous. In particular, $(f_n)_n$ does not converge uniformly.

Let (X,τ) be a topological space, and let

$$C_b(X) = \{f : X \to \mathbb{K} \mid f \text{ is continuous and bounded}\}.$$

Then $C_b(X) \subset \ell_\infty(X)$, so $\|\cdot\|_\infty$ is a norm on $C_b(X)$. If X is compact, then

$$C_b(X) = \{f : X \to \mathbb{K} \mid f \text{ is continuous}\}.$$

Theorem 23.6 If (X, τ) is a topological space, then $(C_b(X), \|\cdot\|_\infty)$ is a Banach space.

Proof. Since $(\ell_\infty(X), \|\cdot\|_\infty)$ is a Banach space by Theorem 22.6, it suffices by Proposition 21.3(2) to show that $C_b(X)$ is closed in $\ell_\infty(X)$. Let $(f_n)_n$ be a sequence in $C_b(X)$ with $f = \lim_{n\to\infty} f_n \in \ell_\infty(X)$. Then $f_n \to f$ uniformly, so f is continuous by Theorem 23.4. Thus $C_b(X)$ is closed in $\ell_\infty(X)$. □

The Relationship Between Compactness and Completeness

Proposition 23.7 Every compact metric space is complete.

Proof. Let (X, d) be compact. Then X is sequentially compact by Theorem 20.1. If $(x_n)_n$ is a Cauchy sequence in X, then $(x_n)_n$ admits a convergent subsequence. Therefore $(x_n)_n$ converges by Lemma 13.4(3). □

Lecture 24 — Monday, November 03, 2014

Recall that every compact metric space is complete.

Proposition 24.1 Let (X, d) be a totally bounded metric space. Then every sequence $(x_n)_n$ in X admits a Cauchy subsequence.

Proof. Since X is totally bounded, it can be covered by finitely many balls of radius 1. Then one of these balls, B_1, contains infinitely many x_n. Let $S_1 = \{n \in \mathbb{N} \mid x_n \in B_1\}$. Similarly, X is covered by finitely many balls of radius $1/2$, so one of them, B_2, contains infinitely many x_n, where $n \in S_1$. Let $S_2 = \{n \in \mathbb{N} \mid x_n \in B_1 \cap B_2\}$. Then $|S_2| = \infty$.

Recursively, we thus obtain a sequence $(B_k)_{k=1}^{\infty}$ of balls, where B_k has radius $1/k$, such that $S_k = \{n \in \mathbb{N} \mid x_n \in B_1 \cap \cdots \cap B_k\}$ is infinite. Thus we can find a sequence $n_1 < n_2 < \ldots$ such that $x_{n_k} \in B_k$ for all $k \in \mathbb{N}, k \geq 1$. Since $\operatorname{diam}(B_k) \to 0$ and since $x_{n_k}, x_{n_j} \in B_K$ if $k, j \geq K$, it follows that (x_{n_k}) is Cauchy. □

We now have a generalization of the Heine–Borel theorem for arbitrary metric spaces.

Theorem 24.2 Let (X, d) be a metric space. The following are equivalent:

1. X is compact.
2. X is complete and totally bounded.

Proof. (1) *implies* (2). Let (X, d) be compact. By Proposition 23.7, X is complete. By Theorem 20.1, X is sequentially compact, and hence X is totally bounded by Proposition 19.3.

(2) *implies* (1). Let X be complete and totally bounded. By Theorem 20.1, it suffices to show that X is sequentially compact. Let $(x_n)_n$ be a sequence in X. By Proposition 24.1, $(x_n)_n$ admits a Cauchy subsequence $(x_{n_k})_k$. Since X is complete, $(x_{n_k})_k$ converges. Thus X is sequentially compact. □

Completions of Metric Spaces

Definition 24.3 Let $(X, d_X), (Y, d_Y)$ be metric spaces. A map $f : X \to Y$ is said to be an **isometry** if it preserves distances; that is, $d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2)$ for all $x_1, x_2 \in X$.

Isometries are always injective. If $f : X \to Y$ is a surjective isometry, then we may identify X and Y as metric spaces.

Definition 24.4 Let (X, d) be a metric space. A **completion** of X is a pair $((X^*, d^*), \rho)$ where (X^*, d^*) is a complete metric space, $\rho : X \to X^*$ is an isometry, and $\overline{\rho(X)} = X^*$.

Example 24.5 Let $\mathbb{Q}$ and $\mathbb{R}$ be equipped with the Euclidean metric, and let $\rho : \mathbb{Q} \to \mathbb{R}$ be the inclusion map. Then $\mathbb{R}$ is complete, $\rho(\mathbb{Q})$ is dense in $\mathbb{R}$, and ρ is an isometry. Thus $\mathbb{R}$ is a completion of $\mathbb{Q}$.

Lemma 24.6 Let (X, d) be a metric space, and let $A \subset X$ be dense. If every Cauchy sequence in A has a limit in X, then X is complete.

Proof. Let $(x_n)_n$ be a Cauchy sequence in X. For any $n \in \mathbb{N}$, choose $a_n \in A$ such that $d(x_n, a_n) < 1/(n+1)$. We claim that (a_n) is Cauchy. Let $\varepsilon > 0$. Since $(x_n)_n$ is Cauchy, there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon/3$ for $n, m \geq N$. Increasing N if necessary, we may also assume that $1/(N+1) < \varepsilon/3$. For $n, m \geq N$,

$$d(a_n, a_m) \leq d(a_n, x_n) + d(x_n, x_m) + d(x_m, a_m) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

By assumption, $x = \lim_{n\to\infty} a_n$ exists in X. Since

$$d(x_n, x) \leq d(x_n, a_n) + d(a_n, x) < \frac{1}{n+1} + d(a_n, x) \to 0,$$

we obtain $\lim_{n\to\infty} x_n = x$. $\square$

Theorem 24.7 Every metric space admits a completion.

Proof. Let (X, d) be a metric space. Let $\mathcal{C}[X]$ denote the set of all Cauchy sequences in X. Define a relation $\sim$ on $\mathcal{C}[X]$ by $(x_n)_n \sim (y_n)_n$ if $d(x_n, y_n) \to 0$ as $n \to \infty$. The metric axioms show that this is an equivalence relation. Define $X^* = \mathcal{C}[X]/\sim$. We want to define $d^* : X^* \times X^* \to \mathbb{R}$ by

$$d^*([(x_n)], [(y_n)]) = \lim_{n\to\infty} d(x_n, y_n).$$

Note that

$$|d(x_n, y_n) - d(x_m, y_m)| \leq |d(x_n, y_n) - d(x_m, y_n)| + |d(x_m, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_n, y_m).$$

Thus $(d(x_n, y_n))_n$ is Cauchy in $\mathbb{R}$, and hence $\lim_{n\to\infty} d(x_n, y_n)$ exists.

Lecture 25 — Wednesday, November 05, 2014

Proof continued. Let $(x_n) \sim (x'_n)$ and $(y_n) \sim (y'_n)$. Then $d(x_n, y_n) \leq d(x_n, x'_n) + d(x'_n, y'_n) + d(y'_n, y_n)$. Interchanging the primed and unprimed sequences gives the reverse inequality up to the same two error terms. Since both error terms tend to zero, $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x'_n, y'_n)$, and so d^* is well-defined. We claim that d^* is a metric on X^*. Indeed, d^* is nonnegative, and $d^*([(x_n)], [(y_n)]) = 0$ if and only if $\lim_{n\to\infty} d(x_n, y_n) = 0$; that is, $(x_n) \sim (y_n)$, so $[(x_n)] = [(y_n)]$. Also,

$$d^*([(x_n)], [(y_n)]) = \lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(y_n, x_n) = d^*([(y_n)], [(x_n)]),$$

so d^* is symmetric. For $[(z_n)] \in X^*$, the triangle inequality in X gives

$$d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n).$$

Taking limits yields

$$d^*([(x_n)], [(z_n)]) \leq d^*([(x_n)], [(y_n)]) + d^*([(y_n)], [(z_n)]).$$

Hence d^* is a metric.

Define $\rho : X \to X^*$, $x \mapsto [(x, x, x, \dots)]$. Then ρ is an isometry. Let $\alpha = [(x_n)] \in X^*$ and let $\varepsilon > 0$. Since (x_n) is Cauchy, there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon/2$ whenever $n, m \geq N$. Then

$$d^*(\rho(x_n), \alpha) = \lim_{m\to\infty} d(x_n, x_m) \leq \frac{\varepsilon}{2} < \varepsilon.$$

So $\rho(X)$ is dense in X^*. Finally, we show that (X^*, d^*) is complete. By Lemma 24.6, it suffices to show that every Cauchy sequence in $\rho(X)$ converges. Let $(\alpha_n)_n$ be a Cauchy sequence in $\rho(X)$. Say $\alpha_n = \rho(x_n)$. Since ρ is an isometry, (x_n) is a Cauchy sequence in X. Let $\alpha = [(x_n)] \in X^*$. Given $\varepsilon > 0$, let $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon/2$ when $n, m \geq N$. So for $n, m \geq N$,

$$d^*(\alpha_n, \alpha) = \lim_{m\to\infty} d(x_n, x_m) \leq \varepsilon/2 < \varepsilon.$$

So $\lim_{n\to\infty} \alpha_n = \alpha$ in X^*, hence X^* is complete. □

The main takeaway from this proof: construct Cauchy sequences and examine equivalence classes. This strategy works for solving many problems in life.

Theorem 25.1 Let (X, d) be a metric space, and assume that $((X_1^*, d_1^*), \rho_1)$ and $((X_2^*, d_2^*), \rho_2)$ are two completions of (X, d). Then there exists a unique surjective isometry $\phi : X_1^* \to X_2^*$ such that $\phi(\rho_1(x)) = \rho_2(x)$ for all $x \in X$. In other words, the diagram in Figure 9 commutes:

$$\begin{tikzcd}
X_1^* \arrow[rr, "\varphi"] & & X_2^* \\
& X \arrow[ul, "\rho_1"] \arrow[ur, "\rho_2"'] &
\end{tikzcd}$$

Figure 9

Proof. Uniqueness follows from the fact that $\rho_1(X)$ is dense in X_1^* and ϕ is assumed to be continuous. To show existence, define $\phi_0 : \rho_1(X) \to X_2^*$ by $\phi_0 = \rho_2 \circ \rho_1^{-1}$. Then ϕ_0 is an isometry; in particular, ϕ_0 is uniformly continuous.

