EL3370 Mathematical Methods in Signals, Systems and Control

Topic 1: Introduction and Preliminaries

Cristian R. Rojas

Division of Decision and Control Systems
KTH Royal Institute of Technology

Mathematical Methods in Signals, Systems and Control

- **Disposition**: 8 credits, 10 topics, 13 lectures
- **Teacher**: Cristian R. Rojas, crro@kth.se
- **Web page**:
  `https://people.kth.se/~crro/Math_Methods2025/Math_Methods.html`
- **Objectives**:
  - Introduction of mathematical tools essential for understanding results from control theory, signal processing and communications.
  - Focus in aspects of functional analysis, specifically basics of Hilbert and Banach spaces. Theory will be complemented with examples from robust control, game theory, model reduction, estimation/filtering theory and system identification.
- **Textbook**: N. Young. *An Introduction to Hilbert Space*. CUP, 1988.
- **Complements**:
  - D. G. Luenberger. *Optimization by Vector Space Methods*. Wiley & Sons, 1969.
  - E. Kreyszig. *Introductory Functional Analysis with Applications*. Wiley & Sons, 1989.

Mathematical Methods in Signals, Systems and Control (cont.)

- **Evaluation:**
  - *Assignments* (80 %)
  - *Project* (20 %): Analysis of a particular application or extension of the theory presented in a recent publication from the areas of control, signal processing or communications, preferably related to the student's own research, with a 15 min presentation of the main ideas/results of that publication.

- **Schedule:**
  1. Introduction
  2. Inner product spaces
  3. Normed spaces
  4. Hilbert and Banach spaces
  5. Orthogonal expansions, classical Fourier series
  6. Estimation and optimization in Hilbert spaces
  7. Dual spaces, Hahn-Banach theorem
  8. Linear operators
  9. Optimization of functionals
  10. Application to H_∞ control theory

- **Expectations:**
  - The contents are not easy. Do not try to do the assignments the day before the deadline.
  - Attending the lectures: not compulsory but highly recommended. This is *not* a long distance course!
  - You develop understanding by working hard on the exercises (in assignments and slides). Do not be afraid to ask the teacher if lost or believe there are missing steps. Also, do not expect that the material is self-contained in the slides: just as in research, dare to check the course book and other references!
  - Regarding the assignments:
    - only results proven in class/slides can be used to solve the problems,
    - the explanations/proofs should be rigorous,
    - deadlines should be met,
    - discussion is encouraged, but the assignments are individual.
  - (Constructive) feedback welcome! The teacher may not know if there are missing concepts/prerequisites/... until it is too late, unless you inform him.

Outline

Outline

Motivation

Functional Analysis = Linear algebra in infinite dimensions ("function spaces")

Historical motivations: Integral equations, foundations of quantum mechanics.

Example (integral equation)

$$\int_0^1 K(x,y)f(y)dy = g(x), \qquad K:[0,1]\times[0,1]\to\mathbb{R}, \quad g:[0,1]\to\mathbb{R} \text{ continuous and known}$$

find: $f:[0,1]\to\mathbb{R}$ continuous.

discretization: $\sum_{j=0}^{n-1} K(i/n, j/n) f_j^n \frac{1}{n} = g(i/n),\ i = 0,\dots,n-1.$

do f_j^n's approximate f for large n? in what sense?

Example:

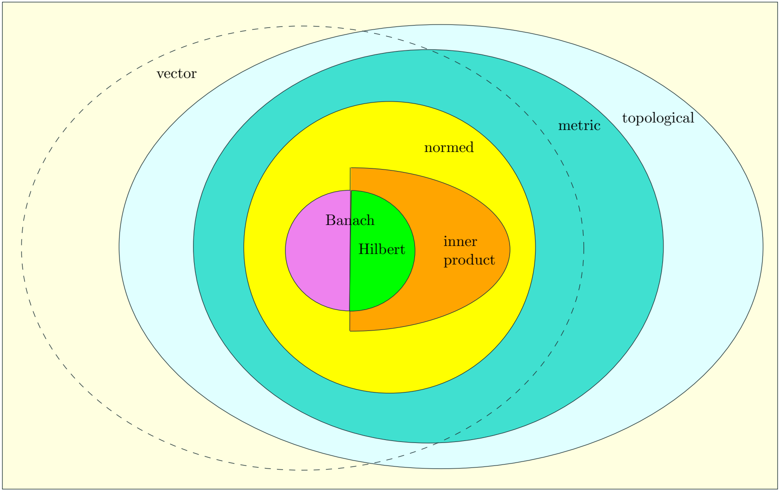
$$\begin{aligned} x_1 + x_2 + x_3 + \cdots &= 1 \\ x_2 + x_3 + \cdots &= 1 \\ x_3 + \cdots &= 1 \\ &\vdots \end{aligned}$$

System of equations with no solution, but truncation (*i.e.*, making $x_n = x_{n+1} = \cdots = 0$ for some n) suggests $x_1 = x_2 = \cdots = 0$!

Not only algebraic aspects matter in ∞ dimensions, but also *topological/analytical* aspects.

The Big Picture

Space = "set with structure" (algebraic, geometrical, order-theoretical, topological, . . .)



Outline

Preliminaries

- **A Dictionary**
- **Sets**
- **Quantifiers**
- **Mappings**
- **Mathematical Induction**
- **Families and Sequences**
- **Countability**
- **Equivalence Relations and Partitions**
- **Order Relations**
- **Supremum and Infimum**
- **Axiom of Choice, Zorn's Lemma, ...**

A Dictionary

Definition: precise, unambiguous description of the meaning of a mathematical term.

Theorem: mathematical statement proved using rigorous mathematical reasoning; in a mathematical paper, this term is often reserved for the most important results.

Lemma: minor result whose sole purpose is to help in proving a theorem. It is a stepping stone on the path to proving a theorem. Some lemmas can take on a life of their own.

Corollary: result in which the (usually short) proof relies heavily on a given theorem.

Proposition: proved, often interesting result, but less important than a theorem.

Conjecture: statement that is unproved, but is believed to be true.

Claim: assertion that is then proved. It is often used like an informal lemma.

Axiom/Postulate: a statement that is assumed to be true without proof; these are the basic building blocks from which all theorems are proved.

Identity: mathematical expression giving the equality of two (often variable) quantities.

Sets

Sets = collection of objects

$\emptyset$:	*empty set*
$a \in A$ ($b \notin A$):	a (b) is (not) an *element* of A
$A = B$:	A and B are *equal*, *i.e.*, $a \in A$ if and only if $a \in B$
$A \subseteq B$:	A is a *subset* of B, *i.e.*, $a \in A$ implies that $a \in B$
$A \subsetneq B$:	A is a *proper subset* of B, *i.e.*, $A \subseteq B$ but $A \neq B$
$A \cup B$:	*union* of A and B, *i.e.*, $\{x : x \in A \text{ or } x \in B\}$
$A \cap B$:	*intersection* of A and B, *i.e.*, $\{x : x \in A \text{ and } x \in B\}$
$A \cap B = \emptyset$:	A and B are *disjoint*
$A \setminus B$:	*difference* of A and B, *i.e.*, $\{x : x \in A \text{ and } x \notin B\}$
A^c:	*complement* of A, *i.e.*, $X \setminus A$, where X = *universe* set (objects of interest)
$\mathscr{P}(A)$:	*power set* of A, *i.e.*, $\{B : B \subseteq A\}$
$A \times B$:	*Cartesian product* of A and B, *i.e.*, $\{(x, y) : x \in A \text{ and } y \in B\}$
$\mathbb{N}$:	set of natural numbers, *i.e.*, $\{1, 2, 3, \dots\}$
$\mathbb{Z}$:	set of integers, *i.e.*, $\{\dots, -2, -1, 0, 1, 2, \dots\}$
$\mathbb{Q}$:	set of rational numbers, *i.e.*, $\{a/b : a, b \in \mathbb{Z}, b \neq 0\}$
$\mathbb{R}$:	set of real numbers
$\mathbb{C}$:	set of complex numbers, *i.e.*, $\{a + bi : a, b \in \mathbb{R}\}$

A set with only one element is called a *singleton*.

Quantifiers

Two types of quantifiers:

- *Universal*: “for all/every” ($\forall$)
- *Existential*: ”there exists/is” ($\exists$)

Example $\lim_{n\to\infty} a_n = a \quad \Leftrightarrow \quad \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geqslant N, |a_n - a| < \varepsilon$

$\Leftrightarrow$ for all $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that for all $n \geqslant N$, $|a_n - a| < \varepsilon$

Negation $\neg(\forall x, P(x)) \quad \Leftrightarrow \quad \exists x, \neg P(x)$ “not for all x, $P(x)$” is equivalent to “there is an x such that $P(x)$ does not hold”

$\neg(\exists x, P(x)) \quad \Leftrightarrow \quad \forall x, \neg P(x)$ “there is no x such that $P(x)$” is equivalent to “for all x, $P(x)$ does not hold”

Quantifiers are not commutative in general

For example, “$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \ldots$” is not the same as “$\exists N \in \mathbb{N}, \forall \varepsilon > 0, \ldots$”, because in the first sentence, N is implicitly a function of ε, *i.e.*, everything to the right of “$\forall \varepsilon > 0$” is a function of ε (a *bound variable*); in the second sentence, N is not a function of ε.

Quantifiers (cont.)

Bound variables are restricted to the domain (or *scope*) of their quantifiers
For example, in the definition of limit, "$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geqslant N, |a_n - a| < \varepsilon$", n is a *bound variable*, which has meaning only after its quantifier ("$\forall n \ldots$"). This means that we can replace "n" by any other label and keep the same meaning of the sentence:
$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall p \geqslant N, |a_p - a| < \varepsilon$.
This is the reason why "$\lim_{n\to\infty} a_n$" and "$\lim_{p\to\infty} a_p$" mean the same.

Writing advice:

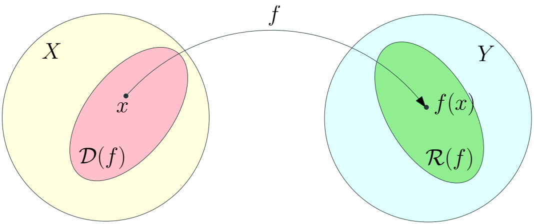
- *Avoid symbols ($\forall$, $\exists$, $\ni$) if possible* (see *P.R. Halmos, "How to write mathematics"*).
- *Avoid "any"*: it is ambiguous in English; can mean "every" or "some", depending on context.

Some abbreviations
"iff" = if and only if
"w.l.o.g." = without loss of generality
"w.r.t." = with respect to
"s.t." = such that

Mappings

Let X, Y be sets, and $A \subseteq X$.
A *mapping* (*function*, *transformation*, *operator*, ...) f from A into Y is a subset R of $A \times Y$, s.t. for every $x \in A$, there is a unique $y \in Y$, denoted $f(x)$ (*image* of x under f), for which $(x, f(x)) \in R$.
$A =: \mathscr{D}(f)$ is the *domain* of f, and Y is its *codomain*.

$$f : \mathscr{D}(f) \to Y$$
$$x \mapsto f(x)$$



$\mathscr{R}(f) := \{y \in Y : y = f(x) \text{ for some } x \in \mathscr{D}(f)\}$: *range* of f
$f(M) := \{y \in Y : y = f(x) \text{ for some } x \in M\}$: *image* of $M \subseteq \mathscr{D}(f)$ under f
$f^{-1}(N) := \{x \in \mathscr{D}(f) : f(x) \in N\}$: *preimage* of N under f

$\mathrm{id}_X : X \to X$: *identity function* on M: $\mathrm{id}_X(x) = x$ for all $x \in X$

Mappings (cont.)

