Math 141: Lecture

Cardinality questions, the complex numbers, inequalities

Bob Hough

September 12, 2016

$\mathbb{N}^2$ is countable

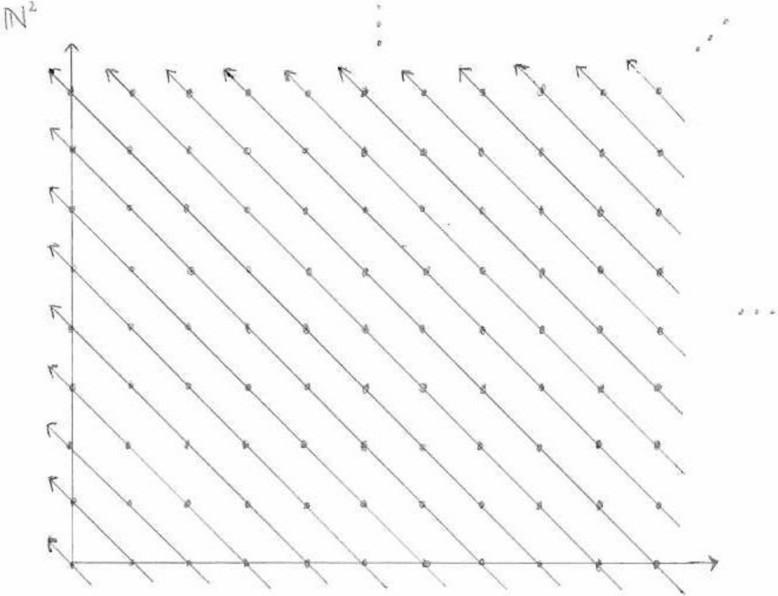
Theorem

Given $(a, b) \in \mathbb{N}^2$, define $s = a + b$. The map $f_3 : \mathbb{N}^2 \to \mathbb{N}$ defined by $f_3(a, b) = \frac{s(s+1)}{2} + b$ is a bijection.

Proof.

Observe that for each $s = 0, 1, 2, \ldots$, f_3 maps the set $\{(a, b) : a + b = s\}$ bijectively onto $\left\{n \in \mathbb{N} : \frac{s(s+1)}{2} \leq n < \frac{(s+1)(s+2)}{2}\right\}$. Since each $n \in \mathbb{N}$ lies in exactly one interval $\frac{s(s+1)}{2} \leq n < \frac{(s+1)(s+2)}{2}$, the claim follows. $\square$

$\mathbb{N}^2$ is countable



The pigeonhole principle

Define $[1] = \{1\}$, and, recursively, for $n \geq 1$, $[n+1] = [n] \cup \{n+1\}$. Thus for natural number $n \geq 1$, $[n] = \{1, 2, 3, ..., n\}$.

Theorem

Let $1 \leq m < n$ be natural numbers. There does not exist an injective function from $[n]$ to $[m]$.

Pigeonhole example

Theorem

Let $n \geq 1$ *and let* $x_1, x_2, \ldots, x_{n+1}$ *be* $n+1$ *real numbers from the half-open interval* $(0, 1]$*. There exist* $1 \leq i < j \leq n+1$ *with* $|x_i - x_j| < \frac{1}{n}$.

Proof.

Form n half-open intervals $\{I_i\}_{i=1}^{n}$, $I_i = \left(\frac{i-1}{n}, \frac{i}{n}\right]$. These intervals are disjoint and their union is $(0, 1]$. Let $f : [n+1] \to [n]$ be defined by letting $f(i)$ be the index of the interval that contains x_i. By the pigeonhole principle, f is not an injection, so there exists some $\ell \in [n]$ and some $1 \leq i < j \leq n+1$ with $f(i) = f(j) = \ell$. It follows that

$$\frac{\ell}{n} < x_i, x_j \leq \frac{\ell + 1}{n}$$

and thus $|x_i - x_j| < \frac{1}{n}$. □

Countable sets

Theorem

Let $S \subset \mathbb{N}$ be a non-empty set. Then either there is an $n \in \mathbb{N}$, $n \geq 1$ such that there is a bijection $f : [n] \to S$ or else there exists a bijection $f : \mathbb{N} \setminus \{0\} \to S$.