We now construct an extension of ϕ_0 to all of X_1^*. Let $u \in X_1^*$. Since $\rho_1(X)$ is dense in X_1^*, there exists a sequence $(u_n)_n \subset \rho_1(X)$ with $u_n \to u$. Then $(u_n)_n$ is Cauchy in X_1^*. Because ϕ_0 is uniformly continuous, $(\phi_0(u_n))_n$ is Cauchy in X_2^*. Since X_2^* is complete, $(\phi_0(u_n))_n$ converges in X_2^*. Define

$$\phi(u) := \lim_{n\to\infty} \phi_0(u_n).$$

To see that this is well-defined, let $(v_n)_n \subset \rho_1(X)$ be another sequence with $v_n \to u$. Then $d_1^*(u_n, v_n) \to 0$. By uniform continuity of ϕ_0, this implies $d_2^*(\phi_0(u_n), \phi_0(v_n)) \to 0$. Hence the limits of $(\phi_0(u_n))_n$ and $(\phi_0(v_n))_n$ agree, so $\phi(u)$ does not depend on the approximating sequence.

Thus $\phi : X_1^* \to X_2^*$ is well-defined. If $u \in \rho_1(X)$, choose $u_n = u$ for all n; then $\phi(u) = \phi_0(u)$, so ϕ extends ϕ_0. We verify that ϕ is an isometry. Let $u, v \in X_1^*$, and choose $(u_n), (v_n) \subset \rho_1(X)$ with $u_n \to u$ and $v_n \to v$. Since ϕ_0 is an isometry on $\rho_1(X)$,

$$d_2^*(\phi_0(u_n), \phi_0(v_n)) = d_1^*(u_n, v_n) \quad \text{for all } n.$$

Taking limits and using continuity of the metric in each space,

$$d_2^*(\phi(u), \phi(v)) = \lim_{n\to\infty} d_2^*(\phi_0(u_n), \phi_0(v_n)) = \lim_{n\to\infty} d_1^*(u_n, v_n) = d_1^*(u, v).$$

Hence ϕ is an isometry. Moreover,

$$\phi(\rho_1(x)) = \phi_0(\rho_1(x)) = \rho_2(x)$$

for $x \in X$. Since X_1^* is complete and ϕ is an isometry, $\phi(X_1^*)$ is complete. Thus, $\phi(X_1^*)$ is closed in X_2^*. Moreover,

$$\text{Range}(\phi) \supseteq \text{Range}(\phi_0) = \text{Range}(\rho_2)$$

, and hence Range(ϕ) is dense in X_2^*. Thus, Range(ϕ) $= X_2^*$; that is, ϕ is surjective. ☐

The Baire Category Theorem

Example 25.2 (Thomae's function) Define $f : \mathbb{R} \to \mathbb{R}$ in the following way:

$$f(x) = \begin{cases} 1/n, & x = m/n \in \mathbb{Q}, n \in \mathbb{N}, n \geq 1, \gcd(m,n) = 1 \\ 0, & x \notin \mathbb{Q} \end{cases}$$

Where is this function continuous? It is continuous at every irrational number and discontinuous at every rational. Indeed, if $x \in \mathbb{Q}$, then there exists a sequence (x_n) in $\mathbb{R} \setminus \mathbb{Q}$ such that $x_n \to x$. Then

$$f(x_n) \to 0 \neq f(x).$$

On the other hand, if $x \in \mathbb{R} \setminus \mathbb{Q}$, let $\varepsilon > 0$ and let $N \in \mathbb{N}$ such that $1/N < \varepsilon$. In $[x-1, x+1]$, there are only finitely many rationals m/n where $n \leq N$. Thus, there exists $\delta > 0$ such that none of these points are contained in $(x-\delta, x+\delta)$. Hence

$$|f(y) - f(x)| = |f(y)| < \varepsilon$$

for all $y \in (x-\delta, x+\delta)$. Thus, f is continuous at x.

Let (X,d) be a metric space and let $f : X \to \mathbb{R}$ be a function. Define the **set of continuity points of f**

$$C(f) = \{x \in X \mid f \text{ is continuous at } x\}.$$

In Example 25.2, $C(f) = \mathbb{R} \setminus \mathbb{Q}$.

Question 25.3 Is there a function $f : \mathbb{R} \to \mathbb{R}$ such that $C(f) = \mathbb{Q}$?

Definition 25.4 Let (X,d) be a metric space with $A \subset X$. A is said to be a $\boldsymbol{G_\delta}$ **set** if there exists a countable collection of open sets $\{U_n\}_{n=0}^\infty$ such that $A = \bigcap_{n=0}^\infty U_n$. A is said to be an $\boldsymbol{F_\sigma}$ **set** if there exists a countable collection of closed sets $\{F_n\}_{n=0}^\infty$ such that $A = \bigcup_{n=0}^\infty F_n$.

Lecture 26 — Monday, November 10, 2014

Recall that (X, d) is a metric space. $A \subset X$ is G_δ if $A = \bigcap_{n=0}^{\infty} U_n$ with each U_n open. $A \subset X$ is F_σ if $A = \bigcup_{n=0}^{\infty} F_n$ with each F_n closed.

Example 26.1 The set $[0, 1]$ is both F_σ and G_δ. Indeed,

$$[0,1] = \bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right) = \bigcup_{n=1}^{\infty} [0,1].$$

Example 26.2

1. $A \subset X$ is a G_δ set if and only if $X \setminus A$ is an F_σ set.
2. Every closed subset of X is a G_δ set, hence every open subset of X is an F_σ set.

Proposition 26.3 Let (X, d) be a metric space and let $f : X \to \mathbb{R}$ be a function. Then $C(f)$ is a G_δ set.

Proof. For each $n \geq 1$, define

$$V_n = \left\{ x \in X : \exists r > 0 \text{ such that } |f(y) - f(z)| < \frac{1}{n} \text{ for all } y, z \in B_r(x) \right\}.$$

We first show that each V_n is open. Let $x \in V_n$. Choose $r > 0$ such that

$$|f(y) - f(z)| < \frac{1}{n} \quad \text{for all } y, z \in B_r(x).$$

If $x' \in B_{r/2}(x)$ and $y, z \in B_{r/2}(x')$, then

$$d(y, x) \leq d(y, x') + d(x', x) < \frac{r}{2} + \frac{r}{2} = r,$$

and similarly $d(z, x) < r$, so $y, z \in B_r(x)$. Hence $|f(y) - f(z)| < 1/n$ for all $y, z \in B_{r/2}(x')$, which shows $x' \in V_n$. Therefore $B_{r/2}(x) \subset V_n$, so V_n is open.

Next, we prove that

$$C(f) = \bigcap_{n=1}^{\infty} V_n.$$

If $x \in C(f)$, then f is continuous at x. Fix $n \geq 1$. Using continuity at x with $\varepsilon = 1/(2n)$, choose $r > 0$ such that

$$|f(w) - f(x)| < \frac{1}{2n} \quad \text{for all } w \in B_r(x).$$

Then for any $y, z \in B_r(x)$,

$$|f(y) - f(z)| \leq |f(y) - f(x)| + |f(x) - f(z)| < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n},$$

so $x \in V_n$. Since n was arbitrary, $x \in \bigcap_{n=1}^{\infty} V_n$.

Conversely, suppose $x \in \bigcap_{n=1}^{\infty} V_n$. Let $\varepsilon > 0$. Choose $n \geq 1$ with $1/n < \varepsilon$. Since $x \in V_n$, there exists $r > 0$ such that

$$|f(y) - f(z)| < \frac{1}{n} \quad \text{for all } y, z \in B_r(x).$$

Taking $z = x$, we get

$$|f(y) - f(x)| < \frac{1}{n} < \varepsilon \quad \text{for all } y \in B_r(x),$$

so f is continuous at x. Thus $x \in C(f)$.

Therefore $C(f) = \bigcap_{n=1}^{\infty} V_n$, where each V_n is open. Hence $C(f)$ is a G_δ set.

$\square$

Theorem 26.4 (Baire category theorem) Let (X, d) be a complete metric space, and let $\{U_n\}_{n=0}^{\infty}$ be a countable collection of open, dense subsets of X. Then $\bigcap_{n=0}^{\infty} U_n$ is dense in X.

Proof. Let $x \in X$ and $\varepsilon > 0$. Set $V = B_\varepsilon(x)$. We want to show that $V \cap \bigcap_{n=0}^{\infty} U_n \neq \emptyset$. Since U_0 is dense, choose $x_0 \in U_0 \cap V$. The set $U_0 \cap V$ is open, so we can choose $\varepsilon_0 \in (0, 1)$ such that

$$\overline{B}_{\varepsilon_0}(x_0) \subset U_0 \cap V.$$

Since U_1 is dense, choose $x_1 \in U_1 \cap B_{\varepsilon_0}(x_0)$. This intersection is open, so we can choose $\varepsilon_1 \in (0, 1/2)$ such that

$$\overline{B}_{\varepsilon_1}(x_1) \subset U_1 \cap B_{\varepsilon_0}(x_0).$$

Continuing recursively, we obtain closed balls $F_n = \overline{B}_{\varepsilon_n}(x_n)$ such that $\varepsilon_n < 1/(n+1)$ and

$$F_n \subset U_n \cap F_{n-1}$$

for every $n \geq 1$, with $F_0 \subset U_0 \cap V$. The F_n are nonempty and closed, they are nested, and $\operatorname{diam}(F_n) \leq 2\varepsilon_n \to 0$. By Cantor's intersection theorem, there exists $z \in \bigcap_{n=0}^{\infty} F_n$. Then $z \in V$ and $z \in U_n$ for every n, so $z \in V \cap \bigcap_{n=0}^{\infty} U_n$. □

Corollary 26.5 $\mathbb{Q}$ is not a G_δ set in $\mathbb{R}$. Hence, there is no $f : \mathbb{R} \to \mathbb{R}$ for which $C(f) = \mathbb{Q}$.

Proof. Assume for a contradiction that $\mathbb{Q}$ is a G_δ set in $\mathbb{R}$. That is, $\mathbb{Q} = \bigcap_{n=0}^{\infty} U_n$, where each U_n is open in $\mathbb{R}$. Since $\mathbb{Q} \subset U_n$, each U_n is dense. For $q \in \mathbb{Q}$, let $V_q = \mathbb{R} \setminus \{q\}$. Then each V_q is open dense. By the Baire category theorem, $\bigcap_{n=0}^{\infty} U_n \cap \bigcap_{q \in \mathbb{Q}} V_q$ is dense in $\mathbb{R}$. But

$$\left(\bigcap_{n=0}^{\infty} U_n \right) \cap \left(\bigcap_{q \in \mathbb{Q}} V_q \right) = \mathbb{Q} \cap (\mathbb{R} \setminus \mathbb{Q}) = \emptyset,$$

a contradiction. □

Definition 26.6 Let (X, d) be a metric space and let $A \subset X$. We say that A is **nowhere dense** in X if $\operatorname{int}(\overline{A}) = \emptyset$. We say A is of **first category in X** (or **meager in X**) if A is a countable union of nowhere dense sets. A is said to be of **second category in X** otherwise.