A mapping f is

- *injective* (*one-to-one*) if for every $x_1, x_2 \in \mathscr{D}(f)$, $x_1 \neq x_2 \quad \Rightarrow \quad f(x_1) \neq f(x_2)$,
- *surjective* (*onto*) if $\mathscr{R}(f) = Y$,
- *bijective* if it is injective and surjective. In this case, there exists an *inverse mapping* $f^{-1}: Y \to X$ given by $f^{-1}(y) = x$ if $f(x) = y$ (*why*?).

Additional functions

Restriction of $f: X \to Y$ to $B \subseteq X$:

$f|_B : B \to Y$ s.t. $f|_B(x) = f(x)$ for all $x \in B$

Extension of $f: X \to Y$ to $C \supseteq X$:

$\tilde{f}: C \to Y$ s.t. $\tilde{f}(x) = f(x)$ for all $x \in X$

Composition of $f: X \to Y$ and $g: Y \to Z$:

$g \circ f: X \to Z$ given by $(g \circ f)(x) = g(f(x))$ for every $x \in X$

Notation. Given a set Y, and a fixed $y \in Y$, a function $f: X \to Y$ s.t. $f(x) = y$ for all $x \in X$ is sometimes denoted by the same symbol y (or a stylized version).
Also, $f(x) \equiv y$ means that $f(x) = y$ for all x.

Mathematical Induction

A standard tool for proving/defining statements that hold "for every $n \in \mathbb{N}$".
To introduce it, consider the *Peano axioms*, which state that $\mathbb{N}$ is a set satisfying

1. $1 \in \mathbb{N}$ (where $1 := \{\emptyset\}$),
2. if $n \in \mathbb{N}$, then $n^+ := n \cup \{n\} \in \mathbb{N}$ (n^+ is the *successor* of n),
3. if $S \subseteq \mathbb{N}$, $1 \in S$ and if $n^+ \in S$ whenever $n \in S$, then $S = \mathbb{N}$
   (*principle of mathematical induction*: establishes the minimality of $\mathbb{N}$),
4. $n^+ \neq 0$ for all $n \in \mathbb{N}$, and
5. if $n, m \in \mathbb{N}$, and $n^+ = m^+$, then $n = m$.

From these axioms it follows that a proposition $P(n)$ holds for every $n \in \mathbb{N}$ if

1. *Base step*: $P(1)$ is true, and
2. *Inductive step*: If $P(n)$ is true, then $P(n^+) = P(n+1)$ is also true.

This procedure is known as *mathematical induction*.

Example. Let $S(n) = 1 + 2 + \cdots + n$ for $n \in \mathbb{N}$; let us prove that $S(n) = n(n+1)/2$ by induction on $n \in \mathbb{N}$. First, $S(1) = 1 = 1 \cdot (1+1)/2$, so the statement is true for $n = 1$. If $S(n) = n(n+1)/2$ for some $n = m$, then $S(m+1) = S(m) + m + 1 = m(m+1)/2 + m + 1 = (m+1)(m+2)/2 = (m+1)([m+1]+1)/2$, so the statement is true for $n = m+1$. Then, by induction, the statement is true for every $n \in \mathbb{N}$.

See bonus slides to a generalization to arbitrary sets: transfinite induction!

Families and Sequences

A *sequence* (x_n) in X is a function $f : \mathbb{N} \to X$ that assigns $x_n = f(n)$ for all $n \in \mathbb{N}$.
$\mathbb{N}$ is the *index set* of the sequence.

Generalization
A *family* $(x_\alpha)_{\alpha \in I}$ in X is a function $f : I \to X$ s.t. $x_\alpha = f(\alpha)$ for all $\alpha \in I$. I = *index set*.
A *subfamily* is obtained by restricting f to a subset of I.

If $\mathscr{F} = (B_\alpha)_{\alpha \in I}$ is a family of subsets of X:

$\bigcup \mathscr{F} := \bigcup_{\alpha \in I} B_\alpha := \{x \in X : x \in B_\alpha \text{ for some } \alpha \in I\}$: *union* of the family $(B_\alpha)_{\alpha \in I}$

$\bigcap \mathscr{F} := \bigcap_{\alpha \in I} B_\alpha := \{x \in X : x \in B_\alpha \text{ for all } \alpha \in I\}$: *intersection* of the family $(B_\alpha)_{\alpha \in I}$

Notation. We will use $(\cdot)$ to denote families, sequences and, in general, ordered/indexed sets. $\{\cdot\}$ is used for general (un-sorted/un-indexed) sets.

Exercise: Let $f : X \to Y$, and (A_α), (B_β) families of subsets of X, Y, respectively. Prove that $f(\bigcup_\alpha A_\alpha) = \bigcup_\alpha f(A_\alpha)$, $f(\bigcap_\alpha A_\alpha) \subseteq \bigcap_\alpha f(A_\alpha)$, $f^{-1}(\bigcup_\beta B_\beta) = \bigcup_\beta f^{-1}(B_\beta)$ and $f^{-1}(\bigcap_\beta B_\beta) = \bigcap_\beta f^{-1}(B_\beta)$. Find examples of strict inclusion in the 2nd relation.

Countability

A set M is *finite* if there exists a bijection $f: \{1, \dots, n\} \to M$ for some $n \in \mathbb{N}$; in this case, the *cardinality* of M is $|M| = n$. Otherwise, M is *infinite*.

A set M is *countable* if it is finite or if there exists a bijection $f: \mathbb{N} \to M$, *i.e.*, we can "enumerate" the set as $M = \{m_1, m_2, m_3, \dots\}$.

Otherwise, M is *uncountably infinite*.

Theorem. The union of a countable number of countable sets is countable.
Proof. W.l.o.g. consider the countably infinite case. Let the sets be $A_n = \{a_n^1, a_n^2, a_n^3, \dots\}$ for $n \in \mathbb{N}$ (assumed w.l.o.g. to be pairwise disjoint, *i.e.*, $A_n \cap A_m = \emptyset$ whenever $n \neq m$), and let $A = \bigcup_{n=1}^{\infty} A_n$. We can enumerate A as $\{a_1^1, a_1^2, a_2^1, a_1^3, a_2^2, a_3^1, \dots\}$, *i.e.*, a_n^i is the $\left[\frac{(n+i-2)(n+i-1)}{2} + n\right]$-th entry of A. □

Corollary. $\mathbb{Q}$ is countable.
Proof. $\mathbb{Q} = \bigcup_{n \in \mathbb{N}} A_n$, where $A_n = \{x \in \mathbb{Q} : x = a/n \text{ for some } a \in \mathbb{Z}\}$ for each $n \in \mathbb{N}$. □

Theorem. $\mathbb{R}$ is uncountably infinite.
Proof. (*Cantor diagonal argument*) We will show that a subset of $\mathbb{R}$, [0.1], is uncountable. Assume, to the contrary, that $\{a^1, a^2, \dots\}$ is an enumeration of $[0,1]$, where $a^n = 0.a_1^n a_2^n a_3^n \dots$ is the decimal expansion of a^n ($n \in \mathbb{N}$); since, *e.g.*, $0.999\cdots = 1$, assume w.l.o.g. that each decimal expansion is infinite. Define a new number $b = 0.b_1 b_2 b_3 \cdots \in [0,1]$, where $b_n = 1$ if $a_n^n \neq 1$ and $b_n = 0$ otherwise for all $n \in \mathbb{N}$. Then $b \notin \{a^1, a^2, \dots\}$, since $b_n \neq a_n^n$ for every n. Thus, $[0,1]$ is uncountable. □

Equivalence Relations and Partitions

If X, Y are sets, $R \subseteq X \times Y$ is a *relation* in $X \times Y$. $(x, y) \in R$ can be written as $R(x, y)$ or xRy.

An *equivalence relation* on X is a relation R on $X \times X$ s.t., for all $x, y, z \in X$,

a) $R(x,x)$ (*reflexivity*),

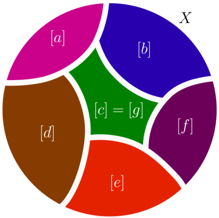
b) $R(x,y) \Rightarrow R(y,x)$ (*symmetry*),

c) $R(x,y)$ and $R(y,z) \Rightarrow R(x,z)$ (*transitivity*).

In this case, $R(x,y)$ is written as $x \sim y$.

A *partition* of a set X is a collection of pairwise disjoint subsets of X, (X_α), s.t. $X = \bigcup_\alpha X_\alpha$.

An equivalence relation R on X induces a partition on X:
For $x \in X$, let $[x] := \{y \in X : x \sim y\}$ (*equivalence class*, or *coset*, of x); any such $y \in [x]$ is a *representative* of $[x]$. $[x], [y]$ are either disjoint or equal, and the union of all equivalence classes is X.
Hence, $\{[x] : x \in X\}$ is a partition of X.
Conversely, every partition of X induces an equivalence relation on X (*how?*).
The set of cosets of X is called *quotient set*, and is denoted X/R.



Example 1: Rational numbers

A natural way to define $\mathbb{Q}$ is as a quotient set of $\{(a,b) \in \mathbb{Z}^2 : b \neq 0\}$, where $(a,b) \sim (c,d)$ iff $ad = bc$; this represents the idea that $a/b = c/d$, even though division is not defined for all pairs of integers. Then, we can define

$$[(a,b)] \pm [(c,d)] := [(ad \pm bc, bd)]$$
$$[(a,b)] \cdot [(c,d)] := [(ac, bd)]$$
$$[(a,b)] \,/\, [(c,d)] := [(ad, bc)].$$

Note. One needs to verify that these operations are well defined! *E.g.*, if $(a_1,b_1) \sim (a_2,b_2)$ and $(c_1,d_1) \sim (c_2,d_2)$, does $(a_1d_1 + b_1c_1, b_1d_1) \sim (a_2d_2 + b_2c_2, b_2d_2)$ hold? This is the same as $[a_1d_1 + b_1c_1]b_2d_2 = [a_2d_2 + b_2c_2]b_1d_1$, or $(a_1b_2)d_1d_2 + b_1b_2(c_1d_2) = (a_2b_1)d_1d_2 + b_1b_2(c_2d_1)$, which holds because $a_1b_2 = a_2b_1$ and $c_1d_2 = c_2d_1$.

Equivalence Relations and Partitions (cont.)

Example 2: Complex numbers

The algebraic way to define $\mathbb{C}$ is as a *quotient ring*: let $\mathbb{R}[x]$ be the set of polynomials in x with real coefficients, together with the operations of addition and multiplication (a *ring*). Let I be the subset of $\mathbb{R}[x]$ consisting of all polynomials of the form $p(x)(x^2+1)$, where $p(x) \in \mathbb{R}[x]$. Then, for $p(x), q(x) \in \mathbb{R}[x]$, let $p(x) \sim q(x)$ iff $p(x) - q(x) \in I$. Since x^2+1 is a polynomial of degree 2, by polynomial division we can always write for $p(x) \in \mathbb{R}[x]$

$$[p(x)] = [a + bx],$$

where $a, b \in \mathbb{R}$. Given $p, q \in \mathbb{R}[x]$, where $[p(x)] = [a_1 + b_1 x]$ and $[q(x)] = [a_2 + b_2 x]$, we can define several operations in $\mathbb{R}[x]/\sim$, such as

$$[p(x)] \pm [q(x)] := [p(x) + q(x)]$$

$$[p(x)] \cdot [q(x)] := [p(x) \cdot q(x)]$$

$$[p(x)]\,/\,[q(x)] := \left[\frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + \frac{b_1 a_2 - a_1 b_2}{a_2^2 + b_2^2} x\right] \quad (\text{if } [q(x)] \neq [0])$$

The last definition ensures that $\left[q(x) \cdot \left(\frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + \frac{b_1 a_2 - a_1 b_2}{a_2^2 + b_2^2} x\right)\right] = [p(x)]$ (*Exercise: verify this!*).