Countable sets

Proof sketch.

- Define a sequence of sets $S_0 = S$ and, for $n \geq 0$, if S_n is non-empty, then S_{n+1} is $S_n \setminus \{\min S_n\}$, otherwise $S_{n+1} = \emptyset$.
- Consider the set A of n for which $S_n = \emptyset$. If A is non-empty it has a least element M.
- Define a function f from $[M]$ (or $\mathbb{N} \setminus \{0\}$ if A is empty) to S, defined by $f(n)$ is the least element of S_{n-1}. Check by induction that if $m < n$ then $f(m) < f(n)$ so f is injective. Use the well-ordering principle to show f is surjective.

$\square$

Countable sets

The previous theorem gives an alternative characterization of the countable sets. A set S is countable if and only if either S is in bijection with $[n]$ for some $n \in \mathbb{N}$, or S is in bijection with $\mathbb{N}$. Thus, the elements of S can be listed in order $s_1, s_2, s_3, \ldots$ in a finite or infinite sequence, and every element of S will be reached.

The reals are uncountable

Theorem

The interval $(0, 1]$ *is uncountable.*

Proof.

We use the characterization of countability from the previous slide. The set $(0, 1]$ is not finite, as it contains $\{\frac{1}{n} : n \in \mathbb{N}, n > 1\}$, so suppose for contradiction that $(0, 1]$ has been enumerated in an infinite sequence $(0, 1] = \{x_n\}_{n \in \mathbb{N}}$. Let x_n have binary expansion $0.a_{1,n}a_{2,n}a_{3,n}...$. $\square$

The reals are uncountable

Proof.

Define another binary expansion $0.a_1a_2a_3...$ by

$$a_n = \begin{cases} 1 & n \text{ odd} \\ 1 - a_{n,n/2} & n \text{ even} \end{cases}$$

Since the sequence $a_1, a_2, ...$ contains an infinite number of 1s, it corresponds to a real number $x \in (0, 1]$. But $x \neq x_n$ for any n, since it differs at the $2n$th term of the binary expansion. $\square$

The technique of this proof is called 'diagonalization'. It plays an important role in mathematical analysis.

Schröder-Bernstein Theorem

The following is a 'gem theorem' from elementary set theory.

Theorem (Schröder-Bernstein Theorem)

Let A and B be sets, and suppose there exist injective maps $f : A \to B$ and $g : B \to A$. Then there exists a bijective map $h : A \to B$.

Proof.

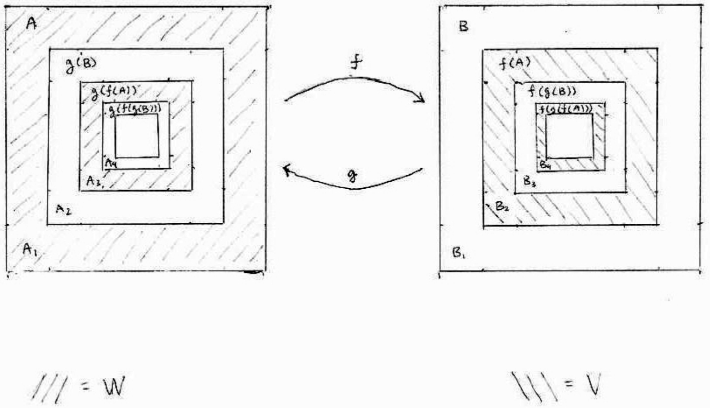
Define for $n \geq 1$, $(g \circ f)^{\circ n} : A \to A$ inductively for $n \geq 1$ by

- $(g \circ f)^{\circ 1} = g \circ f$
- For $n \geq 1$, $(g \circ f)^{\circ n+1} = (g \circ f) \circ (g \circ f)^{\circ n}$.