Since $X \setminus \operatorname{int}(\overline{A}) = \overline{X \setminus \overline{A}}$, it follows that A is nowhere dense if and only if $X \setminus \overline{A}$ is dense.

Example 26.7

1. $\mathbb{Z} \subset \mathbb{R}$ is nowhere dense.
2. $\mathbb{Q} \subset \mathbb{R}$ is not nowhere dense, as $\overline{\mathbb{Q}} = \mathbb{R}$. However, $\mathbb{Q}$ is of first category since $\mathbb{Q} = \bigcup_{q \in \mathbb{Q}} \{q\}$.

Corollary 26.8 (Baire category theorem: version 2) Let (X, d) be a complete metric space. Then X is of second category in itself. In particular, if $X = \bigcup_{n=0}^{\infty} F_n$ where each F_n is closed, then there exists n_0 such that $\operatorname{int}(F_{n_0}) \neq \emptyset$.

Proof. Suppose that $X = \bigcup_{n=0}^\infty A_n$, where each A_n is nowhere dense. Then $X = \bigcup_{n=0}^\infty \overline{A_n}$. Let $U_n = X \setminus \overline{A_n}$. Then each U_n is open and dense; hence, $\bigcap_{n=0}^\infty U_n \neq \emptyset$ by the Baire category theorem. This is a contradiction, because

$$\bigcap_{n=0}^\infty U_n = X \setminus \bigcup_{n=0}^\infty \overline{A_n} = \emptyset.$$

Thus X is of second category in itself. For the final assertion, if every closed set F_n had empty interior, then every F_n would be nowhere dense and $X = \bigcup_n F_n$ would be of first category, a contradiction. □

Theorem 26.9 (Uniform boundedness principle) Let $(X, \|\cdot\|)$ be a Banach space and let $(Y, \|\cdot\|)$ be a normed linear space. Let $\emptyset \neq \mathcal{F} \subset B(X,Y)$. If $\sup_{T\in\mathcal{F}} \|Tx\|_Y < \infty$ for every $x \in X$, then $\sup_{T\in\mathcal{F}} \|T\| < \infty$.

Proof. For $n \in \mathbb{N}$, let

$$F_n = \{x \in X \mid \|Tx\|_Y \leq n \text{ for all } T \in \mathcal{F}\}.$$

Since each $T \in \mathcal{F}$ is bounded and hence continuous, F_n is closed. Moreover, $X = \bigcup_{n=0}^\infty F_n$. By Corollary 26.8, $\operatorname{int}(F_{n_0}) \neq \emptyset$. Thus there exists $x_0 \in F_{n_0}$ and $\varepsilon > 0$ such that $B_\varepsilon(x_0) \subset F_{n_0}$. Let $T \in \mathcal{F}$ and let $z \in X$ with $\|z\| < \varepsilon$. Then $z = ((x_0 + z) - (x_0 - z))/2$, and thus

$$\begin{aligned}
\|Tz\|_Y &= \frac{1}{2}\|T(x_0+z) - T(x_0-z)\|_Y \\
&\leq \frac{1}{2}(\|T(x_0+z)\|_Y + \|T(x_0-z)\|_Y) \\
&\leq \frac{1}{2}(n_0 + n_0) \\
&= n_0.
\end{aligned}$$

For $u \in X$ with $\|u\| \leq 1$, take $z = (\varepsilon/2)u$. Then $\|Tu\|_Y \leq 2n_0/\varepsilon$ for every $T \in \mathcal{F}$, and therefore

$$\sup_{T\in\mathcal{F}} \|T\| \leq \frac{2n_0}{\varepsilon} < \infty.$$

□

Lecture 27 — Wednesday, November 12, 2014

5. Spaces of Continuous Functions

The Weierstrass Approximation Theorem

Lemma 27.1 For all $x \in [0,1]$ and $n \in \mathbb{N}^+$, we have $(1-x^2)^n \geq 1 - nx^2$.

Proof. Let $f(x) = (1-x^2)^n - (1-nx^2)$. Then $f(0) = 0$, and

$$f'(x) = -2xn(1-x^2)^{n-1} + 2nx = 2nx(1-(1-x^2)^{n-1}) \geq 0$$

for all $x \in [0,1]$. Thus, $f(x) \geq 0$ for all $x \in [0,1]$ by the mean value theorem. □

Theorem 27.2 (Weierstrass approximation theorem) Let $a, b \in \mathbb{R}$ with $a < b$, and let $f : [a,b] \to \mathbb{K}$ be continuous. Then there exists a sequence of polynomials $(p_n)_n$ such that $(p_n)_n$ converges uniformly to f on $[a,b]$.

Proof. We first reduce to the case where $[a,b] = [0,1]$. Let $\phi : [0,1] \to [a,b]$, $x \mapsto a + (b-a)x$. If f is continuous on $[a,b]$, then $f \circ \phi$ is continuous on $[0,1]$. If $(p_n)_n$ is a sequence of polynomials which converges to $f \circ \phi$ uniformly on $[0,1]$, then $(p_n \circ \phi^{-1})_n$ is a sequence of polynomials which converges to f uniformly on $[a,b]$. Thus, we may assume that $[a,b] = [0,1]$.

Next, by replacing f with $g : [0,1] \to \mathbb{K}$ defined by

$$x \mapsto f(x) - f(0) - x(f(1) - f(0)),$$

we can assume that $f(0) = f(1) = 0$. By defining $f(x) = 0$ if $x \notin [0,1]$, we may extend f to a uniformly continuous function on $\mathbb{R}$. For $n \in \mathbb{N}^+$, define $Q_n(x) = c_n(1-x^2)^n$, where c_n is chosen so that

$$\int_{-1}^{1} Q_n(x)\,\mathrm{d}x = 1$$

for all $n \in \mathbb{N}^+$. We require an estimate for c_n. By Lemma 27.1, $(1-x^2)^n \geq 1 - nx^2$. Then

$$\begin{aligned}
\int_{-1}^{1} (1-x^2)^n \,\mathrm{d}x &= 2\int_0^1 (1-x^2)^n \,\mathrm{d}x \\
&\geq 2\int_0^{1/\sqrt{n}} (1-nx^2)\,\mathrm{d}x \\
&= 2\left(\frac{1}{\sqrt{n}} - \frac{n}{3}\frac{1}{n\sqrt{n}}\right) \\
&= \frac{2}{\sqrt{n}} - \frac{2}{3\sqrt{n}} \\
&= \frac{4}{3\sqrt{n}} \\
&\geq \frac{1}{\sqrt{n}}.
\end{aligned}$$

Thus $c_n \leq \sqrt{n}$. In particular, if $0 < \delta \leq |x| \leq 1$, then $Q_n(x) \leq \sqrt{n}(1-\delta^2)^n$. Let

$$p_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)\,\mathrm{d}t.$$

By a change of variables,

$$p_n(x) = \int_{x-1}^{x+1} f(u)Q_n(u-x)\,\mathrm{d}u = \int_0^1 f(u)Q_n(u-x)\,\mathrm{d}u$$

for $x \in [0,1]$, because f vanishes outside $[0,1]$. This is a polynomial in x, as $Q_n(u-x)$ is a polynomial in x. Let $M = \|f\|_\infty$ and let $\varepsilon > 0$. Since f is uniformly continuous, there exists $0 < \delta \leq 1$ such that $|f(x) - f(y)| < \varepsilon/2$ whenever $|x-y| < \delta$. Since

$$\int_{-1}^{1} Q_n(t)\,\mathrm{d}t = 1,$$

we deduce that

$$\begin{aligned}
|p_n(x) - f(x)| &= \left|\int_{-1}^{1} (f(x+t) - f(x))Q_n(t)\,\mathrm{d}t\right| \\
&\leq \int_{-1}^{1} |f(x+t) - f(x)|Q_n(t)\,\mathrm{d}t \\
&= \int_{-\delta}^{\delta} |f(x+t) - f(x)|Q_n(t)\,\mathrm{d}t + \int_{-1}^{-\delta} |f(x+t) - f(x)|Q_n(t)\,\mathrm{d}t \\
&\quad + \int_{\delta}^{1} |f(x+t) - f(x)|Q_n(t)\,\mathrm{d}t
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{\varepsilon}{2}\int_{-1}^{1} Q_n(t)\,\mathrm{d}t + 2M\int_{|t|\geq\delta} Q_n(t)\,\mathrm{d}t \\
&\leq \frac{\varepsilon}{2} + 4M\sqrt{n}(1-\delta^2)^n \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

for n sufficiently large. Thus, (p_n) converges to f uniformly on $[0,1]$. (The function Q_n is called an **approximate identity**.) $\square$

Corollary 27.3 If $a, b \in \mathbb{R}$ with $a < b$, then the space $(C([a,b],\mathbb{K}), \|\cdot\|_\infty)$ is separable.

Proof. Let

$$\mathcal{P}_{\mathbb{Q}} = \begin{cases} \{\text{polynomials with coefficients in } \mathbb{Q}\}, & \mathbb{K} = \mathbb{R}, \\ \{\text{polynomials with coefficients in } \mathbb{Q} + i\mathbb{Q}\}, & \mathbb{K} = \mathbb{C}. \end{cases}$$

Then $\mathcal{P}_{\mathbb{Q}}$ is countable, because it is a countable union (over the degree) of finite Cartesian products of a countable set. We claim that $\mathcal{P}_{\mathbb{Q}}$ is dense in $(C([a,b],\mathbb{K}), \|\cdot\|_\infty)$. Let $f \in C([a,b],\mathbb{K})$ and let $\varepsilon > 0$. By the Weierstrass approximation theorem, there exists a polynomial

$$p(x) = \sum_{j=0}^{m} c_j x^j$$

with coefficients in $\mathbb{K}$ such that

$$\|f - p\|_\infty < \frac{\varepsilon}{2}.$$

Set $R = \max\{|a|, |b|\}$. Choose coefficients d_j in $\mathbb{Q}$ (if $\mathbb{K} = \mathbb{R}$) or in $\mathbb{Q} + i\mathbb{Q}$ (if $\mathbb{K} = \mathbb{C}$) such that

$$|c_j - d_j| < \frac{\varepsilon}{2(m+1)(1+R)^j}, \quad 0 \leq j \leq m.$$

Define

$$q(x) = \sum_{j=0}^{m} d_j x^j \in \mathcal{P}_{\mathbb{Q}}.$$

For $x \in [a,b]$, we have $|x| \leq R$, so

$$|p(x) - q(x)| \leq \sum_{j=0}^{m} |c_j - d_j|\,|x|^j \leq \sum_{j=0}^{m} \frac{\varepsilon}{2(m+1)(1+R)^j} R^j \leq \sum_{j=0}^{m} \frac{\varepsilon}{2(m+1)} = \frac{\varepsilon}{2}.$$

Hence $\|p-q\|_\infty \le \varepsilon/2$, and therefore

$$\|f-q\|_\infty \le \|f-p\|_\infty + \|p-q\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $\mathcal{P}_{\mathbb{Q}}$ is countable and dense in $(C([a,b],\mathbb{K}), \|\cdot\|_\infty)$, so this space is separable. $\square$

The Stone–Weierstrass Theorem

Definition 27.4 Let X be a set and let $M \subset \mathbb{K}^X$. We say that M **separates points** of X if whenever $x, y \in X$ with $x \neq y$, there exists $f \in M$ such that $f(x) \neq f(y)$.