These operations coincide with those of complex numbers $a_1 + b_1 i$ and $a_2 + b_2 i$, where i satisfies $i^2 + 1 = 0$ (recall the definition of I). Thus, we can define $\mathbb{C}$ as $\mathbb{R}[x]/\sim$!

Example 3: Grassmannian spaces and Riccati differential equations

Equivalence relations can be employed to study Riccati differential equations (used in estimation and control theory) as linear differential equations!

The *Grassmannian* $\mathbf{Gr}(k,\mathbb{R}^n)$ $(k \leqslant n)$ is the set of all k-dimensional linear subspaces of the vector space $\mathbb{R}^n$. Since each such subspace can be written as span M for some $M \in \mathbb{R}^{n\times k}$ of full column rank, and span $M =$ span M' iff $M = M'N$ for some non-singular $N \in \mathbb{R}^{k\times k}$, we can define the equivalence relation $M \cong M'$ iff $M = M'N$ for some non-singular $N \in \mathbb{R}^{k\times k}$, and represent $\mathbf{Gr}(k,\mathbb{R}^n)$ as the quotient set $\mathbb{R}^{n\times k}/\cong$.

A symmetric Riccati differential equation (in $X = X(t)$) has the form

$$\dot{X} = XDX - XA - A^T X - C,$$

where $A,C,D,X \in \mathbb{R}^{n\times n}$, and in particular C,D,X are symmetric and positive definite.

Example 3: Grassmannian spaces and Riccati differential equations (cont.)

Setting $X(t) =: N(t)M^{-1}(t)$ for some $N(t)$ and $M(t)$, and noting that $\frac{d}{dt}(AB) = \dot{A}B + A\dot{B}$ and

$$MM^{-1} = I \quad \Rightarrow \quad \dot{M}M^{-1} + M\frac{d(M^{-1})}{dt} = 0 \quad \Rightarrow \quad \frac{d(M^{-1})}{dt} = -M^{-1}\dot{M}M^{-1},$$

we can write the Riccati equation as

$$\begin{aligned} \dot{N}M^{-1} - NM^{-1}\dot{M}M^{-1} &= NM^{-1}DNM^{-1} - NM^{-1}A - A^T NM^{-1} - C \\ \Leftrightarrow \quad N^{-1}\dot{N} - M^{-1}\dot{M} &= M^{-1}DN - M^{-1}AM - N^{-1}A^T N - N^{-1}CM \\ \Leftrightarrow \quad N^{-1}(\dot{N} + A^T N + CM) &= M^{-1}(\dot{M} + DN - AM). \end{aligned}$$

Grouping those terms starting with M^{-1} and those starting with N^{-1}, we can force this equation to hold by solving these two separate equations:

$$\begin{aligned} \dot{N} + A^T N + CM &= 0 \\ \dot{M} + DN - AM &= 0 \end{aligned} \quad \Leftrightarrow \quad \begin{bmatrix} \dot{N} \\ \dot{M} \end{bmatrix} = \begin{bmatrix} -A^T & -C \\ -D & A \end{bmatrix} \begin{bmatrix} N \\ M \end{bmatrix}.$$

Example 3: Grassmannian spaces and Riccati differential equations (cont.)

Notice that, using the equivalence relation defined previously,

$$\begin{bmatrix} X \\ I \end{bmatrix} = \begin{bmatrix} NM^{-1} \\ I \end{bmatrix} = \begin{bmatrix} N \\ M \end{bmatrix} M^{-1} \cong \begin{bmatrix} N \\ M \end{bmatrix},$$

so the solution X of the Riccati equation can be identified with an element of $\mathbf{Gr}(n, \mathbb{R}^{2n})$, and its evolution is dictated by a linear differential equation on a Grassmannian space!

$$\begin{bmatrix} \dot{N} \\ \dot{M} \end{bmatrix} = \begin{bmatrix} -A^T & -C \\ -D & A \end{bmatrix} \begin{bmatrix} N \\ M \end{bmatrix}.$$

Order Relations

A *partial order* on a set X is a relation $\leqslant$ on $X \times X$ s.t., for all $x, y, z \in X$,

a) $x \leqslant x$, (*reflexivity*)
b) $x \leqslant y$ and $y \leqslant x \quad \Rightarrow \quad x = y$, (*anti-symmetry*)
c) $x \leqslant y$ and $y \leqslant z \quad \Rightarrow \quad x \leqslant z$. (*transitivity*)

$x \leqslant y$ can also be written as $y \geqslant x$. $(X, \leqslant)$ is a *partially ordered set* (or *poset*).
If all $x, y \in X$ are *comparable*, *i.e.*, either $x \leqslant y$ or $y \leqslant x$ (or both), then $\leqslant$ is a *linear* (or *total*) *order*, and $(X, \leqslant)$ is called a *chain* (or *totally ordered set*).

Examples

- $(\mathbb{R}, \leqslant)$ is a totally ordered set.
- If A is a set and $X \subseteq \mathscr{P}(A)$, then $(X, \subseteq)$ is a partially ordered set.
- $\mathbb{R}^n$ is a partially ordered set with *element-wise ordering* (*i.e.*, given $x, y \in \mathbb{R}^n$, $x \leqslant y$ iff $x_i \leqslant y_i$ for all $i = 1, \dots, n$).
- If $\mathbb{S}^n$ is the set of real symmetric $n \times n$ matrices, $A, B \in \mathbb{S}^n$ and $A \preceq B$ iff $B - A$ is positive semi-definite, then $(\mathbb{S}^n, \preceq)$ is a partially (but not totally) ordered set.
- *Majorization*: For $x, y \in \mathbb{R}^n$, $x \preceq y$ iff $\sum_{i=1}^k \bar{x}_i \leqslant \sum_{i=1}^k \bar{y}_i$ for all $k = 1, \dots, n$, where $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$ is a reordering of the entries of x in descending order.

Order Relations (cont.)

Distinguished elements of a poset $(X, \leqslant)$

- If $a \in X$ is s.t. $a \leqslant x$ ($x \leqslant a$) for all $x \in X$, it is the *least/minimum* (*greatest/maximum*) element of X; it is unique, but not guaranteed to exist.
- If every non-empty $S \subseteq X$ has a minimum element, then X is *well ordered*.
- An $a \in X$ s.t. $x \leqslant a$ ($a \leqslant x$) for $x \in X$ only if $x = a$ is a *minimal* (*maximal*) element of X.
- An $a \in X$ s.t. $a \leqslant x$ ($x \leqslant a$) for all $x \in E \subseteq X$ is a *lower* (*upper*) *bound* of the set E in X.
- The *supremum* (*infimum*) $a \in X$ of a set $E \subseteq X$, denoted $\sup E$ ($\inf E$), is its least upper bound (greatest lower bound); it is not guaranteed to exist for general posets.

In a linearly ordered set, the concepts of minimal and least (maximal and greatest) coincide.

Even in a chain, a set may have a supremum (infimum) but no maximum (minimum). *E.g.*, the set $(0,1) \subseteq \mathbb{R}$ has infimum 0 and supremum 1, but no minimum nor maximum, since 0 and 1 do not belong to that interval.

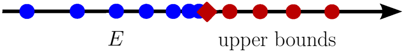
Least Upper Bound Property of $\mathbb{R}$

A set $E \subseteq \mathbb{R}$ is *bounded from above* (*below*) if there is an $M \in \mathbb{R}$ s.t. $x \leqslant M$ ($M \leqslant x$) for all $x \in E$; M is an *upper* (*lower*) *bound* of E.

E is *bounded* if it is bounded from above and bounded from below.

In contrast to $\mathbb{Q}$ (set of rational numbers), $\mathbb{R}$ possesses the *least upper bound property*:

"If $E \neq \emptyset$ is bounded from above, then it has a *supremum* in $\mathbb{R}$, $\sup E$."



The least upper bound property of $\mathbb{R}$ implies:

- The *completeness* of $\mathbb{R}$ (*see bonus slides*).
- The *Archimedean property* of $\mathbb{R}$ (and of $\mathbb{Q}$):

  "For every positive $x \in \mathbb{R}$, there is an $n \in \mathbb{N}$ s.t. $x > 1/n$."

  (Otherwise, if $n \leqslant 1/x$ for all $n \in \mathbb{N}$, then $\sup \mathbb{N} \leqslant 1/x < \infty$, so by definition of sup, there is an $m \in \mathbb{N}$ s.t. $\sup \mathbb{N} - 1 \leqslant m \leqslant \sup \mathbb{N}$, but then $\sup \mathbb{N} \leqslant m + 1$, a contradiction!)

For a function $f : X \to \mathbb{R}$, we denote $\sup_{x \in X} f(x) := \sup f(X)$ and $\inf_{x \in X} f(x) := \inf f(X)$.

A real sequence $(x_n)_{n\in\mathbb{N}}$ is *monotone* if it is *non-decreasing* ($x_n \leqslant x_{n+1}$ for all $n \in \mathbb{N}$) or *non-increasing* ($x_n \geqslant x_{n+1}$ for all $n \in \mathbb{N}$).

Theorem. A real monotone sequence $(x_n)_{n\in\mathbb{N}}$ has a limit in $[-\infty,\infty]$ ($:= \mathbb{R} \cup \{\pm\infty\}$).

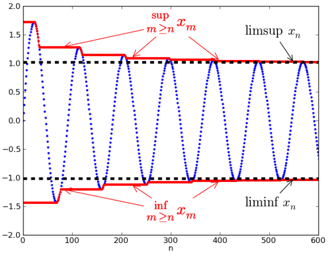
Proof. Assume w.l.o.g. that $(x_n)_{n\in\mathbb{N}}$ is non-decreasing, and consider the set $X = \{x_n : n \in \mathbb{N}\}$. If X is not bounded from above, then for every $M > 0$ there is an $x_N \geqslant M$, so $x_n \geqslant M$ for all $n \geqslant N$; thus $\lim_{n\to\infty} x_n = +\infty$. Otherwise, set $x = \sup X$. For every $\varepsilon > 0$ there is an $N \in \mathbb{N}$ s.t. $x - \varepsilon < x_N \leqslant x$, and hence $x - \varepsilon < x_n \leqslant x$ for all $n \geqslant N$; thus, $x = \lim_{n\to\infty} x_n$. □

Given a real sequence $(x_n)_{n\in\mathbb{N}}$, $(\sup_{m\geqslant n} x_m)_{n\in\mathbb{N}}$ is non-increasing and $(\inf_{m\geqslant n} x_m)_{n\in\mathbb{N}}$ is non-decreasing, hence

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty}\left(\sup_{m\geqslant n} x_m\right),$$

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty}\left(\inf_{m\geqslant n} x_m\right) \text{ exist (in } [-\infty,\infty]).$$

In fact, $\lim_{n\to\infty} x_n$ exists iff $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$.



Axiom of Choice, Zorn's Lemma, ...

Sometimes one needs to appeal to a non-constructive axiom (or one of its equivalents).

A *choice function* for a set X is a function f that assigns to each non-empty subset E of X an element of E: $f(E) \in E$.