Similarly define $(f \circ g)^{\circ n} : B \to B$ for $n \geq 1$.

□

Schröder-Bernstein Theorem



Schröder-Bernstein Theorem

Proof.

Define nested sequence of sets $S_1 \supset S_2 \supset S_3 \supset ...$,

- $S_1 = A$, $S_2 = g(B)$
- For $n \geq 1$, $S_{2n+1} = (g \circ f)^{\circ n}(A)$, $S_{2n+2} = (g \circ f)^{\circ n} \circ g(B)$.

Similarly define $T_1 \supset T_2 \supset T_3 \supset ...$,

- $T_1 = B$, $T_2 = f(A)$
- For $n \geq 1$, $T_{2n+1} = (f \circ g)^{\circ n}(B)$, $T_{2n+2} = (f \circ g)^{\circ n} \circ f(A)$

For $n = 1, 2, 3, ...$, $f : S_n \to T_{n+1}$ and $g : T_n \to S_{n+1}$ are both bijections. □

Schröder-Bernstein Theorem

Proof.

Define for $n = 1, 2, 3, ...,$

$$A_n = S_n \setminus S_{n+1}, \qquad B_n = T_n \setminus T_{n+1}$$

The A_n are pairwise disjoint, as are the B_n. Also, $f : A_n \to B_{n+1}$ and $g : B_n \to A_{n+1}$ are bijections. □

Schröder-Bernstein Theorem

Proof.

Define $W = \bigcup_{n=1}^{\infty} A_{2n-1}$, $V = \bigcup_{n=1}^{\infty} B_{2n}$. Thus $f : W \to V$ is a bijection. We claim $g : B \setminus V \to A \setminus W$ is a bijection.

- If $x \in B \setminus V$, then $x \notin B_{2n}$ for all n, whence $g(x) \notin A_{2n-1}$ for all n, so $g(x) \in A \setminus W$, so the function is well defined.
- If $y \in A \setminus W$, then $y \notin A_{2n-1}$ for all n. In particular, $y \notin A_1$ so $y \in g(B)$ has a pre-image x. It follows that $x \in B \setminus V$, whence the map is surjective.
- The map is injective because g is injective on all of B.

$\square$

Schröder-Bernstein Theorem

Proof.

Define $h : A \to B$ by

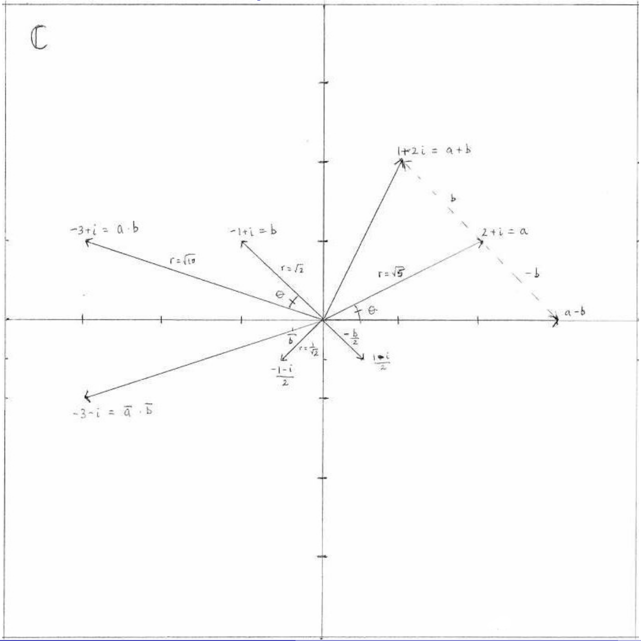
$$h(y) = \begin{cases} f(y) & y \in W \\ g^{-1}(y) & y \in A \setminus W \end{cases}$$