Example 27.5

1. The set of all real polynomials $\mathbb{R}[x]$ separates the points of $\mathbb{R}$. In fact, the polynomial $f(x) = x$ separates points.
2. Let (X, d) be a metric space. Then the set

$$C(X, \mathbb{K}) = \{f : X \to \mathbb{K} \mid f \text{ is continuous}\}$$

separates the points of X. Indeed, if $a, b \in X$ with $a \neq b$, then $f(x) = d(x, a)$ is continuous and $f(a) = 0 \neq f(b)$.

Lecture 28 — Friday, November 14, 2014

Let (X, τ) be a topological space and let $f, g \in C(X, \mathbb{R})$. Define

$$(f \wedge g)(x) := \min(f(x), g(x)) \qquad \text{and} \qquad (f \vee g)(x) := \max(f(x), g(x)).$$

Since

$$(f \wedge g)(x) = \frac{f(x) + g(x) - |f(x) - g(x)|}{2} \qquad \text{and} \qquad (f \vee g)(x) = \frac{f(x) + g(x) + |f(x) - g(x)|}{2}$$

for all $x \in X$, we have $f \vee g, f \wedge g \in C(X, \mathbb{R})$.

Definition 28.1 Let (X, τ) be a topological space. A subspace $L \subset C(X, \mathbb{R})$ is said to be a **lattice** if $f \wedge g \in L$ and $f \vee g \in L$ whenever $f, g \in L$.

Remark 28.2

1. Since $f \wedge g = -((-f) \vee (-g))$, it suffices to demand that $f \vee g \in L$ whenever $f, g \in L$.
2. The equations for $f \wedge g$ and $f \vee g$ show that if $L \subset C(X, \mathbb{R})$ is a subspace such that $f \in L \implies |f| \in L$, then L is a lattice.

Example 28.3 A function $f : [a, b] \to \mathbb{R}$ is said to be **piecewise linear** if there exists a partition

$$a = t_0 < t_1 < \cdots < t_n = b$$

such that $f|_{[t_{i-1}, t_i]}$ is affine linear for $1 \leq i \leq n$. Then

$$\{f \in C([a, b], \mathbb{R}) : f \text{ is piecewise linear}\}$$

is a lattice.

Theorem 28.4 (Stone–Weierstrass theorem) Let (X, τ) be a compact Hausdorff space. Suppose that $L \subset C(X, \mathbb{R})$ is a subspace such that the following hold:

1. The constant function 1 is in L.
2. L separates points.
3. L is a lattice.

Then L is dense in $(C(X, \mathbb{R}), \|\cdot\|_\infty)$.

Proof. Let $h \in C(X, \mathbb{R})$ and let $\varepsilon > 0$. We want to show that there exists $f \in L$ such that $\|f - h\|_\infty < \varepsilon$ (that is, L is dense).

Step 1: We show that for any two points $s, t \in X$ with $s \neq t$, there exists $f_{s,t} \in L$ such that $f_{s,t}(s) = h(s)$ and $f_{s,t}(t) = h(t)$. Since L separates points, there exists $g \in L$ such that $g(s) \neq g(t)$.

Define

$$f_{s,t}(x) = h(t) + (h(s) - h(t))\frac{g(x) - g(t)}{g(s) - g(t)}$$

for $x \in X$. Then $f_{s,t} \in L$, since $1 \in L$ and L is a subspace, and the construction gives the required values at s and t.

Step 2: Let $t \in X$. Define $f_{t,t}$ to be the constant function $h(t)$, which belongs to L because $1 \in L$. We show that there exists $g_t \in L$ such that $g_t(t) = h(t)$ and $g_t(x) > h(x) - \varepsilon$ for all $x \in X$. For any $s \in X$, the function $f_{s,t} - h$ is continuous and $f_{s,t}(s) = h(s)$. Thus there exists a neighborhood U_s of s such that

$$h(x) - \varepsilon < f_{s,t}(x) < h(x) + \varepsilon$$

for all $x \in U_s$. Since X is compact, the open cover $\{U_s\}_{s \in X}$ admits a finite subcover $\{U_{s_1}, \dots, U_{s_k}\}$. Let $g_t = f_{s_1,t} \vee \cdots \vee f_{s_k,t} \in L$. Then $g_t(t) = \max\{f_{s_i,t}(t)\} = h(t)$. If $x \in X$, then $x \in U_{s_j}$ for some $j \in \{1, \dots, k\}$. Hence

$$g_t(x) \geq f_{s_j,t}(x) > h(x) - \varepsilon.$$

Step 3: We show that there exists $f \in L$ with $\|f - h\|_\infty < \varepsilon$. For $t \in X$, the function $g_t - h$ is continuous and $g_t(t) = h(t)$. Thus there exists a neighborhood V_t of t such that

$$h(x) - \varepsilon < g_t(x) < h(x) + \varepsilon$$

for all $x \in V_t$. Since X is compact, the open cover $\{V_t\}_{t \in X}$ admits a finite subcover $\{V_{t_1}, \dots, V_{t_n}\}$. Let $f = g_{t_1} \wedge \cdots \wedge g_{t_n} \in L$. If $x \in X$, then

$$f(x) = \min\{g_{t_i}(x) : 1 \leq i \leq n\} > h(x) - \varepsilon.$$

If $x \in V_{t_j}$ for some $j \in \{1, \dots, n\}$, then

$$f(x) \leq g_{t_j}(x) < h(x) + \varepsilon.$$

Thus $\|f - h\|_\infty < \varepsilon$. □

Corollary 28.5 Let $L = \{f \in C([a,b], \mathbb{R}) \mid f \text{ is piecewise linear}\}$. Then L is dense in $(C([a,b], \mathbb{R}), \|\cdot\|_\infty)$.

Proof. The space L is a lattice, contains 1, and separates points because it contains the identity function. It is therefore dense by the Stone–Weierstrass theorem. □

Lecture 29 — Friday, November 14, 2014

Definition 29.1 Let (X,τ) be a topological space. A subspace $A \subset C(X,\mathbb{R})$ is said to be a **subalgebra** if $fg \in A$ whenever $f, g \in A$.

Example 29.2 Let $K \subset \mathbb{R}^n$ be compact. Then the set of all real-valued polynomial functions on K is a subalgebra of $C(K,\mathbb{R})$.

Theorem 29.3 (Stone–Weierstrass theorem: real version) Let (X,τ) be a compact Hausdorff space and let $A \subset C(X,\mathbb{R})$ be a subspace such that the following hold:

1. A is a subalgebra.
2. A separates the points of X.
3. A contains the constant function 1.

Then A is dense in $(C(X,\mathbb{R}), \|\cdot\|_\infty)$.

Proof. We first show that $\overline{A} \subset C(X,\mathbb{R})$ is a subalgebra. If $f, g \in \overline{A}$, then there are sequences $(f_n)_n$ in A and $(g_n)_n$ in A such that $f_n \to f$ and $g_n \to g$ with respect to $\|\cdot\|_\infty$. Then $f_n + g_n \to f + g \in \overline{A}$. Thus $\overline{A}$ is a subspace. Moreover,

$$\|f_n g_n - fg\|_\infty = \|f_n g_n - f_n g + f_n g - fg\|_\infty \leq \|f_n\|_\infty \|g_n - g\|_\infty + \|g\|_\infty \|f_n - f\|_\infty \to 0$$

as $\|f_n\|_\infty$ is bounded. Thus, $fg \in \overline{A}$. Hence, $\overline{A}$ is a subalgebra of $C(X,\mathbb{R})$. Moreover, it separates points because A does, and contains the constant 1.

We claim that $\overline{A}$ is a lattice. Let $f \in \overline{A}$. We must show that $|f| \in \overline{A}$. Let $M = \|f\|_\infty$. If $M = 0$, then $f = 0$ and the claim is immediate. Assume $M > 0$, and let $\varepsilon > 0$. By the Weierstrass approximation theorem, there exists a polynomial $p \in \mathbb{R}[x]$ such that $|p(x) - |x|| < \varepsilon$ for all $x \in [-M, M]$. Then $\|p(f) - |f|\|_\infty < \varepsilon$. Since $\overline{A}$ is a subalgebra containing 1, we have $p(f) \in \overline{A}$. Since $\varepsilon > 0$ was arbitrary, $|f| \in \overline{A}$. By the original lattice version of the Stone–Weierstrass theorem, $\overline{A} = C(X,\mathbb{R})$. $\square$

Corollary 29.4 Let $K \subset \mathbb{R}^n$ be compact and let $f : K \to \mathbb{R}$ be continuous. Then there exists a sequence of real-valued polynomials $(p_n)_n$ which converges to f uniformly on K.

Proof. The set of all real-valued polynomial functions on K is a subalgebra and contains 1. It also separates points: if two points of K differ, then one of the coordinate functions $x \mapsto x_j$ distinguishes them. The Stone–Weierstrass theorem now gives the conclusion. □

Example 29.5 Let $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ and let A be the set of all functions on $\mathbb{T}$ of the form

$$f(z) = \sum_{k=0}^{n} c_k z^k,$$

where $n \in \mathbb{N}$ and $c_k \in \mathbb{C}$ for $0 \le k \le n$. Then A is a subalgebra of $C(\mathbb{T}, \mathbb{C})$, $1 \in A$, and A separates points. Note that if

$$f = \sum_{k=0}^{n} c_k z^k \in A,$$

then

$$\int_0^{2\pi} f(e^{it}) e^{it} \, \mathrm{d}t = \sum_{k=0}^{n} c_k \int_0^{2\pi} e^{i(k+1)t} \, \mathrm{d}t = 0,$$

because each integral in the sum vanishes. Since every $f \in \overline{A}$ is a uniform limit of functions of A,

$$\int_0^{2\pi} f(e^{it}) e^{it} \, \mathrm{d}t = 0$$

for all $f \in \overline{A}$. Let $g(z) = \bar{z}$, the complex conjugate. Then g is certainly continuous, and

$$\int_0^{2\pi} g(e^{it}) e^{it} \, \mathrm{d}t = \int_0^{2\pi} e^{-it} e^{it} \, \mathrm{d}t = \int_0^{2\pi} \mathrm{d}t = 2\pi \neq 0.$$

So $g \notin \overline{A}$, so A is not dense in $C(\mathbb{T}, \mathbb{C})$.