- **Axiom of Choice**: For every set there is a choice function.
- **Zorn's Lemma**: Suppose that every chain inside a non-empty partially ordered set X has an upper bound in X. Then X has at least one maximal element.
- **Hausdorff's Maximal Principle**: Every partially ordered set contains a maximal chain (*i.e.*, one which is not contained in a larger chain).
- **Well-Ordering Theorem**: Every set can be well-ordered, *i.e.*, an order relation can be found for a set s.t. the resulting poset is well-ordered.

These axioms cannot be proven/disproven from the other axioms of standard (Zermelo-Fraenkel) set theory, and they are equivalent to each other (*see bonus slides*).

Note. These axioms are implicitly used in many derivations in analysis, *e.g.*, whenever one creates an infinite sequence without directly specifying its elements. However, in some proofs we will explicitly appeal to them.

Application: Existence of bases of vector spaces

In a vector space V, a *Hamel basis* (or simply, *basis*) is a linearly independent subset $B \subseteq V$ (*i.e.*, every *finite* subset of B is linearly independent), s.t. every $x \in V$ can be written as a linear combination of *finitely* many elements of B.

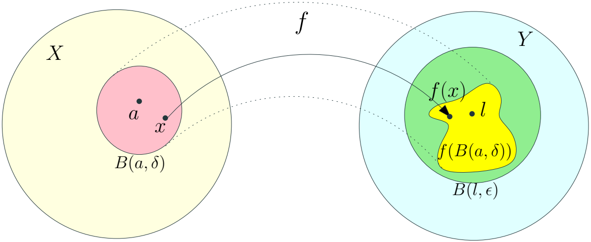
Theorem. Every vector space $V \neq \{0\}$ has a Hamel basis.

Proof. Let M be the set of all linearly independent subsets of V. Since $V \neq \{0\}$, there is an element $x \neq 0$ in V, so $\{x\} \in M$ and hence M is non-empty. Set inclusion defines a partial order in M. Every chain in M has an upper bound in M, namely, the union of all the elements in the chain (*why is this set in M?*). By Zorn's lemma, M has a maximal element B. We will show that B is a Hamel basis for V. Indeed, since $B \in M$, it is a linearly independent subset of V. Also, let Y be the set of all elements of V which can be written as finite linear combinations of elements of B. Then, $Y = V$, since otherwise there would be a $z \in V \setminus Y$, and $B \cup \{y\}$ would be a linearly independent set larger than B, contradicting the maximality of B. □

Outline

Motivation

Let $f : \mathbb{R}^n \to \mathbb{R}^n$.



$$\lim_{x \to a} f(x) = l$$

$\Leftrightarrow$ "For every $\varepsilon > 0$ there exists a $\delta > 0$ s.t. $\|x - a\| < \delta \quad \Rightarrow \quad \|f(x) - l\| < \varepsilon$."

$\Leftrightarrow$ "For every $B(l, \varepsilon)$ there exists a $B(a, \delta)$ s.t. $f(B(a, \delta)) \subseteq B(l, \varepsilon)$."

Exact distances do not matter for limits, only balls $B(x_0, r) := \{x \in \mathbb{R}^n : \|x - x_0\| < r\}$ *!*

In fact, we can replace the balls by more general sets: *open sets*!

Point Set Topology (cont.)

Definition

Let X be a set. A *topology* τ on X is a collections of subsets of X s.t.

a) $\emptyset, X \in \tau$,

b) if $(U_\alpha)_{\alpha \in I}$ is a family of elements of τ, then $\bigcup_{\alpha \in I} U_\alpha \in \tau$, (arbitrary unions)

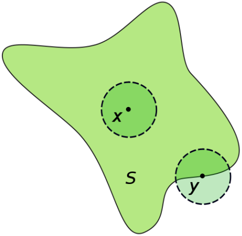
c) if $U_1, \ldots, U_n \in \tau$, then $\bigcap_{i=1}^{n} U_i \in \tau$. (finite intersections)

The elements of τ are called *open sets* of the *topological space* (X, τ).
The subsets $K \subseteq X$ s.t. K^c is open are called *closed sets*.
A *neighborhood* (*nbd*) of a point $x \in X$ is an open set containing x.

Interior, closure, boundary

Given a set $S \subseteq X$:

- *Interior* of S (int S): set of all the points in S which have a nbd contained in S.
- *Closure* of S ($\overline{S}$): set of all points x in X s.t. every nbd of x contains at least one point in S.
- *Boundary* of S (∂S): $\partial S := \overline{S} \setminus \text{int } S$.



Exercise: Show that $\overline{S}$ is the intersection of all closed sets containing S (*i.e.*, the smallest closed set containing S), and that int S is the union of all open sets contained in S.
Exercise: Prove that a set S is open iff $S = \text{int } S$, and that S is closed iff $S = \overline{S}$.

Limits of sequences
A sequence (x_n) in X is *convergent* if there is an $x \in X$ (*limit* of (x_n)) s.t for every nbd U of x, there is an $N \in \mathbb{N}$ s.t. $x_n \in U$ for all $n \geqslant N$. Here we write "$x_n \to x$".

Limits (topological version)
Given topological spaces (X, τ_X), (Y, τ_Y), and $f : X \to Y$:

$$\lim_{x \to a} f(x) = l \quad \overset{\text{def}}{\Longleftrightarrow} \quad \text{for every nbd } V \text{ of } l \text{, there is a nbd } U \text{ of } a \text{ s.t. } f(U \setminus \{a\}) \subseteq V$$

$$\Longrightarrow \quad \text{for every } (x_n) \text{ in } X \setminus \{a\} \text{ s.t. } x_n \to a \text{, it holds that } f(x_n) \to l.$$

f is *continuous* at $a \in X$ if $\lim_{x \to a} f(x) = f(a)$, and it is *continuous* if this holds for every $a \in X$.

In general: $\boxed{f : X \to Y \text{ is } \textit{continuous} \text{ iff } f^{-1}(V) \text{ is open for every open set } V \subseteq Y.}$ (*why?*)

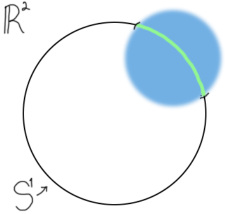
Exercise: Prove that if (x_n) lies in a closed set $K \subseteq X$, and converges to $x \in X$, then $x \in K$.

Point Set Topology (cont.)

$x \in X$ is an *accumulation* (*cluster* or *limit*) point of a subset A of a topological space X if every nbd of x contains points of A other than x (**note**: x does not need to belong to A).
In fact, $A \subseteq X$ is closed iff it contains the set of its accumulation points (*why?*).
Also, the closure of $A \subseteq X$ is the union of A and all of its accumulation points (*why?*).

Relative topology

If (X, τ) is a topological space and $A \subseteq X$, we can construct a topology ν for A, called the *relative topology* of τ to A, as follows: $U \in \nu$ iff $U = V \cap A$ for some $V \in \tau$.
(A, ν) is called a *subspace* of (X, τ).



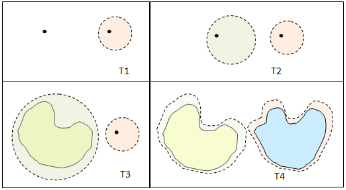
A property of a set $A \subseteq X$ is said to be *topological* if it can be defined only in terms of A and its relative topology; we will see later two important topological properties: compactness and connectedness.

A *homeomorphism* is a function $f : A \to B$ between topological spaces, which is continuous and bijective, with a continuous inverse; homeomorphisms preserve topological properties.

Point Set Topology (cont.)

We typically restrict attention to topological spaces (X,τ) satisfying additional properties:

Separation axioms



T_1: For every distinct $x,y \in X$, there is a nbd of x not containing y.

T_2: (*Hausdorff*) For every distinct $x,y \in X$, there are disjoint nbd's of x and y.

T_3: T_1 + *regular* (*i.e.*, for all $x \in X$ and closed set C not including x, there are disjoint open sets U,V s.t. $x \in U$ and $C \subseteq V$).

T_4: T_1 + *normal* (*i.e.*, for every disjoint closed sets $C,D \subseteq X$, there are disjoint open sets $U,V \subseteq X$ s.t. $C \subseteq U$ and $D \subseteq V$).

Axioms of countability

- *First-countable*: Every $x \in X$ has a *countable nbd base* (*i.e.*, a sequence $N_1 \supseteq N_2 \supseteq \cdots$ of nbd's of x s.t. for every nbd U of x there is an $N_i \subseteq U$).
- *Second-countable*: There is a *countable base* (*i.e.*, a countable family of open sets $\mathscr{U} = (U_i)$ s.t. every open set U is the union of some subfamily of $\mathscr{U}$).

Most topologies in the course are Hausdorff and first-countable, except for weak*-convergence (which is Hausdorff but not first-countable in general).

Examples of topologies

Our best source of examples will come from metric spaces, but here are simple examples:

- Given a set X, suppose we regard *every* subset of X as open. This leads to the *discrete topology* of X: every subset is both open and closed, so every subset coincides with its closure. This space is T_4.
- As another extreme example, given a set X, let $\tau = \{\emptyset, X\}$ (*trivial*/*indiscrete topology*); the closure of every non-empty subset is X.
- Let $X = \{a, b\}$ and $\tau = \{\emptyset, \{b\}, \{a, b\}\}$; the closed sets are then $\emptyset$, X and $\{a\}$, and the closure of $\{b\}$ is X. This space is not T_1 (*why?*).
- Consider $[0, 1]$ with the usual topology. The family of open sets $((x - 1/n, x + 1/n))_{x \in \mathbb{Q} \cap [0,1], n \in \mathbb{N}}$ is countable and every open set is a union of sets from this family, so $[0, 1]$ is second-countable. Also, given $x, y \in [0, 1]$, $x < y$, $(x - \varepsilon, x + \varepsilon) \cap [0, 1]$ and $(y - \varepsilon, y + \varepsilon) \cap [0, 1]$ (with $\varepsilon = (y - x)/2$) are disjoint nbd's of x and y, so $[0, 1]$ is Hausdorff.

Sequential definitions

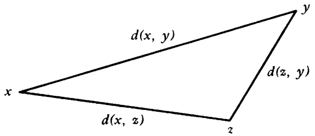
For first-countable spaces (which include metric, normed and inner-product spaces), several topological definitions can be re-written in terms of sequences. For example:

1. $\text{int } S = \{x \in S : \text{for every } x_n \to x \text{ there is an } N \in \mathbb{N} \text{ s.t. } x_n \in S \text{ if } n \geqslant N\}$ (why?).
2. $x \in X$ is a limit point of S iff there exists a sequence in $S \setminus \{x\}$ which converges to x, *i.e.*, x can be approximated arbitrarily well by elements of S other than itself (*why?*).
3. $\overline{S} = \{x \in X : \text{there is a sequence in } S \text{ with limit } x\}$, *i.e.*, $\overline{S}$ contains S and all its limit points (*why?*).
4. S is closed iff every convergent sequence in S has a limit in S (*why?*).
5. $\lim_{x \to a} f(x) = l$ iff for every sequence (x_n) in X with limit a, it holds that $f(x_n) \to l$ (*why?*).
6. $f : X \to Y$ is continuous at a iff $f(x_n) \to f(a)$ for every sequence (x_n) in X with limit a (*why?*).

Given a set X, a *metric* d is a function $d: X \times X \to \mathbb{R}_0^+$ s.t. for all $x, y, z \in X$,

a) $d(x, y) = 0$ iff $x = y$,
b) $d(x, y) = d(y, x)$ (*symmetry*),
c) $d(x, z) \leqslant d(x, y) + d(y, z)$ (*triangle inequality*).

(X, d) is a *metric space*.