to complete the proof. $\square$

The field of complex numbers

- The field $\mathbb{C}$ of complex numbers consists of pairs of real numbers (a, b) written $a + bi$.
- The complexes originate from solving the equation $x^2 = -1$. The solution is the imaginary number $i = \sqrt{-1}$.
- Addition and subtraction are performed coordinatewise $(a + bi) + (a' + b'i) = a + a' + (b + b')i$.
- Multiplication is performed by treating i as a variable whose square is -1:
$$(a + bi)(c + di) = ac - bd + (ad + bc)i.$$
- The reals embed by mapping $a \mapsto (a, 0)$. All of their usual properties apply.

The field of complex numbers



The field of complex numbers

The complex numbers have a beautiful geometric interpretation.

- Addition and subtraction are performed as vector addition and vector subtraction in a two dimensional space.

The field of complex numbers

- Multiplication and division are most easily performed in polar coordinates. Given $z = a + bi$,

$$r = \sqrt{a^2 + b^2}, \qquad \theta = \sin^{-1}\left(\frac{b}{\sqrt{a^2+b^2}}\right) + 2n\pi, n \in \mathbb{Z}$$
$$a = r\cos\theta, \qquad b = r\sin\theta.$$

- Given $z_1 = r_1 \cos\theta_1 + ir_1 \sin\theta_1$, $z_2 = r_2\cos\theta_2 + ir_2\sin\theta_2$,

$$r(z_1 z_2) = r_1 r_2, \qquad \theta(z_1 z_2) = \theta_1 + \theta_2.$$

  To multiply, multiply the radii and add the angles.

- If $z_2 \neq 0$,

$$r\left(\frac{z_1}{z_2}\right) = \frac{r(z_1)}{r(z_2)}, \qquad \theta\left(\frac{z_1}{z_2}\right) = \theta_1 - \theta_2.$$

The field of complex numbers

To check the multiplication rule, write $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$, $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$,

$$\begin{aligned} z_1 z_2 &= r_1 r_2(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) \\ &\qquad + i r_1 r_2(\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2) \\ &= r_1 r_2(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)). \end{aligned}$$

Theorem (de Moirve's Theorem)

Let $\theta \in \mathbb{R}$ and $n \in \mathbb{Z}$. Then

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta).$$

Euler's formula

A beautiful formula due to Euler is as follows: for $\theta \in \mathbb{R}$,

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

This permits writing complex number $a + bi \neq 0$ as

$$a + bi = e^{\log r + i\theta}.$$

Thus all non-zero complex numbers may be expressed as the exponential of a complex number. The addition rule of the exponential function includes the law of multiplication.

We will prove Euler's formula rigorously by the end of the course.

The field of complex numbers

The *conjugate* of complex number $z = a + bi$ is $\overline{z} = a - bi$. In polar coordinates, $\theta(\overline{z}) = -\theta(z)$. One has $r^2 = z\overline{z}$, whence $\frac{1}{z} = \frac{\overline{z}}{r^2}$. This satisfies

- $\overline{z_1} + \overline{z_2} = \overline{z_1 + z_2}$
- $\overline{z_1} \cdot \overline{z_2} = \overline{z_1 z_2}$.

The fundamental theorem of algebra

Gauss proved the following theorem about the complex numbers.

Theorem (Fundamental theorem of algebra)

Let $P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ be a polynomial of degree at least 1 with complex coefficients. The equation $P(z) = 0$ has a complex solution.

A highlight of the course is a rigorous proof of this theorem.