Definition 29.6 Let (X, τ) be a topological space and let $S \subset C(X, \mathbb{C})$ be a subspace. We say that S is **self-adjoint** if $f \in S$ implies $\bar{f} \in S$, where $\bar{f}(x) = \overline{f(x)}$ is the complex conjugate function.

If S is a self-adjoint subspace of $C(X, \mathbb{C})$ and $f \in S$, then

$$\mathrm{Re}(f) = \frac{f + \bar{f}}{2} \in S \qquad \text{and} \qquad \mathrm{Im}(f) = \frac{f - \bar{f}}{2i} \in S.$$

Lecture 30 — Wednesday, November 19, 2014

Recall that a subspace $S \subset C(X, \mathbb{C})$ is self-adjoint if $f \in S$ implies $\bar{f} \in S$. In this case, $f \in S$ implies $\mathrm{Re}(f), \mathrm{Im}(f) \in S$.

Theorem 30.1 (Stone–Weierstrass theorem: complex version) Let (X, τ) be a compact Hausdorff space and let $A \subset C(X, \mathbb{C})$ be a subspace such that the following hold:

1. A is a subalgebra of $C(X, \mathbb{C})$.
2. A is self-adjoint.
3. A separates points.
4. A contains 1.

Then A is dense in $(C(X, \mathbb{C}), \|\cdot\|_\infty)$.

Proof. Let

$$A_{\mathbb{R}} = \{f \in A \mid f(X) \subset \mathbb{R}\}.$$

Then $A_{\mathbb{R}}$ is a subalgebra of $C(X, \mathbb{R})$ and $1 \in A_{\mathbb{R}}$. If $x \neq y \in X$, then there exists $f \in A$ such that $f(x) \neq f(y)$. Then $\mathrm{Re}(f)(x) \neq \mathrm{Re}(f)(y)$ or $\mathrm{Im}(f)(x) \neq \mathrm{Im}(f)(y)$. Since A is self-adjoint, $\mathrm{Re}(f), \mathrm{Im}(f) \in A_{\mathbb{R}}$. By the real version of the Stone–Weierstrass theorem, $A_{\mathbb{R}}$ is dense in $C(X, \mathbb{R})$.

Let $f \in C(X, \mathbb{C})$ and $\varepsilon > 0$. Then $\mathrm{Re}(f), \mathrm{Im}(f) \in C(X, \mathbb{R})$, so there exist $g_1, g_2 \in A_{\mathbb{R}}$ such that $\|\mathrm{Re}(f) - g_1\|_\infty < \varepsilon/2$ and $\|\mathrm{Im}(f) - g_2\|_\infty < \varepsilon/2$. Let $g = g_1 + ig_2 \in A$. Then

$$\|f - g\|_\infty = \|\mathrm{Re}(f) + i\mathrm{Im}(f) - (g_1 + ig_2)\|_\infty \leq \|\mathrm{Re}(f) - g_1\|_\infty + \|\mathrm{Im}(f) - g_2\|_\infty < \varepsilon.$$

□

Corollary 30.2 Let A be the set of all functions on $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ of the form

$$f(z) = \sum_{k=-n}^{n} c_k z^k,$$

where $n \in \mathbb{N}, c_k \in \mathbb{C}$. Then A is dense in $(C(\mathbb{T}, \mathbb{C}), \|\cdot\|_\infty)$.

Proof. Let

$$A = \left\{ f : \mathbb{T} \to \mathbb{C} \;\middle|\; f(z) = \sum_{k=-n}^{n} c_k z^k \text{ for some } n \in \mathbb{N},\ c_k \in \mathbb{C} \right\}.$$

Then A is a subalgebra of $C(\mathbb{T}, \mathbb{C})$: if $f, g \in A$, then $f + g, \lambda f \in A$, and also $fg \in A$ because

$$\left(\sum_{k=-m}^{m} a_k z^k \right) \left(\sum_{j=-n}^{n} b_j z^j \right) = \sum_{\ell=-(m+n)}^{m+n} d_\ell z^\ell$$

for suitable coefficients $d_\ell \in \mathbb{C}$. Also $1 \in A$ (take $n = 0$ and $c_0 = 1$), and A separates points because the identity function $z \mapsto z$ belongs to A.

It remains to show that A is self-adjoint. Let $f(z) = \sum_{k=-n}^{n} c_k z^k \in A$. For $z \in \mathbb{T}$, we have $\overline{z} = z^{-1}$, so

$$\overline{f(z)} = \sum_{k=-n}^{n} \overline{c_k}\, \overline{z}^k = \sum_{k=-n}^{n} \overline{c_k}\, z^{-k} = \sum_{k=-n}^{n} d_k z^k,$$

where $d_k = \overline{c_{-k}} \in \mathbb{C}$. Hence $\overline{f} \in A$.

Therefore all hypotheses of Theorem 30.1 are satisfied, so A is dense in $(C(\mathbb{T}, \mathbb{C}), \|\cdot\|_\infty)$. $\square$

Let

$$C_{\text{per}}([0, 2\pi], \mathbb{K}) = \{f \in C([0, 2\pi], \mathbb{K}) \mid f(0) = f(2\pi)\}.$$

This space can be identified with $C(\mathbb{T}, \mathbb{K})$ in the following way: define $\phi : [0, 2\pi] \to \mathbb{T}$ by $t \mapsto e^{it}$. Then $T : C(\mathbb{T}, \mathbb{K}) \to C_{\text{per}}([0, 2\pi], \mathbb{K})$ defined by $f \mapsto f \circ \phi$ is an **isometric isomorphism** (that is, a linear surjective isometry). In particular, if $A \subset C(\mathbb{T}, \mathbb{K})$ is dense, then $T(A)$ is dense in $C_{\text{per}}([0, 2\pi], \mathbb{K})$.

Corollary 30.3

1. The set of functions on $[0, 2\pi]$ of the form
$$f(t) = \sum_{k=-n}^{n} c_k e^{ikt}$$
where $n \in \mathbb{N}, c_k \in \mathbb{C}$ for $-n \le k \le n$, is dense in $C_{\text{per}}([0, 2\pi], \mathbb{C})$.
2. The set of all functions on $[0, 2\pi]$ of the form
$$f(t) = a_0 + \sum_{k=1}^{n} a_k \sin(kt) + \sum_{k=1}^{n} b_k \cos(kt)$$
where $n \in \mathbb{N}, a_0, a_k, b_k \in \mathbb{R}$, is dense in $C_{\text{per}}([0, 2\pi], \mathbb{R})$ with respect to $\|\cdot\|_\infty$.

Proof. For (1), let
$$\mathcal{A} = \left\{ z \mapsto \sum_{k=-n}^{n} c_k z^k \;\middle|\; n \in \mathbb{N},\ c_k \in \mathbb{C} \right\} \subset C(\mathbb{T}, \mathbb{C}).$$

By Corollary 30.2, $\mathcal{A}$ is dense in $(C(\mathbb{T}, \mathbb{C}), \|\cdot\|_\infty)$. Let $T : C(\mathbb{T}, \mathbb{C}) \to C_{\text{per}}([0, 2\pi], \mathbb{C})$ be the isometric isomorphism $T(f) = f \circ \phi$, where $\phi(t) = e^{it}$. Since isometric isomorphisms preserve density, $T(\mathcal{A})$ is dense in $C_{\text{per}}([0, 2\pi], \mathbb{C})$. For $f(z) = \sum_{k=-n}^{n} c_k z^k \in \mathcal{A}$, we have
$$T(f)(t) = f(e^{it}) = \sum_{k=-n}^{n} c_k e^{ikt}.$$

Hence (1) follows.

For (2), let $u \in C_{\text{per}}([0, 2\pi], \mathbb{R})$ and let $\varepsilon > 0$. Viewing u as a complex-valued function, part (1) gives coefficients $c_k \in \mathbb{C}$ and $n \in \mathbb{N}$ such that
$$\left\| u - \sum_{k=-n}^{n} c_k e^{ikt} \right\|_\infty < \varepsilon.$$

Define
$$p(t) = \operatorname{Re}\left(\sum_{k=-n}^{n} c_k e^{ikt} \right).$$

Then $p \in C_{\text{per}}([0, 2\pi], \mathbb{R})$ and

$$\|u - p\|_\infty \leq \left\| u - \sum_{k=-n}^{n} c_k e^{ikt} \right\|_\infty < \varepsilon,$$

because u is real-valued and $|\operatorname{Re}(w)| \leq |w|$ for all $w \in \mathbb{C}$.

It remains to identify the form of p. Write $c_k = \alpha_k + i\beta_k$ with $\alpha_k, \beta_k \in \mathbb{R}$. Since

$$\operatorname{Re}(c_k e^{ikt}) = \alpha_k \cos(kt) - \beta_k \sin(kt),$$

we obtain

$$p(t) = a_0 + \sum_{k=1}^{n} a_k \sin(kt) + \sum_{k=1}^{n} b_k \cos(kt)$$

for suitable real coefficients a_0, a_k, b_k. Therefore (2) holds. $\square$

The Arzelà-Ascoli Theorem

Definition 30.4 Let $X \neq \emptyset$ be a set and let $\mathcal{F} \subset \mathbb{K}^X$. We say that $\mathcal{F}$ is **bounded at every point** (or **pointwise bounded**) if for every $x \in X$, $\sup_{f \in \mathcal{F}} |f(x)| < \infty$. We say $\mathcal{F}$ is **uniformly bounded** if there exists $M \geq 0$ such that $\sup_{f \in \mathcal{F}} |f(x)| \leq M$ for all $x \in X$.

Lecture 31 — Wednesday, November 19, 2014

Recall that $\mathcal{F} \subset \mathbb{K}^X$ is pointwise bounded if for every $x \in X$, $\sup_{f \in \mathcal{F}} |f(x)| < \infty$. The collection $\mathcal{F}$ is uniformly bounded if there exists $M \geq 0$ such that $\sup_{f \in \mathcal{F}} |f(x)| \leq M$ for all $x \in X$.

Example 31.1 Consider $C([0, 1], \mathbb{R})$.

1. For $n \in \mathbb{N}$, define $f_n \in C([0, 1], \mathbb{R})$ by $f_n(x) = x^n$ for $x \in [0, 1]$. Let $\mathcal{F} = \{f_n\}_{n \in \mathbb{N}}$. Then $\mathcal{F}$ is uniformly bounded by $M = 1$; that is, $\|f_n\|_\infty = 1$ for $n \in \mathbb{N}$. However, the sequence (f_n) does not admit a subsequence which converges in $(C([0, 1], \mathbb{R}), \|\cdot\|_\infty)$.