A metric space defines a topology based on its *balls* $B(x, \lambda) := \{y \in X : d(x, y) < \lambda\}$.
A set U is *open* if for every point $x \in U$, there is a ball $B(x, \lambda) \subseteq U$ for some $\lambda > 0$.
A set U is *bounded* if there is an $M > 0$ s.t. $d(x, y) < M$ for all $x, y \in U$.

Examples

1. Real line: $X = \mathbb{R}$, $d(x, y) = |x - y|$.
2. Euclidean space: $X = \mathbb{R}^n$, $d(x, y) = \|x - y\|_2 = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$.

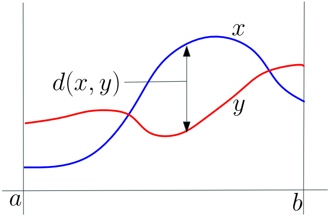
Examples (cont.)

3. Function space: $X = C[a,b] := C([a,b]) := \{f : [a,b] \to \mathbb{R} : f \text{ is continuous}\}$

$$d(f,g) = \max_{t\in[a,b]} |f(t) - g(t)|.$$

Proof of triangle inequality: Let $x, y, z \in C[a,b]$, then

$$\begin{aligned} d(x,z) &= \max_{t\in[a,b]} |x(t) - z(t)| \\ &= \max_{t\in[a,b]} |x(t) - y(t) + y(t) - z(t)| \\ &\leqslant \max_{t\in[a,b]} \{|x(t) - y(t)| + |y(t) - z(t)|\} \\ &\leqslant \max_{t\in[a,b]} |x(t) - y(t)| + \max_{r\in[a,b]} |y(r) - z(r)| \\ &= d(x,y) + d(y,z). \end{aligned}$$



□

The topology of a metric space is always T_4 and first-countable (*home work!*).

Notation. In general, if X, Y are topological spaces, $C(X,Y)$ is the space of continuous functions from X to Y; in particular, $C(X) := C(X,\mathbb{R})$ (or $:= C(X,\mathbb{C})$, depending on context).

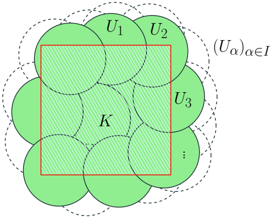
Compactness

Goal

Generalize the result that a continuous $f:[a,b]\to\mathbb{R}$ achieves its maximum in $[a,b]$.

Definition

A set $K\subseteq X$ is *compact* if for every family of open sets $(U_\alpha)_{\alpha\in I}$ s.t. $K\subseteq\bigcup_{\alpha\in I}U_\alpha$ (*open cover of* K), there is a finite subfamily $(U_1,\dots,U_n)$ of $(U_\alpha)_{\alpha\in I}$ (*sub-cover*) s.t. $K\subseteq\bigcup_{i=1}^n U_i$.



Alternative characterization (*finite intersection property*)

If $(C_\alpha)_{\alpha\in I}$ is a family of closed subsets of K s.t. every finite number of them has a nonempty intersection, then $\bigcap_{\alpha\in I}C_\alpha$ is nonempty (*home work!*).

Sequential compactness (equivalent to compactness for metric spaces; *see bonus slides*)

Every sequence $(x_n)_{n\in\mathbb{N}}$ in K has a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with limit in K.

Heine-Borel Theorem (*see bonus slides for proof*)

$K\subseteq\mathbb{R}^n$ is compact iff it is closed and bounded.

Compactness (cont.)

Examples of non-compact sets

- $\mathbb{Q}$ *is non-compact*: it is not bounded (take $(x_n)_{n\in\mathbb{N}}$ with $x_n = n$: no subsequence is convergent (*why?*)). It is also not closed in $\mathbb{R}$, since every nbd of an irrational number contains rational numbers, so $\mathbb{Q} \neq \overline{\mathbb{Q}} = \mathbb{R}$; this implies also that no bounded interval in $\mathbb{Q}$ is compact either.
- $(0,1)$ *is not compact*: consider the open sets $U_n = (0, 1-1/n) \subseteq (0,1)$, $n \in \mathbb{N}$. Every $x \in (0,1)$ belongs to at least some U_n (just take n large enough so that $1 - 1/n > x$, or $n > 1/(1-x)$, hence $\bigcup_{n\in\mathbb{N}} U_n = (0,1)$. However, no finite subfamily covers $(0,1)$: assume $(U_{n_i})_{i=1,\dots,N}$ is such a subcover, where $U_{n_1} \subseteq U_{n_2} \subseteq \cdots \subseteq U_{n_N}$; but then $\bigcup_{i=1}^N U_{n_i} = U_{n_N}$, which does not contain $[1-1/n_N, 1) \subseteq (0,1)$.

  From the sequential characterization perspective, notice that the sequence with $x_n = 1 - 1/(n+1) \in (0,1)$ is convergent in $\mathbb{R}$ to $1 \notin (0,1)$, so every subsequence is also convergent to 1, *i.e.*, there is no convergent subsequence in $(0,1)$.

Compactness and continuity

Let $f : X \to Y$ be continuous, where X and Y are topological spaces. Then, if $K \subseteq X$ is compact, so is $f(K)$.

Proof. Let $(V_\alpha)_{\alpha \in I}$ be an open cover of $f(K)$. Let $U_\alpha := f^{-1}(V_\alpha)$ for all $\alpha \in I$. Then $(U_\alpha)_{\alpha \in I}$ is an open cover of K. Since K is compact, there is a finite subcover, say, $U_{\alpha_1}, \ldots, U_{\alpha_n}$. Then $V_{\alpha_1}, \ldots, V_{\alpha_n}$ is a finite subcover of $f(K)$. This means that $f(K)$ is compact. □

Corollary (Weierstrass' theorem)

Let $f : K \to \mathbb{R}$ be continuous, where K is compact. Then f achieves its maximum.

Proof. We know that $f(K) \subseteq \mathbb{R}$ is compact, *i.e.*, closed and bounded. Since $f(K)$ is bounded, let $M = \sup f(K) < \infty$, and consider a sequence (y_n) in $f(K)$ converging to M. As $f(K)$ is compact, $M \in f(K)$. Then there is an $x \in f^{-1}(K)$, which achieves the maximum. □

Compactness and uniform continuity (Heine-Cantor)

Given metric spaces (X, d_X) and (Y, d_Y), if X is compact and $f : X \to Y$ is continuous, then f is *uniformly continuous*, *i.e.*, given $\varepsilon > 0$, there is a $\delta > 0$ s.t. for all $x, y \in X$ s.t $d_X(x, y) < \delta$, we have $d_Y(f(x), f(y)) < \varepsilon$ (*see bonus slides for proof*).

Application: Fundamental Theorem of Algebra

Let $P(z) = \sum_{k=0}^{n} a_k z^k$, where $n \in \mathbb{N}$ and $a_0, \dots, a_n \in \mathbb{C}$, $a_n \neq 0$. Then $P(z) = 0$ for some $z \in \mathbb{C}$.

Proof. Assume w.l.o.g. that $a_n = 1$, and let $\mu = \inf\{|P(z)| : z \in \mathbb{C}\} \in \mathbb{R}$. If $|z| = R$, then

$$|P(z)| \geqslant R^n [1 - |a_{n-1}|R^{-1} - \cdots - |a_0|R^{-n}] \xrightarrow{R \to \infty} \infty,$$

so there is an R_0 s.t. $|P(z)| > \mu$ if $|z| > R_0$. Since $z \mapsto |P(z)|$ is continuous on $\{z \in \mathbb{C} : |z| \leqslant R_0\}$, which is compact, it follows that $|P(z_0)| = \mu$ for some z_0. Assume that $\mu > 0$. Let $Q(z) = P(z + z_0)/P(z_0)$, which satisfies $Q(0) = 1$, $|Q(z)| \geqslant 1$, and is of the form $Q(z) = 1 + b_k z^k + \cdots + b_n z^n$, where $b_k \neq 0$. Choose $\theta \in \mathbb{R}$ and $r > 0$ so that $e^{ik\theta} b_k = -|b_k|$ and $r^k |b_k| < 1$. Then, $|1 + b_k r^k e^{ik\theta}| = 1 - r^k |b_k|$, so

$$|Q(re^{i\theta})| \leqslant 1 - r^k \{|b_k| - r|b_{k+1}| - \cdots - r^{n-k}|b_n|\}.$$

Thus, for r small enough, $|Q(re^{i\theta})| < 1$: a contradiction. Thus, $\mu = 0$ and hence $P(z_0) = 0$. □

Compactness (cont.)

Example: Arzelà-Ascoli theorem and differential equations
The concept of compactness can be well characterized for some function spaces:

$\mathscr{F} \subseteq C(X)$, where X is a metric space, is:

- *pointwisely bounded* if for every $x \in X$ there is an $M_x > 0$ s.t. $|f(x)| \leqslant M_x$ for all $f \in \mathscr{F}$;
- *equicontinuous* if for every $\varepsilon > 0$ and $x \in X$ there is a $\delta > 0$ s.t. for all $f \in \mathscr{F}$ and $y \in X$ for which $d(x, y) < \delta$, it holds that $|f(x) - f(y)| < \varepsilon$.

Theorem (Arzelà-Ascoli) (*See bonus slides for proof*)
Let X be a compact metric space. Then, every sequence in $\mathscr{F} \subseteq C(X)$ has a uniformly convergent subsequence in $C(X)$ iff $\mathscr{F}$ is pointwisely bounded and equicontinuous.

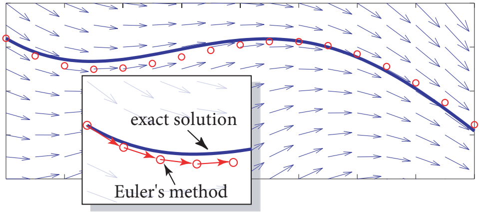
The next result uses Arzelà-Ascoli's theorem to establish existence of solutions to ODEs:

Cauchy-Peano existence theorem. Let $I = [t_0, t_0 + \beta] \subseteq \mathbb{R}$, $\Omega = \overline{B(x^0, r)} \subseteq \mathbb{R}^n$, and suppose $f : I \times \Omega \to \mathbb{R}^n$ is continuous. Then there is a solution to the ODE $\dot{x}(t) = f(t, x(t))$, $x(t_0) = x^0$, on $C([t_0, t_0 + \alpha])$, where $\alpha = \min\{\beta, r/M\}$ and $M = \max_{(t,x) \in I \times \Omega} |f(t, x)|$.

Compactness (cont.)

Example: Arzelà-Ascoli theorem and differential equations (cont.)

Proof. The idea is to construct approximations to the solution via forward Euler's method:



For each $k \in \mathbb{N}$, partition $I_\alpha := [t_0, t_0 + \alpha]$ into k subintervals of equal length α/k. Set $x(t_0) = x^0$, and inductively define $x_k(t) = x_k(t_{l-1}) + f(t_{l-1}, x_k(t_{t-1}))(t - t_{l-1})$ on (t_{l-1}, t_l), where $t_l = t_0 + \alpha l/k$. One can check that $\|x_k(t) - x_k(\tau)\| \leqslant M|t - \tau|$ for $t, \tau \in I_\alpha$, so $\|x_k(t) - x^0\| \leqslant r$ on I_α. This bound also implies that $\{x_k\}_{k \in \mathbb{N}}$ is equicontinuous and pointwisely bounded, hence by Arzelà-Ascoli's theorem there exists a $x_* \in C(I_\alpha)$ and a subsequence $(x_{k_i})_{i \in \mathbb{N}}$ converging uniformly to x_*.