The division algorithm

Theorem (The divison algorithm for polynomials)

Let $\mathbf{F}$ *be a field, and let* $D(x)$ *be a non-zero polynomial with coefficients in* $\mathbf{F}$. *For any polynomial* $B(x)$ *with coefficients in* $\mathbf{F}$ *there exist unique polynomials* $Q(x)$, $R(x)$ *with coefficients in* $\mathbf{F}$, *such that*

$$B(x) = Q(x)D(x) + R(x)$$

with $\deg R(x) < \deg D(x)$. *[We use the convention that* 0 *has negative degree.]*

For a proof, see HW#3.

The fundamental theorem of algebra

Theorem (Fundamental theorem of algebra, variant)

Let $P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$, $a_n \neq 0$ *be a polynomial with complex coefficients. There are complex numbers* $z_1, ..., z_n$ *such that*

$$P(z) = a_n \prod_{j=1}^{n} (z - z_j).$$

The fundamental theorem of algebra

Proof.

The proof is by induction, assuming the FTA stated previously.

- Base case: If $n = 0$ then the statement is true with an empty product.
- Inductive step: Suppose for all polynomials of degree n, a factorization of the given type is true. Let $P(z)$ by a polynomial of degree $n + 1$ with leading coefficient a_{n+1}. Let $z_{n+1} \in \mathbb{C}$ be a root of $P(z)$.
- By the division algorithm,

$$P(z) = (z - z_{n+1})Q(z) + R(z)$$

  where $R(z)$ is a constant. We have $R(z) = 0$ (choose $z = z_{n+1}$).
- The leading coefficient of $Q(z)$ is a_{n+1}, so the theorem follows from the inductive assumption.

$\square$

The distance function on $\mathbb{R}$

Define a distance function $d(\cdot, \cdot)$ on $\mathbb{R}$ by

$$d(x, y) = |x - y|.$$

This satisfies the following properties.

- Symmetry: For all $x, y \in \mathbb{R}$, $d(x, y) = d(y, x)$
- Nondegeneracy: $d(x, y) = 0$ implies $x = y$.
- Triangle inequality: For all $x, y, z \in \mathbb{R}$,

$$d(x, y) + d(y, z) \geq d(x, z).$$

A function which satisfies the listed properties above on a set is called a 'metric.' The set together with the metric is called a 'metric space.'

Proof of the triangle inequality

Theorem

Let x, y, z be real numbers. Then $|x - y| + |y - z| \geq |x - z|$.

Proof.

We may assume that $x \geq y$, since otherwise, replace x, y, z with their negatives. There are three cases to consider:

1. $z \geq x$: The LHS is $x - y + z - y = x + z - 2y$. The RHS is $z - x$. The claim reduces to $x + z - 2y \geq z - x$ or $2x \geq 2y$, which is true.
2. $y \leq z < x$: The LHS is $x - y + z - y = x + z - 2y$. The RHS is $x - z$. The inequality reduces to $x + z - 2y \geq x - z$ or $2z \geq 2y$, which is true.
3. $z < y$: The LHS is $x - y + y - z = x - z$. The RHS is also $x - z$, so equality holds.

$\square$

The Cauchy-Schwarz inequality

Theorem

Let $n \geq 1$ be an integer, and let $x_1, ..., x_n$, $y_1, ..., y_n$ be real numbers. We have

$$\left(\sum_{j=1}^{n} x_j y_j\right)^2 \leq \left(\sum_{j=1}^{n} x_j^2\right)\left(\sum_{j=1}^{n} y_j^2\right).$$

Equality holds if and only if there is a real number λ such that either $(x_1, ..., x_n) = \lambda(y_1, ..., y_n)$ or $\lambda(x_1, ..., x_n) = (y_1, ..., y_n)$.

The Cauchy-Schwarz inequality

Proof.