Indeed, if (f_{n_k}) is a subsequence of (f_n), then

$$\lim_{k\to\infty} f_{n_k}(x) = \begin{cases} 0, & x \in [0,1), \\ 1, & x = 1. \end{cases}$$

Thus the pointwise limit of (f_{n_k}) is not continuous. Since convergence in $\|\cdot\|_\infty$ implies pointwise convergence, (f_{n_k}) cannot converge in $(C([0,1],\mathbb{R}), \|\cdot\|_\infty)$. Hence the unit ball of $C([0,1],\mathbb{R})$ is not compact. More generally, the unit ball of an infinite-dimensional normed vector space is never compact.

2. Let $g_n(x) = \sin(nx)$ for $x \in [0,1]$. Then $\|g_n\|_\infty \leq 1$ for all $n \in \mathbb{N}$, so $\{g_n \mid n \in \mathbb{N}\}$ is uniformly bounded. Using measure theory, we can show that the sequence $(g_n)_n$ does not admit a pointwise convergent subsequence.

Definition 31.2 Let (X, d_X) and (Y, d_Y) be metric spaces. A collection $\mathcal{F}$ of functions from X to Y is said to be **equicontinuous at $\boldsymbol{x_0 \in X}$** if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $d_X(x, x_0) < \delta$ implies $d_Y(f(x), f(x_0)) < \varepsilon$, for all $f \in \mathcal{F}$. We say that $\mathcal{F}$ is **equicontinuous** if $\mathcal{F}$ is equicontinuous at each $x_0 \in X$. We say that $\mathcal{F}$ is **uniformly equicontinuous** if for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $d_X(x, x') < \delta$, then $d_Y(f(x), f(x')) < \varepsilon$ for all $f \in \mathcal{F}$.

The key word is "all": one horizontal tolerance works simultaneously for every curve in the family. Figure 10 illustrates this common δ at x_0 for three different functions.

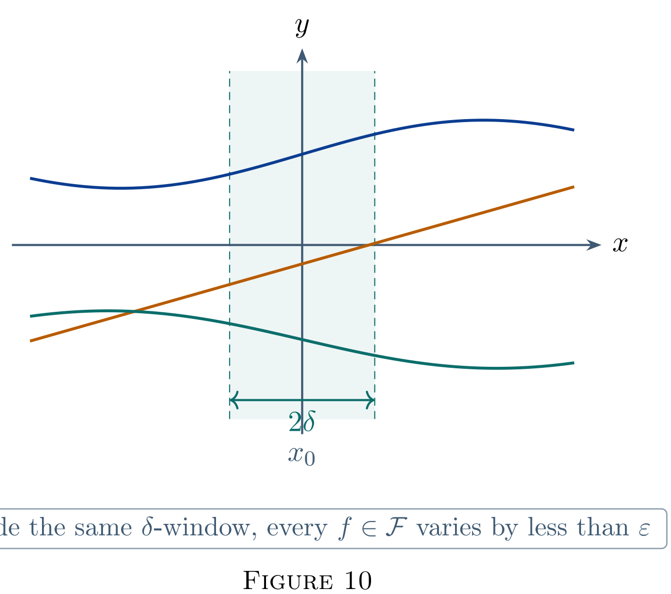


Figure 10

Example 31.3

1. Every finite collection of continuous functions is equicontinuous. Similarly, every finite collection of uniformly continuous functions is uniformly equicontinuous.
2. The collection $\{f_n\}_{n\in\mathbb{N}}$ from Example 31.1 is not equicontinuous at 1. Indeed, let $\varepsilon = 1/2$ and assume there exists $\delta \in (0,1)$ such that if $1-\delta < x \leq 1$, then $|f_n(x) - f_n(1)| < 1/2$ for all $n \in \mathbb{N}$. Since $f_n(1) = 1$ for all $n \in \mathbb{N}$, taking $x = 1-\delta/2$ gives
$$f_n(x) = (1-\delta/2)^n \to 0$$
as $n \to \infty$, a contradiction. Hence $\{f_n \mid n \in \mathbb{N}\}$ is not equicontinuous at 1.

Definition 31.4 Let (X, d_X) and (Y, d_Y) be metric spaces. We say a function $f : X \to Y$ is **Lipschitz continuous** if there exists $L \geq 0$ such that
$$d_Y(f(x), f(x')) \leq L \cdot d_X(x, x')$$
for all $x, x' \in X$. Any such L is called a **Lipschitz constant**.

Every Lipschitz continuous function is uniformly continuous. If $\mathcal{F}$ is a family of Lipschitz continuous functions with uniformly bounded Lipschitz constants, then $\mathcal{F}$ is uniformly equicontinuous.

Example 31.5 Let $f : [a,b] \to \mathbb{R}$ be continuous and differentiable on (a,b). If there exists $L \geq 0$ such that $|f'(x)| \leq L$ for all $x \in (a,b)$, then f is Lipschitz continuous with Lipschitz constant L by the mean value theorem. In particular, if $\mathcal{F} \subset C([a,b],\mathbb{R})$ is a family of functions differentiable on (a,b) and there exists $L \geq 0$ such that $|f'(x)| \leq L$ for all $x \in (a,b)$ and all $f \in \mathcal{F}$, then $\mathcal{F}$ is uniformly equicontinuous.

Definition 31.6 Let (X,d) be a metric space and let $A \subset X$. We say that A is **relatively compact** in X if $\overline{A}$ is compact.

Example 31.7 By the Heine–Borel theorem, a subset of $\mathbb{R}^n$ is relatively compact if and only if it is bounded.

Proposition 31.8 Let (X, d) be a complete metric space and let $A \subset X$. The following are equivalent:

1. A is relatively compact.
2. A is totally bounded.
3. Every sequence in A admits a subsequence which converges in X.

Proof. (1) *implies* (3). If A is relatively compact, then $\overline{A}$ is compact, hence sequentially compact because we are in a metric space. Every sequence in A is a sequence in $\overline{A}$, so it has a subsequence converging in $\overline{A} \subset X$.

(3) *implies* (2). Suppose A is not totally bounded. Then there exists $\varepsilon > 0$ such that no finite collection of ε-balls covers A. Choose $x_1 \in A$ and, inductively, choose

$$x_{n+1} \in A \setminus \bigcup_{j=1}^{n} B_\varepsilon(x_j).$$

Then $d(x_m, x_n) \geq \varepsilon$ for $m \neq n$, so (x_n) has no Cauchy subsequence and therefore no convergent subsequence, contradicting (3).

(2) *implies* (1). Since A is totally bounded, $\overline{A}$ is totally bounded. Since X is complete and $\overline{A}$ is closed, $\overline{A}$ is complete. A complete totally bounded metric space is compact, so $\overline{A}$ is compact and hence A is relatively compact. □

Lecture 32 — Monday, November 24, 2014

Recall that if $\mathcal{F}$ is a collection of functions from X to Y, then $\mathcal{F}$ is equicontinuous at $x_0 \in X$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $d_X(x, x_0) < \delta$ implies $d_Y(f(x), f(x_0)) < \varepsilon$ for all $f \in \mathcal{F}$. The collection $\mathcal{F}$ is uniformly equicontinuous if δ can be chosen independently of x_0.

Lemma 32.1 Let (X, d_X) be a compact metric space and let (Y, d_Y) be a metric space. If $\mathcal{F}$ is an equicontinuous collection of functions from $X \to Y$, then $\mathcal{F}$ is uniformly equicontinuous.

Proof. Let $\varepsilon > 0$. Since $\mathcal{F}$ is equicontinuous, we can find for every $x \in X$ a $\delta_x > 0$ such that $d_X(x, x') < \delta_x$ implies $d_Y(f(x), f(x')) < \varepsilon/2$ for all $f \in \mathcal{F}$. Then $\{B_{\delta_x}(x)\}_{x \in X}$ is an open cover of X. By Lebesgue's covering lemma, this open cover admits a Lebesgue number $\delta > 0$. If $d_X(x, x') < \delta$, then $\operatorname{diam}\{x, x'\} < \delta$, so there exists some $x_0 \in X$ such that $x, x' \in B_{\delta_{x_0}}(x_0)$. Thus

$$d_Y(f(x), f(x')) \leq d_Y(f(x), f(x_0)) + d_Y(f(x_0), f(x')) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for every $f \in \mathcal{F}$. Thus, $\mathcal{F}$ is uniformly equicontinuous. $\square$

Lemma 32.2 Let A be a countable set and let $(f_n)_n$ be a sequence of functions from $A \to \mathbb{K}$ which is pointwise bounded. Then $(f_n)_n$ admits a subsequence $(f_{n_k})_k$ such that $(f_{n_k})_k$ converges pointwise (that is, $(f_{n_k}(a))_k$ converges for all $a \in A$).

Proof. If A is finite, successive applications of the Bolzano–Weierstrass theorem, one for each point of A, give the desired subsequence. Otherwise, write $A = \{a_n\}_{n=1}^{\infty}$. By assumption, $(f_n(a_1))_n$ is a bounded sequence in $\mathbb{K}$. So by the Bolzano–Weierstrass theorem, there exists a subsequence $(f_{1,n})_n$ of $(f_n)_n$ such that $(f_{1,n}(a_1))_n$ converges. By the same argument, we may choose a subsequence $(f_{2,n})_n$ of $(f_{1,n})_n$ such that $(f_{2,n}(a_2))_n$ converges. Observe that $(f_{2,n}(a_1))$ still converges. Proceeding recursively, there exists, for each $k \geq 1$, a subsequence $(f_{k,n})_n$ of $(f_{k-1,n})_n$ such that $(f_{k,n}(a_j))_n$ converges for $1 \leq j \leq k$. Consider the diagonal sequence $(f_{n,n})_n$. Then $(f_{n,n})_n$ is a subsequence of (f_n) such that $(f_{n,n}(a_j))$ converges for all $j \in \mathbb{N}$. $\square$

Theorem 32.3 (Arzelà–Ascoli theorem) Let (X, d) be a compact metric space and let $\mathcal{F} \subset C(X, \mathbb{K})$. The following are equivalent:

1. $\mathcal{F}$ is relatively compact in $(C(X, \mathbb{K}), \|\cdot\|_\infty)$.
2. $\mathcal{F}$ is uniformly bounded and uniformly equicontinuous.
3. $\mathcal{F}$ is equicontinuous and pointwise bounded.

Proof. (1) *implies* (2). Assume $\mathcal{F}$ is relatively compact. Then $\overline{\mathcal{F}}$ is compact and hence bounded in $(C(X, \mathbb{K}), \|\cdot\|_\infty)$. Thus there exists $M \geq 0$ such that $\|f\|_\infty \leq M$ for all $f \in \mathcal{F}$. Hence $|f(x)| \leq M$ for all $x \in X$ and all $f \in \mathcal{F}$. Thus $\mathcal{F}$ is uniformly bounded. Let $\varepsilon > 0$. Since $\mathcal{F}$ is relatively compact, it is totally bounded, so there exists a finite $(\varepsilon/3)$-net $\{f_1, \ldots, f_k\}$. Then each f_j is uniformly continuous, since X is compact. Hence $\{f_1, \ldots, f_k\}$ is uniformly equicontinuous.