Example: Arzelà-Ascoli theorem and differential equations (cont.)

It remains to prove that x_* solves the ODE. Since x_k is piecewise linear, it can be written as

$$x_k(t) = x^0 + \int_{t_0}^{t} \dot{x}_k(\tau) d\tau.$$

Define $\Delta_k(t) := \dot{x}_k(t) - f(t, x_k(t))$ on $I_\alpha \setminus \{t_0, \dots, t_k\}$, and $\Delta_k(t_l) = 0$ for $l = 0, \dots, k$. Observe that $\max_{t \in I_\alpha} \|\Delta_k(t)\| \xrightarrow{k \to \infty} 0$, due to the uniform continuity of f on the compact set $I_\alpha \times \Omega$ (by Heine-Cantor), and that since $|t - t_{l-1}| < \alpha/k$, $\|x_k(t) - x_k(t_{l-1})\| \leqslant M|t - t_{l-1}| < M\alpha/k$, and for $k \in \mathbb{N}$, $l \in \{1, \dots, k\}$ and $t \in (t_{l-1}, t_l)$, $\|\dot{x}_k(t) - f(t, x_k(t))\| = \|f(t_{l-1}, x_k(t_{l-1})) - f(t, x_k(t))\|$. Thus,

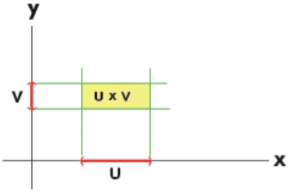
$$x_{k_i}(t) = x^0 + \int_{t_0}^{t} \dot{x}_{k_i}(\tau) d\tau = x^0 + \int_{t_0}^{t} [f(\tau, x_{k_i}(\tau)) + \Delta_{k_i}(\tau)] d\tau.$$

Taking $i \to \infty$ and using the uniform continuity of f yields $x_*(t) = x_0 + \int_{t_0}^{t} f(\tau, x_*(\tau)) d\tau$, so x_* satisfies the ODE. $\square$

Product Topologies

Product topologies

To define continuity of operations like sum and product, we need to define the topology of Cartesian products $X \times Y$, where (X, τ_X) and (Y, τ_Y) are topological spaces. The standard way is to define the *product topology*: a set $W \subseteq X \times Y$ is open iff for every $(x, y) \in X \times Y$ there are nbd's of x and y, $U \in \tau_X$ and $V \in \tau_Y$, s.t. $U \times V \subseteq W$.



This definition extends naturally to the product of a finite number of topological spaces.

See bonus slides for extensions to arbitrary families of topological spaces.

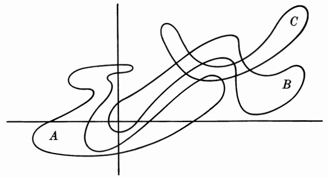
Connectedness

Another important topological notion is the property that a set may or not be decomposed into separate pieces. This will allow us to generalize the mean value theorem.

Definition. A topological space X is *connected* if it is not the union of 2 disjoint non-empty open sets.

Theorem. $E \subseteq \mathbb{R}$ is connected iff:
for every $x, y \in E$, $z \in \mathbb{R}$, if $x < z < y$ then $z \in E$.

Proof. Assume there are $x, y \in E$ and $z \in \mathbb{R}$, s.t. $x < z < y$, but $z \notin E$. Then can be decomposed as $E = \{E \cap (-\infty, z)\} \cup \{E \cap (z, \infty)\}$, where the two sets in brackets are disjoint, open and non-empty, so E is disconnected.



$A, B, C, A \cup C, B \cup C, A \cup B \cup C$ are connected;

$A \cup B$ is not

Conversely, assume E is disconnected, *i.e.*, $E = A \cup B$, where A and B are disjoint, open and non-empty. Pick $a \in A$ and $b \in B$, assuming w.l.o.g. that $a < b$. Let $x = \sup A \cap [a, b]$. If $x \in A$, then there exists an $\varepsilon > 0$ s.t. $x + \varepsilon \notin B$ and $x + \varepsilon < b$ (since A is open), so $a < x + \varepsilon < b$ and $x + \varepsilon \notin E$; in case $x \notin A$, it cannot belong to B (since B is open, and every nbd of x contains points of A), so $a < x < b$ and $x \notin E$. $\square$

Thus, all connected subsets of $\mathbb{R}$ are of the form (a, b), $(a, b]$, $[a, b)$ or $[a, b]$!

Connectedness (cont.)

Connectedness and continuity

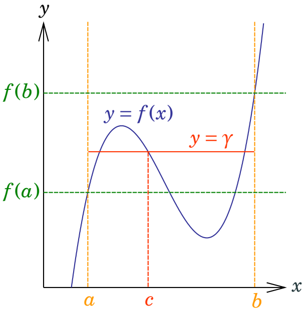
Let $f: X \to Y$ be continuous between topological spaces X and Y. Then, if $K \subseteq X$ is connected, so is $f(K)$.

Proof. If $f(K)$ were disconnected, then $f(K) = A \cup B$ for some disjoint non-empty open sets A and B. By continuity, $f^{-1}(A)$ and $f^{-1}(B)$ are open, disjoint and non-empty, and $K \subseteq f^{-1}(A) \cup f^{-1}(B)$, so K would be disconnected. $\square$

A direct consequence of the previous theorems is

Intermediate value theorem

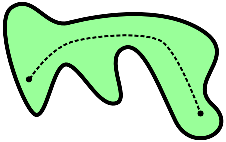
If $a, b \in \mathbb{R}$, $a < b$, and $f: [a,b] \to \mathbb{R}$ is continuous, then for every $\gamma \in \mathbb{R}$ between $f(a)$ and $f(b)$, there exists a $c \in (a,b)$ s.t. $f(c) = \gamma$.



This is the basis of the *bisection method* for solving $f(x) = 0$, which constructs a sequence of subintervals of decreasing length at whose ends f takes alternate signs; the intermediate value theorem guarantees that these intervals contain one root of f!

Connectedness (cont.)

A stronger concept of connectedness is *path-connectedness*:



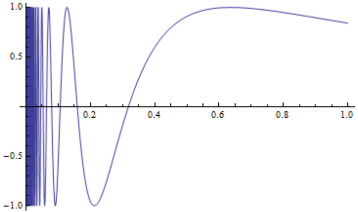
A topological space X is *path-connected* if, for every $x, y \in X$ there exists a *path* from x to y, *i.e.*, a continuous function $f: [0,1] \to X$ s.t. $f(0) = x$ and $f(1) = y$.

A path-connected space is necessarily connected (*home work!*), but not conversely.

Counterexample: "topologist's sine curve"

$$T = \{(x, \sin(1/x)) : x \in (0,1]\} \cup \{(0,0)\}.$$

(*Why is it a counterexample?*)



However, every *open* connected subset U of $\mathbb{R}^n$ is path-connected, and, moreover, a path between every two points in U made of a finite number of straight segments (contained in U) can always be found (*why?*).

Next Topic

Inner Product Spaces

Outline

Bonus: Equivalent Forms of the Axiom of Choice

Hausdorff's maximal principle ⇒ Zorn's lemma

Assume that every chain inside a partially ordered set X has an upper bound. By the maximal principle, X has a maximal chain C. Then C has an upper bound $m \in X$, but due to the maximality of C, $m \in C$ (otherwise m could be attached to C, contradicting its maximality), so it is the largest element of C. Furthermore, since C is a maximal chain, m is a maximal element of X, as otherwise there would exist an $m' \geqslant m$, $m' \neq m$, and $C \cup \{m'\}$ would be a larger chain than C. □

Zorn's lemma ⇒ Well-ordering theorem

Let X be a set, and consider the set S of all the well-ordered subsets of X. S contains $\emptyset$ and all singletons of X, so it is non-empty. If $A, B \in S$, we say that $A \leqslant B$ if (i) A is a subset of B, (ii) there is an element of B, x, s.t. $A = \{a \in B : a < x\}$, and (iii) for all $x, y \in A$, $x \leqslant y$ in A iff $x \leqslant y$ in B. For every chain $C \subseteq S$, $\bigcup C$ is an upper bound of C, so by Zorn's lemma, there is a maximal well-ordered set $E \in S$. Then, E contains all of X, since otherwise there would be an $x \in X \setminus E$, and $E \cup \{x\}$ would be a larger well-ordered set in S (where we define $e < x$ for all $e \in E$). This concludes the proof. □

Well-ordering theorem ⇒ Axiom of choice

Let X be a set. By the well-ordering theorem, X can be turned into a well-ordered poset, *i.e.*, every non-empty set $E \subseteq X$ has a minimum element. Let f be a function that assigns to each such E its minimum element; then, f is a choice function for X. This establishes the axiom of choice. □

Axiom of choice $\Rightarrow$ Hausdorff's maximal principle (*difficult bit*)
Let $(X, \leqslant)$, with $X \neq \emptyset$, be partially ordered, and consider the collection P of all chains in X, with set inclusion as partial order. If $C \in P$ is non-maximal, then there is a chain in X larger than C, so we can choose, using the axiom of choice, an $x \in X \setminus C$ s.t. $C \cup \{x\}$ is a chain; set $C^+ := C \cup \{x\}$. If C is maximal, set $C^+ := C$. Note also that the union of a chain $\Gamma \subseteq P$ (*i.e.*, a chain of chains in X) is again a chain in X, which corresponds to an upper bound of Γ.

A subset $N \subseteq P$ is said to be a *tower* if it satisfies the following properties:

(i) If $C \in N$, then $C^+ \in N$,

(ii) If $\Gamma \subseteq N$ is a chain, then $\bigcup \Gamma$ is in N.

Note that P itself is a tower, and that the intersection of a family of towers is a tower. In particular, the intersection of all tower subsets of P is the smallest tower; call it M.

We will show that M is a chain; with this fact, from property (ii) we know that $m = \bigcup M$ is the largest chain in M, and by property (i), $m^+ = m$, so m is a maximal chain in X.

Call an element $C \in M$ *comparable* if for every $D \in M$ either $D \subseteq C$ or $C \subseteq D$ holds. To prove that M is a chain, we must show that every element of M is comparable. For this we need the following lemmas:

Lemma 1. Assume $C \in M$ is comparable. If $D \in M$ and $D \subseteq C$ but $D \neq C$, then $D^+ \subseteq C$.
Proof. Assume $C \subseteq D^+$ but $C \neq D$. Then $D \subseteq C \subseteq D^+$ but C is different from D and D^+, contradicting the fact that D^+ is constructed by adjoining a single element of X to D. □

Lemma 2. Assume $C \in M$ is comparable. For every $D \in M$, either $D \subseteq C$ or $D \supseteq C^+$.
Proof. Let $N = \{D \in M : D \subseteq C \text{ or } D \supseteq C^+\}$; we will show that $N = M$. Since M is the smallest tower, it suffices to show that N is a tower. Given $D \in N$, we have that (1) $D \subseteq C$ but $D \neq C$, (2) $D = C$ or (3) $D \supseteq C^+$. In (1), $D^+ \subseteq C$ by Lemma 1, so $D^+ \in N$; in (2) and (3), $D^+ \supseteq C^+$ so again $D^+ \in N$. Next, assume $\Gamma \subseteq N$ is a chain, and let $E = \bigcup \Gamma$. If every $D \in \Gamma$ is a subset of C, then $E \subseteq C$ so $E \in N$; otherwise, some $D \in \Gamma$ contains C^+, so $E \supseteq C^+$ and again $E \in N$. Thus, both properties (i) and (ii) hold, hence N is a tower. □

Let M^* be the set of all comparable elements of M. We will show that M^* is a tower, so it has to be equal to M (as M is the smallest tower). If C is comparable and $D \in M$, by Lemma 2 either $D \subseteq C$ or $C^+ \subseteq D$. In both cases D is comparable to C^+, so C^+ is comparable. Next, assume $\Gamma \subseteq M^*$ is a chain, and let $D = \bigcup \Gamma$. Given $E \in M$, either $C \subseteq E$ for all $C \in \Gamma$, so $D \subseteq E$, or else $E \subseteq C$ for some $C \in \Gamma$, so $E \subseteq D$. Thus D is comparable, so M^* is indeed a tower. □

Bonus: Transfinite Induction

This is a generalization of mathematical induction to arbitrary well-ordered sets; recall that by the axiom of choice, every set can be well-ordered.