Write the RHS minus the LHS as (expand the square)

$$\begin{aligned}\sum_{j,k=1}^{n}(x_j^2 y_k^2 - x_j x_k y_j y_k) &= \frac{1}{2}\sum_{j,k=1}^{n}(x_j^2 y_k^2 - 2x_j x_k y_j y_k + x_k^2 y_j^2) \\ &= \frac{1}{2}\sum_{j,k=1}^{n}(x_j y_k - x_k y_j)^2 \geq 0.\end{aligned}$$

Equality holds only if each term is zero. Suppose without loss of generality that $x_1 \neq 0$. Then for each j, $x_1 y_j = x_j y_1$ implies

$$y_j = \frac{y_1}{x_1} x_j,$$

which gives the claimed condition with $\lambda = \frac{y_1}{x_1}$. $\square$

The Cauchy-Schwarz inequality

Claim: Let $x_1, x_2, ..., x_n$ be n real numbers satisfying $\frac{1}{n}\sum_{i=1}^n x_i = 1$. Then

$$\frac{1}{n}\sum_{i=1}^n x_i^2 \geq 1.$$

Proof: Apply Cauchy-Schwarz to $\underline{x} = (x_1, ..., x_n)$ and $\underline{y} = (1/n, ..., 1/n)$ to find

$$1 = \left(\frac{1}{n}\sum_{i=1}^n x_i\right)^2 \leq \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n \frac{1}{n^2}\right) = \frac{1}{n}\sum_{i=1}^n x_i^2.$$

The Euclidean distance function on $\mathbb{R}^n$

Euclidean n-space is the set of n-tuples of real numbers

$$\mathbb{R}^n = \{(x_1, ..., x_n) : x_1, ..., x_n \in \mathbb{R}\}$$

together with the distance function

$$d((x_1, ..., x_n), (y_1, ..., y_n)) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}.$$

This distance satisfies the conditions of metric, i.e. is symmetric, non-degenerate and satisfies the triangle inequality.

Proof of the triangle inequality

Theorem

Let $\underline{x} = (x_1, ..., x_n)$, $\underline{y} = (y_1, ..., y_n)$, $\underline{z} = (z_1, ..., z_n)$ *be three points of* $\mathbb{R}^n$. *We have*

$$d(\underline{x}, \underline{y}) + d(\underline{y}, \underline{z}) \geq d(\underline{x}, \underline{z}).$$

Proof.

It's equivalent to check the inequality $\sqrt{\sum_{i=1}^n x_i^2} + \sqrt{\sum_{i=1}^n y_i^2} \geq \sqrt{\sum_{i=1}^n (x_i + y_i)^2}$ since we can replace $\underline{x} - \underline{y}$ with $\underline{x}$ and $\underline{y} - \underline{z}$ with $\underline{y}$. To check the above, square both sides.

$$\mathrm{LHS}^2 = \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 + 2\sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2}.$$

$$\mathrm{RHS}^2 = \sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 + 2\sum_{i=1}^n x_i y_i.$$

Proof of the triangle inequality

Proof.

Thus, by Cauchy-Schwarz

$$\mathrm{LHS}^2 - \mathrm{RHS}^2 = 2\left(\sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2} - \sum_{i=1}^n x_i y_i\right) \geq 0.$$

Since both LHS and RHS are positive, $\mathrm{LHS} \geq \mathrm{RHS}$ follows from $\mathrm{LHS}^2 \geq \mathrm{RHS}^2$. □

Other distance functions on $\mathbb{R}^n$

Several other distance functions on $\mathbb{R}^n$ are popular, including

- The Manhattan distance: $d(\underline{x}, \underline{y}) = \sum_{i=1}^n |x_i - y_i|$.
- The chessboard distance: $d(\underline{x}, \underline{y}) = \max_i |x_i - y_i|$.
- For each $1 < p < \infty$, the ℓ^p distance: $d(\underline{x}, \underline{y}) = \left(\sum_{i=1}^n |x_i - y_i|^p\right)^{\frac{1}{p}}$.

The triangle inequalities in the first two cases reduce to the triangle inequality on $\mathbb{R}$. In the third case, the triangle inequality is known as Minkowski's inequality. We may prove this later in the course.