Thus, there exists $\delta > 0$ such that if $d(x, x') < \delta$, then $|f_j(x) - f_j(x')| < \varepsilon/3$ for all $1 \leq j \leq k$. If $f \in \mathcal{F}$, then there exists $j \in \{1, \ldots, k\}$ such that $\|f - f_j\|_\infty < \varepsilon/3$. If $d(x, x') < \delta$, then

$$|f(x) - f(x')| \leq |f(x) - f_j(x)| + |f_j(x) - f_j(x')| + |f_j(x') - f(x')| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

(2) *implies* (3). Uniform boundedness implies pointwise boundedness, and uniform equicontinuity implies equicontinuity.

(3) *implies* (1). This is highly nontrivial. By Proposition 31.8, it suffices to show that every sequence in $\mathcal{F}$ has a subsequence converging in $(C(X, \mathbb{K}), \|\cdot\|_\infty)$. Let (f_n) be a sequence in $\mathcal{F}$. Since X is compact, it is sequentially compact by Theorem 20.1, and then totally bounded by Proposition 19.3. For each $m \geq 1$, choose a finite $(1/m)$-net E_m. Then

$$A := \bigcup_{m=1}^{\infty} E_m$$

is countable. It is dense in X: if $x \in X$ and $\varepsilon > 0$, choose m with $1/m < \varepsilon$; then some $a \in E_m \subset A$ satisfies $d(x, a) < 1/m < \varepsilon$. Hence X is separable. By Lemma 32.2, there exists a subsequence $(g_n)_n$ of $(f_n)_n$ such that $(g_n(a))_n$ converges for all $a \in A$. By Lemma 32.1, $\mathcal{F}$ is uniformly equicontinuous. Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that if $d(x, x') < \delta$, then $|f(x) - f(x')| < \varepsilon/3$ for all $f \in \mathcal{F}$. Let $\{x_1, \ldots, x_k\}$ be a finite $(\delta/2)$-net for X. Since A is dense, for each $1 \leq j \leq k$ there exists $a_j \in A$ such that $d(a_j, x_j) < \delta/2$. Since $(g_n(a_j))_n$ converges for $1 \leq j \leq k$, there exists $N \in \mathbb{N}$ such that $|g_n(a_j) - g_m(a_j)| < \varepsilon/3$ for all $m, n \geq N$ and all $1 \leq j \leq k$. Let $x \in X$. Then there exists $j \in \{1, \ldots, k\}$ such that $d(x, x_j) < \delta/2$. Hence $d(x, a_j) < \delta$, and

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(a_j)| + |g_n(a_j) - g_m(a_j)| + |g_m(a_j) - g_m(x)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

for all $m, n \geq N$. Thus $(g_n)_n$ is Cauchy in $(C(X, \mathbb{K}), \|\cdot\|_\infty)$, and hence converges uniformly because $C(X, \mathbb{K})$ is complete. $\square$

Lecture 33 — Wednesday, November 26, 2014

Urysohn's Lemma and the Tietze Extension Theorem

Definition 33.1 Let (X, τ) be a Hausdorff topological space. We say that (X, τ) is **normal** if for any two closed and disjoint subsets $F_1, F_2 \subset X$, there are disjoint open subsets $U_1, U_2 \subset X$ such that $F_1 \subset U_1$ and $F_2 \subset U_2$.

Proposition 33.2 Compact Hausdorff spaces are normal. Metric spaces are normal as well.

Proof. First let X be compact and Hausdorff, and let $F_1, F_2 \subset X$ be closed and disjoint. For each $x \in F_1$ and $y \in F_2$, choose disjoint open sets $U_{x,y}$ and $V_{x,y}$ such that $x \in U_{x,y}$ and $y \in V_{x,y}$. For fixed x, the sets $V_{x,y}$ cover F_2, so compactness gives $y_1, \ldots, y_n \in F_2$ with

$$F_2 \subset \bigcup_{j=1}^{n} V_{x,y_j}.$$

Set $U_x = \bigcap_{j=1}^n U_{x,y_j}$ and $V_x = \bigcup_{j=1}^n V_{x,y_j}$. Then U_x and V_x are disjoint open sets with $x \in U_x$ and $F_2 \subset V_x$. The sets U_x cover F_1, so compactness gives $x_1, \ldots, x_m \in F_1$ such that $F_1 \subset \bigcup_{i=1}^m U_{x_i}$. Then

$$U = \bigcup_{i=1}^{m} U_{x_i} \qquad \text{and} \qquad V = \bigcap_{i=1}^{m} V_{x_i}$$

are disjoint open sets with $F_1 \subset U$ and $F_2 \subset V$. Thus X is normal.

Now let (X, d) be a metric space and let $F_1, F_2 \subset X$ be closed and disjoint. If one of the sets is empty, the result is immediate. Otherwise define

$$U_1 = \{x \in X \mid d(x, F_1) < d(x, F_2)\} \qquad \text{and} \qquad U_2 = \{x \in X \mid d(x, F_2) < d(x, F_1)\}.$$

The functions $x \mapsto d(x, F_i)$ are 1-Lipschitz, so these sets are open. They are disjoint; moreover, closedness gives $d(x, F_2) > 0 = d(x, F_1)$ for $x \in F_1$, and similarly with the indices reversed. Thus U_1 and U_2 contain F_1 and F_2, respectively, and metric spaces are normal. $\square$

Lemma 33.3 Let (X, τ) be a Hausdorff space. The following are equivalent:

1. X is normal.
2. If $F \subset X$ is closed and $U \subset X$ is open such that $F \subset U$, then there exists another open set V such that $F \subset V \subset \overline{V} \subset U$.

Proof. (1) *implies* (2). Assume X is normal. Let $F \subset U$ be as in (2). Let $F' = X \setminus U$. Then F' is closed and $F \cap F' = \emptyset$. By normality of X, there are disjoint open sets V, V' such that $F \subset V, F' \subset V'$. Then $V \subset X \setminus V'$; hence $\overline{V} \subset X \setminus V' \subset X \setminus F' = U$.

(2) *implies* (1). Assume (2) holds, and let $F_1, F_2 \subset X$ be closed and disjoint. Then $X \setminus F_2$ is open and $F_1 \subset X \setminus F_2$. By (2), there exists an open set $U_1 \subset X$ such that $F_1 \subset U_1 \subset \overline{U_1} \subset X \setminus F_2$. Let $U_2 = X \setminus \overline{U_1}$. Then U_2 is open, $U_1 \cap U_2 = \emptyset$, and $F_2 \subset U_2$. Thus, X is normal. □

Theorem 33.4 (Urysohn's lemma) Let (X, τ) be a normal topological space and let $F_1, F_2 \subset X$ be closed and disjoint. Then there exists a continuous function $f : X \to [0, 1]$ such that $f(x) = 0$ for all $x \in F_1$, and $f(x) = 1$ for all $x \in F_2$.

Proof. Let $F = F_1$ and $U = X \setminus F_2$. Then U is open and $F \subset U$. By Lemma 33.3, there exists an open set $U_{1/2}$ such that

$$F \subset U_{1/2} \subset \overline{U_{1/2}} \subset U.$$

By the same argument, there exist open sets $U_{1/4}, U_{3/4} \subset X$ such that

$$F \subset U_{1/4} \subset \overline{U_{1/4}} \subset U_{1/2} \subset \overline{U_{1/2}} \subset U_{3/4} \subset \overline{U_{3/4}} \subset U.$$

Let

$$D = \left\{ \frac{m}{2^n} \mid 1 \leq m \leq 2^n - 1, n \in \mathbb{N} \right\}$$

be the set of dyadic rationals in $(0, 1)$. Continuing recursively, we obtain for each $t \in D$ an open set $U_t \subset X$ such that

$$F \subset U_s \subset \overline{U_s} \subset U_t \subset U$$

if $s < t$. Set $U_1 = X$. Define $f : X \to \mathbb{R}$ by

$$f(x) = \inf\{t \in D \cup \{1\} \mid x \in U_t\}.$$

Then $f(x) \in [0,1]$. If $x \in F$, then $x \in U_t$ for all $t \in D$, hence $f(x) = 0$. If $x \in F_2 = X \setminus U$, then $x \notin U_t$ for all $t \in D$, so $f(x) = \inf\{1\} = 1$.

It remains to show that f is continuous; that is, preimages of open sets are open. Note that every open subset of $[0,1]$ is a union of sets of the following form:

$$\begin{gathered}(a,b), \text{ where } 0 < a < b < 1\\ [0,a), \text{ where } 0 < a < 1\\ (b,1], \text{ where } 0 < b < 1\end{gathered}$$

Furthermore, $(a,b) = (a,1] \cap [0,b)$. Thus, it suffices to show that $f^{-1}([0,a))$ and $f^{-1}((a,1])$ are open for all $a \in (0,1)$. We claim that

$$f^{-1}([0,a)) = \bigcup_{\substack{t<a\\ t\in D}} U_t.$$

If $f(x) < a$, choose $t \in D$ with $f(x) < t < a$. By the definition of the infimum, there is some $s < t$ such that $x \in U_s$. Since $s < t$, we have $U_s \subset U_t$, and hence $x \in U_t$. Conversely, if $x \in U_t$ for some $t \in D$ with $t < a$, then $f(x) \leq t < a$. Hence $f^{-1}([0,a))$ is open. Next, we claim that

$$f^{-1}((a,1]) = \bigcup_{\substack{t>a\\ t\in D}} X \setminus \overline{U_t}.$$

Indeed, if $f(x) > a$, then there are $t, s \in D$ such that

$$a < t < s < f(x).$$

Then $x \notin U_s$. Since $\overline{U_t} \subset U_s$, we have $x \in X \setminus \overline{U_t}$. Conversely, suppose $x \in X \setminus \overline{U_t}$ for some $t > a$. If $x \in U_s$ for some $s < t$, then $x \in U_s \subset \overline{U_s} \subset U_t \subset \overline{U_t}$, a contradiction. Every index s for which $x \in U_s$ must therefore satisfy $s \geq t$, so $f(x) \geq t > a$. Thus $f^{-1}((a,1])$ is open, and f is continuous. $\square$

Lecture 34 — Friday, November 28, 2014

Recall that Urysohn's lemma states that if (X,τ) is normal and $F_1, F_2 \subset X$ are closed and disjoint, then there exists a continuous function $f : X \to [0,1]$ such that $f\big|_{F_1} = 0$ and $f\big|_{F_2} = 1$.