Transfinite Induction. Let X be a well-ordered set, and $P(x)$ a proposition defined for all $x \in X$. Assume that if $P(y)$ is true for all $y < x$, then $P(x)$ is true. Then, $P(x)$ is true for all $x \in X$.

Proof. Let E be the set of all $x \in X$ for which $P(x)$ is false, and assume that E is non-empty. Since X is well-ordered, E has a minimum element, say, e. Then, $P(y)$ is true for all $y < e$, so $P(e)$ must be true, contradicting the assumption that E is non-empty. This concludes the proof. □

Note. Transfinite induction does not require a base case ($n = 1$), but its inductive step needs $P(y)$ to hold for all $y < x$, whereas in ordinary induction the case "$n - 1$" suffices. Reason: there are well-ordered sets that contain elements without an "predecessor"; *e.g.*, in $X = \mathbb{N} \cup \{\mathbb{N}\}$, where $x < \mathbb{N}$ for all $x \in \mathbb{N}$, there is no $x \in X$ s.t. $x + 1 = \mathbb{N}$, so ordinary induction can never "reach" the last element in X.

If $X = \mathbb{N}$ with its usual order, transfinite induction corresponds to *complete*/*strong induction*.

This result can sometimes yield more intuitive proofs than Zorn's lemma.
Exercise: Prove by transfinite induction that every vector space has a Hamel basis.

Bonus: Compactness vs. Sequential Compactness

Theorem. The following statements are equivalent for a metric space K:

(1) K is compact.

(2) K is sequentially compact.

Proof

(1) $\Longrightarrow$ (2): Assume that $(x_n)_{n\in\mathbb{N}}$ is a sequence in K which has no convergent subsequence. Then, for every $i \in \mathbb{N}$, let U_{x_i} be a nbd of x_n s.t. $x_n \notin U_{x_i}$ for all $n \neq i$ (such a nbd exists since otherwise there would be a subsequence of $(x_n)_{n\in\mathbb{N}}$ converging to x_i). Also, for every $x \in K \setminus \{x_1, x_2, \dots\}$, let U_x be a nbd of x contained in $K \setminus \{x_1, x_2, \dots\}$ (such a nbd exists for the same reason as before). Thus, $(U_x)_{x\in K}$ is an open cover of K, and since K is compact by (1), there is a finite subcover $\{U_1, \dots, U_N\}$ of K; since each x_n can belong to at most one such U_i, the sequence $(x_n)_{n\in\mathbb{N}}$ can take at most N values, which contradicts the assumption that it has no convergent subsequence. Therefore, K is sequentially compact.

(2) $\Longrightarrow$ (1): For this we need the following preliminary results:

Lemma (Lebesgue's Number Lemma). Let K be a sequentially compact metric space, and $(U_\alpha)_{\alpha\in I}$ an open cover of K. Then, there is an $r>0$ (*Lebesgue number of the cover*) s.t. for every $x\in K$, $B(x,r)$ belongs to at least one U_α.

Proof. Assume that the claim is false, *i.e.*, that for every $n\in\mathbb{N}$, there exists an $x_n\in K$ s.t. $B(x_n,1/n)\cap U_\alpha^c\neq\emptyset$ for each $\alpha\in I$. Let x be the limit of some convergent subsequence of (x_n). Pick some $\alpha\in I$ s.t. $x\in U_\alpha$, and choose some $r>0$ for which $B(x,r)\subseteq U_\alpha$. Next, select n large enough so that $1/n<r/2$ and $d(x,x_n)<r/2$. It follows that $B(x_n,1/n)\subseteq B(x,r)\subseteq U_\alpha$, contrary to the selection of x_n. This establishes the result. □

Lemma. A sequentially compact metric space K is *totally bounded*, *i.e.*, for every $\varepsilon>0$ there is a finite set $\{x_1,\dots,x_N\}\subseteq K$ s.t. $K\subseteq\bigcup_{i=1}^N B(x_i,\varepsilon)$ (ε-*net* of K).

Proof. Assume there is an $\varepsilon>0$ for which no ε-net of K exists. Take some $x_1\in K$ and, inductively, let $x_{i+1}\in K\setminus\bigcup_{k=1}^i B(x_i,\varepsilon)$ for $i\in\mathbb{N}$. By construction, $d(x_i,x_j)>\varepsilon$ for all $i\neq j$, so there is no convergent subsequence of $(x_i)_{i\in\mathbb{N}}$, which contradicts the sequential compactness of K. □

(2) $\Longrightarrow$ (1) (cont.):

Let $(U_\alpha)_{\alpha\in I}$ be an open cover of K, and denote by r the Lebesgue number of this cover. Then, by total boundedness of K is a finite set of balls $\{B(x_i,r)\colon i=1,\dots,N\}$ which cover K, each of which is completely contained inside one U_α, say, $B(x_i,r)\subseteq U_{\alpha_i}$. Thus, $\{U_{\alpha_i}\colon i=1,\dots,N\}$ is a finite subcover of K, so K is compact. □

Bonus: Completeness of the Real Numbers

A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X,d) is said to be a *Cauchy sequence* if, for every $\varepsilon > 0$, there is an $N \in \mathbb{N}$ s.t. $d(x_n, x_m) < \varepsilon$ for every $n, m \geqslant N$.
(X,d) is a *complete* metric space if every Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ in X has a limit (in X).

Theorem. The real line, $(\mathbb{R}, d)$ with $d(x,y) = |x-y|$, is a complete metric space.
Proof. Take a real Cauchy sequence $(x_n)_{n\in\mathbb{N}}$. For $\varepsilon = 1$, there is an $n \in \mathbb{N}$ s.t. $|x_n - x_N| < 1$ for all $n \geqslant N$; thus $|x_n| \leqslant \max\{|x_1|, \dots, |x_{N-1}|, |x_N| + 1\}$ for all n, so $\limsup_{n\to\infty} x_n$ exists and is finite. Let $x = \limsup_{n\to\infty} x_n$. Then, given $\varepsilon = 1/k > 0$ (for some $k \in \mathbb{N}$), there is an $N_1 = N_1(k) \in \mathbb{N}$ s.t. $x \leqslant \sup_{m\geqslant N_1} x_m \leqslant x + \varepsilon$, so there is an $N_2 = N_2(k) \geqslant N_1$ for which

$$x - 1/k \leqslant \sup_{m\geqslant N_1} x_m - 1/k \leqslant x_{N_2} \leqslant \sup_{m\geqslant N_1} x_m \leqslant x + 1/k,$$

i.e., $|x_{N_2} - x| \leqslant 1/k$, hence there is a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ which converges to x. However, given $\varepsilon > 0$, there is an $N \in \mathbb{N}$ s.t. $|x_n - x_m| < \varepsilon/2$ for all $n, m \geqslant N$, hence picking N large enough so that $|x_{n_k} - x| < \varepsilon/2$ for all $n_k \geqslant N$ shows that $|x_n - x| \leqslant |x_n - x_{n_k}| + |x_{n_k} - x| < \varepsilon$ for all $n \geqslant N$; this means that $x_n \to x$. □

Bonus: Proof of Heine-Borel Theorem

$K \subseteq \mathbb{R}^n$ compact $\implies$ K closed and bounded:
Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in K with limit $x \in \mathbb{R}^n$. Since K is compact, there is a convergent subsequence $(x_{k_i})_{i\in\mathbb{N}}$ with limit in K; however, every subsequence should converge to the same limit, hence $x \in K$. Thus, by the second Exercise in Slide 34, K is closed.
If K were unbounded, one could create a sequence as follows: pick $x_1 \in K$, and for every $i > 1$, choose $x_i \in K$ s.t. $\|x_i\| \geqslant \|x_{i-1}\| + 1$. This sequence does not have a convergent subsequence, since $\|x_i - x_j\| \geqslant \big|\|x_i\| - \|x_j\|\big| \geqslant 1$ for all $i, j \in \mathbb{N}$, which contradicts the assumption that K is compact. Hence, K must be bounded.

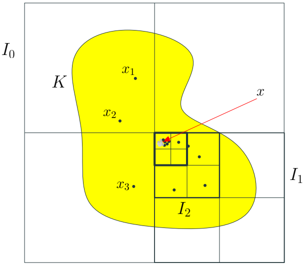
$K \subseteq \mathbb{R}^n$ closed and bounded $\implies$ K compact:
Consider a sequence $(x_k)_{k\in\mathbb{N}}$ in K. Since K is bounded, $K \subseteq [-M, M]^n =: I_0$ for some $M > 0$. Partition I_0 into 2^n hyper-cubes ($[-M, 0]^n$, $[0, M]^n$, *etc.*). At least one of these should contain an infinitude of x_k's; call such subset I_1. By further partitioning I_1 into hyper-cubes, and choosing one with an infinitude of x_k's as $I_2, \ldots$, we build a sequence of sets $I_0 \supset I_1 \supset \cdots$.

Bonus: Proof of Heine-Borel Theorem (cont.)

Define a subsequence $(x_{k_i})_{i\in\mathbb{N}}$ s.t. $x_{k_i} \in I_i$ for all i. The l-th component of the sequence, $(x^l_{k_i})_{i\in\mathbb{N}}$, satisfies $|x^l_{k_i} - x^l_{k_j}| \leqslant M \cdot 2^{-N}$ whenever $i, j \geqslant N$, so $(x^l_{k_i})_{i\in\mathbb{N}}$ is a Cauchy sequence, and hence convergent (since $\mathbb{R}$ is complete; see Slide 60); define $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ where $x^l = \lim_{i\to\infty} x^l_{k_i}$. Now,

$$\|x - x_{k_i}\|_2 = \sum_{l=1}^{n} \left(x^l - x^l_{k_i}\right)^2 \xrightarrow{i\to\infty} 0,$$

so $(x_{k_i})_{i\in\mathbb{N}}$ is convergent with limit $x \in \mathbb{R}^n$.



Furthermore, since K is closed, then, by the second Exercise in Slide 34, $x \in K$. Thus, every sequence in K has a convergent subsequence in K, so K is compact. □

Bonus: Proof of Heine-Cantor Theorem

Theorem (Heine-Cantor). Given metric spaces (X, d_X) and (Y, d_Y), if X is compact and $f : X \to Y$ is continuous, then f is uniformly continuous.