Corollary 34.1 Let (X,τ) be a normal topological space and let $F_1, F_2 \subset X$ be closed and disjoint. If $a, b \in \mathbb{R}$ with $a < b$, then there exists a continuous function $f : X \to [a,b]$ such that $f\big|_{F_1} = a$ and $f\big|_{F_2} = b$.

Proof. By Urysohn's lemma, there exists $g : X \to [0,1]$ such that g is continuous, $g\big|_{F_1} = 0$ and $g\big|_{F_2} = 1$. Let $f : X \to [a,b]$, $x \mapsto (b-a)g(x) + a$. Then f satisfies the statement of the corollary. □

Remark 34.2 Let (X,τ) be a topological space and $Y \subset X$. The relative topology τ_Y is defined by $\tau_Y = \{U \cap Y \mid U \in \tau\}$. If Y is closed in X and $F \subset Y$ is closed with respect to τ_Y, then F is closed in X (that is, with respect to τ). Indeed, $Y \setminus F$ is open in Y, so there exists some $U \in \tau$ such that $Y \setminus F = U \cap Y$. Then $F = Y \setminus (U \cap Y) = (X \setminus U) \cap Y$ is τ-closed. So being closed is transitive.

Theorem 34.3 (Tietze extension theorem) Let (X,τ) be a normal topological space and let $Y \subset X$ be closed. Suppose $a, b \in \mathbb{R}$ with $a < b$, and let $f : Y \to [a,b]$ be continuous. Then there exists a continuous function $g : X \to [a,b]$ such that $g\big|_Y = f$.

Proof. It suffices to consider the case where $a = -1$ and $b = 1$. Indeed, if $f : Y \to [a,b]$ is continuous, then

$$\tilde{f}(x) = \frac{2f(x) - a - b}{b - a}$$

defines a continuous function $\tilde{f} : Y \to [-1,1]$. If $\tilde{g} : X \to [-1,1]$ is a continuous extension of $\tilde{f}$, then

$$g(x) = \frac{1}{2}\big((b-a)\tilde{g}(x) + a + b\big)$$

is a continuous extension of f to X with values in $[a,b]$. Thus we may assume that $f : Y \to [-1,1]$.

We recursively construct continuous functions $g_n : X \to \mathbb{R}$ and residuals $r_n : Y \to \mathbb{R}$. Set $r_0 = f$. Suppose

$$r_n(Y) \subset \left[-\left(\frac{2}{3}\right)^n, \left(\frac{2}{3}\right)^n\right].$$

Let

$$F_n = r_n^{-1}\left(\left[-\left(\frac{2}{3}\right)^n, -\frac{1}{3}\left(\frac{2}{3}\right)^n\right]\right) \qquad \text{and} \qquad G_n = r_n^{-1}\left(\left[\frac{1}{3}\left(\frac{2}{3}\right)^n, \left(\frac{2}{3}\right)^n\right]\right).$$

Then F_n and G_n are closed in X and disjoint, by Remark 34.2. By Corollary 34.1, there exists a continuous function

$$g_n : X \to \left[-\frac{1}{3}\left(\frac{2}{3}\right)^n, \frac{1}{3}\left(\frac{2}{3}\right)^n\right]$$

such that

$$g_n\Big|_{F_n} = -\frac{1}{3}\left(\frac{2}{3}\right)^n \qquad \text{and} \qquad g_n\Big|_{G_n} = \frac{1}{3}\left(\frac{2}{3}\right)^n.$$

Define $r_{n+1} = r_n - g_n\Big|_Y$. If $r_n(y)$ lies in the lower third, middle third, or upper third of its range, respectively, then the defining bounds on $g_n(y)$ show in each case that $|r_{n+1}(y)| \leq (2/3)^{n+1}$. Hence

$$r_{n+1}(Y) \subset \left[-\left(\frac{2}{3}\right)^{n+1}, \left(\frac{2}{3}\right)^{n+1}\right].$$

The construction gives

$$\|g_n\|_\infty \leq \frac{1}{3}\left(\frac{2}{3}\right)^n.$$

Thus

$$\sum_{n=0}^{\infty} \|g_n\|_\infty \leq \sum_{n=0}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^n = 1.$$

Therefore the series $\displaystyle\sum_{n=0}^{\infty} g_n$ converges absolutely in the Banach space $(C_b(X,\mathbb{R}), \|\cdot\|_\infty)$. Define

$$g = \sum_{n=0}^{\infty} g_n \in C_b(X,\mathbb{R}).$$

Then

$$\|g\|_\infty \leq \sum_{n=0}^{\infty} \|g_n\|_\infty \leq 1.$$

Thus $g(X) \subset [-1,1]$. For $y \in Y$,

$$0 \leq |f(y) - g(y)| = \lim_{n\to\infty} \left| f(y) - \sum_{j=0}^{n} g_j(y) \right| = \lim_{n\to\infty} |r_{n+1}(y)| \leq \lim_{n\to\infty} \left(\frac{2}{3}\right)^{n+1} = 0.$$

Hence $g\Big|_Y = f$. □

Theorem 34.4 Let (X, τ) be a normal topological space, let $Y \subset X$ be closed, and let $f : Y \to \mathbb{R}$ be continuous. Then there exists a continuous function $g : X \to \mathbb{R}$ such that $g\Big|_Y = f$.

Proof. Let $\tilde{f} = \arctan(f)$. Then $\tilde{f} : Y \to (-\pi/2, \pi/2)$ is continuous. By the Tietze extension theorem, there exists a continuous function $\tilde{g} : X \to [-\pi/2, \pi/2]$ such that $\tilde{g}\Big|_Y = \tilde{f}$. Let

$$F = \tilde{g}^{-1}\left(\left\{-\frac{\pi}{2}, \frac{\pi}{2}\right\}\right).$$

Then F is closed and $F \cap Y = \emptyset$. By Urysohn's lemma, there exists a continuous function $h : X \to [0,1]$ such that $h\Big|_F = 0$ and $h\Big|_Y = 1$. Then $h\tilde{g}$ takes values in $(-\pi/2, \pi/2)$, agrees with $\tilde{f}$ on Y, and is continuous. Hence

$$g = \tan(h\tilde{g})$$

is continuous on X and satisfies $g\Big|_Y = f$. □

Lecture 35 — Monday, December 01, 2014

Theorem 35.1 Let (X, τ) be a Hausdorff space. The following are equivalent:

1. X is normal.
2. If $F_1, F_2 \subset X$ are closed and disjoint, then there exists a continuous function $f : X \to [0,1]$ such that $f\Big|_{F_1} = 0$ and $f\Big|_{F_2} = 1$.

3. If $Y \subset X$ is closed, $a, b \in \mathbb{R}$ with $a < b$, and $f : Y \to [a, b]$ is continuous, then there exists a continuous function $g : X \to [a, b]$ such that $g\Big|_Y = f$.

Proof. (1) *implies* (2). This is Urysohn's lemma.

(2) *implies* (1). Let F_1, F_2 be closed and disjoint, and let f be as in (2). Define $U_1 = f^{-1}([0, 1/2))$ and $U_2 = f^{-1}((1/2, 1])$. Then U_1 and U_2 are open, $F_1 \subset U_1$, $F_2 \subset U_2$, and $U_1 \cap U_2 = \emptyset$.

(1) *implies* (3). This is the Tietze extension theorem.

(3) *implies* (1). Let $F_1, F_2 \subset X$ be closed and disjoint. Let $Y = F_1 \cup F_2$, which is closed. Let $f_0 : Y \to [0, 1]$ be defined by $f_0(x) = 0$ if $x \in F_1$ and $f_0(x) = 1$ if $x \in F_2$. Note that F_1 and F_2 are both open and closed in Y. Hence, f_0 is continuous on Y. By (3), there exists $f : X \to [0, 1]$ continuous such that $f\Big|_Y = f_0$. Thus f satisfies (2). □

Corollary 35.2 Let (X, d) be a metric space and assume that every continuous function $f : X \to \mathbb{R}$ is bounded. Then X is compact.

Proof. Suppose that X is not compact. We will construct a continuous unbounded function on X. Since X is not compact, X does not have the Bolzano–Weierstrass property, so there exists an infinite subset $A \subset X$ with no accumulation points. Replacing A with a subset of A, we may assume that A is countably infinite, say $A = \{a_n\}_{n=1}^{\infty}$ where $a_n \neq a_m$ if $n \neq m$. It follows that A is closed. Moreover, the relative topology on A is the discrete topology: if $a \in A$, there exists $\delta > 0$ such that $B_\delta(a) \setminus \{a\}$ is disjoint from A, and hence $B_\delta(a) \cap A = \{a\}$. Thus every function on A is continuous. Define $f : A \to \mathbb{R}$ by $f(a_n) = n$. By the Tietze extension theorem, f extends to a continuous function $g : X \to \mathbb{R}$. Since $g(a_n) = n$ for all n, the function g is unbounded, which is a contradiction. □

Appendix: Lecture and Tutorial Index

1. Lecture 1 — Monday, September 08, 2014
2. Lecture 2 — Wednesday, September 10, 2014
3. Lecture 3 — Friday, September 12, 2014
4. Lecture 4 — Monday, September 15, 2014
5. Lecture 5 — Wednesday, September 17, 2014
6. Lecture 6 — Friday, September 19, 2014
7. Lecture 7 — Monday, September 22, 2014
8. Lecture 8 — Wednesday, September 24, 2014
9. Lecture 9 — Friday, September 26, 2014
10. Lecture 10 — Monday, September 29, 2014
11. Lecture 11 — Wednesday, October 01, 2014
12. Lecture 12 — Friday, October 03, 2014
13. Lecture 13 — Monday, October 06, 2014
14. Lecture 14 — Wednesday, October 08, 2014
15. Lecture 15 — Friday, October 10, 2014
16. Lecture 16 — Wednesday, October 15, 2014
17. Lecture 17 — Friday, October 17, 2014
18. Lecture 18 — Monday, October 20, 2014
19. Lecture 19 — Friday, October 24, 2014
20. Lecture 20 — Monday, October 27, 2014
21. Lecture 21 — Wednesday, October 29, 2014
22. Lecture 22 — Wednesday, October 29, 2014
23. Lecture 23 — Friday, October 31, 2014
24. Lecture 24 — Monday, November 03, 2014
25. Lecture 25 — Wednesday, November 05, 2014
26. Lecture 26 — Monday, November 10, 2014

27. Lecture 27 — Wednesday, November 12, 2014
28. Lecture 28 — Friday, November 14, 2014
29. Lecture 29 — Friday, November 14, 2014
30. Lecture 30 — Wednesday, November 19, 2014
31. Lecture 31 — Wednesday, November 19, 2014
32. Lecture 32 — Monday, November 24, 2014
33. Lecture 33 — Wednesday, November 26, 2014
34. Lecture 34 — Friday, November 28, 2014
35. Lecture 35 — Monday, December 01, 2014