Proof
Assume that f is not uniformly continuous. Then there is an $\varepsilon > 0$ s.t. for all $n \in \mathbb{N}$ there are points $x_n, y_n \in X$ for which $d_X(x_n, y_n) < 1/n$ but $d_Y(f(x_n), f(y_n)) > \varepsilon$. Since X is compact, (x_n) has a subsequence (x_{n_k}) converging to, say, x. Since

$$d_X(y_{n_k}, x) \leqslant d_X(y_{n_k}, x_{n_k}) + d_X(x_{n_k}, x) \to 0 \quad \text{as } k \to \infty,$$

(y_{n_k}) is also converging to x. Therefore, due to the continuity of f,

$$d_Y(f(x_{n_k}), f(y_{n_k})) \leqslant d_Y(f(x_{n_k}), f(x)) + d_Y(f(x), f(y_{n_k})) \to 0,$$

which contradicts the fact that $d_Y(f(x_n), f(y_n)) > \varepsilon$ for all n. This proves that f is uniformly continuous. $\square$

Theorem (Arzelà-Ascoli)

Let X be a compact metric space. Then, every sequence in $\mathscr{F} \subseteq C(X)$ has a uniformly convergent subsequence in $C(X)$ iff $\mathscr{F}$ is pointwisely bounded and equicontinuous.

Proof.

($\Leftarrow$) Assume $\mathscr{F}$ is pointwisely bounded and equicontinuous. Since X is compact, there is a finite set of balls, $\{B(x_i^n, 1/n)\colon i = 1, \dots, K(n)\}$, which covers X for every $n \in \mathbb{N}$. The set $\{x_i^n\}_{i,n}$ is countable and *dense*, *i.e.*, for every $\varepsilon > 0$ and $x \in X$, there is an x_i^n s.t. $d(x, x_i^n) < \varepsilon$; let $D = (x_i)_{i\in\mathbb{N}}$ be an enumeration of $\{x_i^n\}$.

Consider a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathscr{F}$. We can construct a convergent subsequence on $C((x_i)_{i\in\mathbb{N}})$ as follows: Due to uniform boundedness, $(f_n(x_1))_{n\in\mathbb{N}}$ is bounded, so it has a convergent subsequence $(f_{1,n}(x_1))_{n\in\mathbb{N}}$. Now, $(f_{1,n}(x_2))_{n\in\mathbb{N}}$ is also bounded, so it has a convergent subsequence $(f_{2,n}(x_2))_{n\in\mathbb{N}}$. Proceeding in this way, we construct several sequences $(f_{i,n})_{n\in\mathbb{N}}$ for $i \in \mathbb{N}$; the "diagonal" subsequence $(f_{n,n})_{n\in\mathbb{N}}$ is thus convergent for all x_i's. We will now show that it is uniformly convergent on X. Fix some $\varepsilon > 0$. By equicontinuity, for every $x \in X$ there is a $\delta_x > 0$ s.t. for all $y_1, y_2 \in B(x, \delta_x)$ and $n \in \mathbb{N}$, $d(y_1, y_2) < \delta$ implies $|f_{n,n}(y_1) - f_{n,n}(y_2)| < \varepsilon/3$; the family $(B(x, \delta_x))_{x\in X}$ covers X, so by compactness there is a finite subcover $\{B(x^k, \delta_{x^k})\colon k = 1, \dots, M\}$.

Bonus: Proof of Arzelà-Ascoli Theorem (cont.)

Pick some $y_k \in B(x^k, \delta_{x^k}) \cap D$ for each k, and take $\delta := \min\{\delta_{x^1}, \dots, \delta_{x^M}\}$. Fix $x \in X$ and pick the respective $y_k \in B(x, \delta)$. There exists a $N \in \mathbb{N}$ s.t. for all $n, m \geqslant N$, $|f_{n,n}(s) - f_{m,m}(s)| < \varepsilon/3$ for all $s \in \{y_1, \dots, y_M\}$ (take N as the maximum of the N's needed for $|f_{n,n}(y_k) - f_{m,m}(y_k)| < \varepsilon/3$, as k runs from 1 to M). Note that N does *not* depend on x. Then, if $n, m \geqslant N$,

$$|f_{n,n}(x) - f_{m,m}(x)| \leqslant |f_{n,n}(x) - f_{n,n}(y_k)| + |f_{n,n}(y_k) - f_{m,m}(y_k)| + |f_{m,m}(y_k) - f_{m,m}(x)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Thus, $(f_{n,n}(x))_{n\in\mathbb{N}}$ is a Cauchy sequence for every $x \in X$ (see Slide 60), so it converges to, say, $f(x)$. Taking $m \to \infty$ in the above inequality yields $|f_{n,n}(x) - f(x)| < \varepsilon$, so $(f_{n,n})_{n\in\mathbb{N}}$ is uniformly convergent to f. Finally, by equicontinuity, for every $\varepsilon > 0$ and $x \in X$, there is a $\delta > 0$ s.t. for all $y \in X$, $d(x, y) < \delta$ implies $|f_{n,n}(x) - f_{n,n}(y)| < \varepsilon$; hence taking $n \to \infty$ yields $|f(x) - f(y)| < \varepsilon$, so $f \in C(X)$.

($\Rightarrow$) Assume every sequence in $\mathscr{F} \subseteq C(X)$ has a uniformly convergent subsequence in $C(X)$. If $\mathscr{F}$ is not pointwisely bounded, then there is an $x \in X$ and a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathbb{F}$ with $M_n := |f_n(x)|$ s.t. $M_{n+1} > M_n + 1$; hence $|f_n(x) - f_m(x)| \geqslant |M_n - M_m| > 1$ for all $n, m \in \mathbb{N}$, $n \neq m$, so there is no convergent subsequence, a contradiction. Thus, $\mathscr{F}$ is pointwisely bounded.
To prove that $\mathscr{F}$ is equicontinuous, fix $\varepsilon > 0$. Since $\overline{\mathscr{F}}$, the closure of $\mathscr{F}$ in $C(X)$, is compact, $\mathscr{F}$ can be covered by finitely many balls $B(f_i, \varepsilon/3)$, $i = 1, \dots, K$. Choose $x \in X$. For each f_i, there is a $\delta_i > 0$ s.t. $d(x, y) < \delta_i$ implies $|f_i(x) - f_i(y)| < \varepsilon/3$. Let $\delta := \min\{\delta_1, \dots, \delta_K\}$. Every $f \in \mathscr{F}$ belongs to some $B(f_i, \varepsilon/3)$, so $d(f_i, f) < \varepsilon/3$. Then, if $d(x, y) < \delta$,

$$|f(x) - f(y)| \leqslant |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)| \leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

which shows that $\mathscr{F}$ is equicontinuous. □

Bonus: More on Product Topologies

Extension to Cartesian products of an arbitrary family $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in I}$ *of topological spaces:*
$\prod_{\alpha \in I} X_\alpha$ is the set of all functions x on I s.t. $x_\alpha \in X_\alpha$ for each $\alpha \in I$;
a set $W \subseteq \prod_{\alpha \in I} X_\alpha$ is open iff for every $x \in \prod_{\alpha \in I} X_\alpha$ there are nbd's of x_α, $U_\alpha \in \tau_\alpha$, for each $\alpha \in I$, s.t. $U_\alpha \neq X_\alpha$ for only *finitely* many α's, and $\prod_{\alpha \in I} U_\alpha \subseteq W$.

Note. The reason for this strange definition is that the product topology is then the weakest topology (*i.e.*, with fewest open sets) s.t. each projection function $\pi_\alpha : \Pi_{\beta \in I} X_\beta \to X_\alpha$ defined by $\pi_\alpha(x) = x_\alpha$, is continuous.

Bonus: More on Product Topologies (cont.)

Product topologies are linked to "pointwise convergence":
E.g., for $I = \mathbb{R}$ and $X_\alpha = \mathbb{R}$ for all $\alpha \in I$, $\prod_{\alpha \in I} X_\alpha$ is the set of functions $f : \mathbb{R} \to \mathbb{R}$; the open sets are generated by sets of the form $\{f : |f(x) - f_0(x)| < \varepsilon\}$ (*i.e.*, open balls in τ_α for $\alpha \in \mathbb{R}$) and finite intersections of them.

Products of Hausdorff topologies are also Hausdorff: if $f, g \in \prod_{\alpha \in I} X_\alpha$ are distinct, then $f_\beta \neq g_\beta$ for some $\beta \in I$, so f and g have respective nbd's $\{h : |f_\beta - h_\beta| < \varepsilon\}$ and $\{h : |g_\beta - h_\beta| < \varepsilon\}$, with $\varepsilon < |f_\beta - g_\beta|/2$, that do not contain the other point.

An important result concerning products of compact sets is the following:

Tychonoff's theorem
Every product of compact spaces is compact.

Proof

Let $(K_\alpha)_{\alpha\in I}$ be a family of compact spaces, and $\mathscr{F}$ a family of closed sets in $K = \prod_{\alpha\in I} K_\alpha$ having the finite intersection property, *FIP* (*i.e.*, the intersection of every finite subfamily of $\mathscr{F}$ is nonempty). We will show that $\bigcap\mathscr{F} \neq \emptyset$ to establish that K is compact.

Zorn's lemma allows us to extend $\mathscr{F}$ to a family $\mathscr{F}_0$ of (not necessarily closed) subsets of K which is maximal with respect to the FIP. The projections $\pi_\alpha: K \to K_\alpha$ (given by $\pi_\alpha(x) = x_\alpha$) of the sets of $\mathscr{F}_0$ onto K_α form a family $\mathscr{F}_0^\alpha$ of sets in K_α having the FIP, and, since K_α is compact, there is a point p_α which is in the closure of every set of $\mathscr{F}_0^\alpha$. Let p be point in K whose α-th coordinate $\pi_\alpha(p)$ is p_α for each $\alpha \in I$. We will show that p is in the closure of every set of $\mathscr{F}_0$, and therefore is in every set of $\mathscr{F}$, which will finish the proof.

Accordingly, let U be a nbd of p in K. There are, by definition of product topology, $\alpha_1, \ldots, \alpha_n$ and open sets $U_{\alpha_i} \subseteq K_{\alpha_i}$, $i = 1, \ldots, n$, s.t. $p \in \bigcap_{i=1}^n \pi_{\alpha_i}^{-1}(U_{\alpha_i}) \subseteq U$. Hence, $p_{\alpha_i} \in U_{\alpha_i}$, so U_{α_i} intersects every set of $\mathscr{F}_0^{\alpha_i}$. But then $\pi_{\alpha_i}^{-1}(U_{\alpha_i})$ intersects every set of $\mathscr{F}_0$, so it belongs to $\mathscr{F}_0$ (due to maximality). Similarly, $\bigcap_{i=1}^n \pi_{\alpha_i}^{-1}(U_{\alpha_i}) \in \mathscr{F}_0$, and so $U \in \mathscr{F}_0$. Thus U intersects every set of $\mathscr{F}_0$, and since U was an arbitrary nbd of p, p must be in the closure of every set of $\mathscr{F}_0$. □

Bonus: Topological Proof of the Infinitude of Primes

This is a curious topological proof that the number of primes is infinite, due to H. Furstenberg (1955). Try to understand it and fill in the gaps:

Theorem. There are infinitely many primes.

Proof. Define in $\mathbb{Z}$ a topology τ based on arithmetic progressions

$$S_{a,b} = \{b + an : n \in \mathbb{Z}\} = (\dots, b - 2a, b - a, b, b + a, b + 2a, \dots),$$

where a set U is open iff $U = \emptyset$ or if for every $x \in U$ there is an $a \in \mathbb{Z}$ s.t. $S_{a,x} \subseteq U$.
For every prime p, $M_p = S_{p,0}$ consists of all multiples of p. M_p is closed because its complement in $\mathbb{Z}$ is the union of the other arithmetic progressions with difference p. Now let A be the union of all the progressions M_p. If the number of primes is finite, then A is a finite union of closed sets, hence closed. However, all integers except -1 and 1 are multiples of some prime, so the complement of A is $\{-1, 1\}$, which is obviously not open, thus A cannot be closed. This contradiction shows that A cannot be a finite union of M_p's, so there are infinitely many primes. $\square$