MTH202: Set theory and discrete mathematics

LECTURE NOTES

Why axioms?

In the early 20th century, sets were described as “well-defined collections of objects”. This leads to contradictions like the Russell’s paradox.

Surely the set of all sets that do not belong to themselves is a well-defined collection. Call it Y. So $Y = \{x : x \notin x\}$. Now either $Y \in Y$ or $Y \notin Y$. But each case implies the other (Why?). So we get a contradiction!

Clearly, something has gone wrong. We must be more precise about the notion of sets. This can be done via an axiomatic theory of sets called ZFC (shorthand for Zermelo-Fraenkel set theory with the axiom of choice). As is common in any axiomatic theory (like Euclid’s axioms for plane geometry), sets and membership are “primitive notions” and the axioms describe the precise rules to reason with them.

Axioms of ZFC

Most of the ZFC axioms describe how to construct new sets out of old.

- **Axiom of empty set**: There is a set with no members.

$$(\exists X)(\forall y)(y \notin X)$$

- **Axiom of extensionality**: Two sets are equal iff they have the same members.

$$(\forall X)(\forall Y)[(X \subseteq Y \& Y \subseteq X) \implies (X = Y)]$$

Extensionality implies that there is a unique empty set which we denote by $\emptyset$ (and later by 0).

Pairing and Union

- **Axiom of pairing**: For any two sets x, y, there is a set whose members are x, y.

$$(\forall x)(\forall y)(\exists Z)(Z = \{x, y\})$$

- **Axiom of union**: For every family $\mathcal{F}$ of set, there is a set whose members are the members of members of $\mathcal{F}$.

$$(\forall \mathcal{F})(\exists Y)[Y = \{v : (\exists X \in \mathcal{F})(v \in X)\}]$$

We write $\bigcup \mathcal{F}$ to denote the union of the sets in $\mathcal{F}$. If $X_1, X_2, \dots, X_n$ are sets, we define

$$X_1 \cup X_2 \cup \cdots \cup X_n = \bigcup \{X_1, X_2, \dots, X_n\}$$

Comprehension scheme

The **axiom of comprehension** says that for any set X and a "first-order property" $\phi(v)$, there is a subset Y of X whose members are precisely those members v of X which satisfy the property $\phi(v)$.

$$(\forall X)(\exists Y)(Y = \{v \in X : \phi(v)\})$$

So axiom of comprehension is a really an axiom scheme as we get one axiom for each "property" $\phi(v)$. We won't go into the precise definition of "first-order property" since we won't need it.

Using comprehension

During the course of these lectures, we'll sometimes introduce new sets via the expression $\{x : \phi(x)\}$. As noted before, for some properties $\phi(x)$ (like $x \notin x$) there is no such set. Therefore, on such occasions, one must check that that the axioms of ZFC guarantee the existence of such sets. For example, define the **difference of two sets** by

$$A \setminus B = \{x : x \in A \ \& \ x \notin B\}$$

This is a set since it equals $\{x \in A : x \notin B\}$ which exists by comprehension.

Intersection

If $\mathcal{F}$ is a **nonempty** collection of sets, then we define

$$\bigcap \mathcal{F} = \{y : (\forall X \in \mathcal{F})(y \in X)\}$$

To see that $\bigcap \mathcal{F}$ exists, using the fact that $\mathcal{F} \neq \emptyset$, fix an arbitrary $Z \in \mathcal{F}$ and apply comprehension to conclude that $\bigcap \mathcal{F} = \{v \in Z : (\forall X \in \mathcal{F})(x \in X)\}$ exists. Define

$$X_1 \cap X_2 \cap \cdots \cap X_n = \bigcap \{X_1, X_2, \ldots, X_n\}$$

Two sets are **disjoint** iff their intersection is the emptyset. We say that $\mathcal{F}$ is a **disjoint family** iff for every $A \neq B$ in $\mathcal{F}$, $A \cap B = \emptyset$.

Replacement scheme

Suppose X is a set and $\phi(x, y)$ is a property such that for every $x \in X$, there is a **unique** set y for which $\phi(x, y)$ holds. Then we can form the set

$$\{y : (\exists x \in X)(\phi(x, y))\}$$

We'll say more on this later when we discuss transfinite recursion.

Power set

The **power set axiom** says that for every set X, there is a set that contains all subsets of X.

$$(\forall X)(\exists Y)(Y = \{S : S \subseteq X\})$$

We denote that power set of X by $\mathcal{P}(X)$.

Natural numbers and the axiom of infinity

Definition (Natural numbers)

- $0 = \emptyset$
- $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$ $\dots$
- $n + 1 = n \cup \{n\}$

A set X is **inductive** iff $0 \in X$ and for every $x \in X$, $x \cup \{x\} \in X$. The **axiom of infinity** says that there is an inductive set. We define ω to be the intersection of all inductive sets.

Definition (The set of natural numbers)

$$\omega = \{0, 1, 2, \dots, n, n + 1, \dots\}$$

Other axioms

The remaining two axioms are

- **Axiom of choice**
- **Axiom of foundation**

We'll introduce the **axiom of choice** later. For the purposes of this course, we can safely ignore the **axiom of foundation**.

The axioms of ZFC

- **Axiom of empty set**
- **Axiom of extensionality**
- **Axiom of pairing**
- **Axiom of union**
- **Axiom scheme of comprehension**
- **Axiom scheme of replacement**
- **Axiom of power set**
- **Axiom of infinity**
- **Axiom of choice**
- **Axiom of foundation**

Ordered pairs

Definition

The ordered pair with first coordinate x and second coordinate y *is defined by*

$$(x, y) = \{\{x\}, \{x, y\}\}$$

Note that (x, y) exists by the axiom of pairing. The key property of ordered pairs is the following.

Proposition

If $(x, y) = (a, b)$*, then* $x = a$ *and* $y = b$*.*

The proof is left as an exercise.

Cartesian products

Definition

The cartesian product $X \times Y$ *is defined to be the set of all ordered pairs whose first the coordinate is in* X *and second coordinate is in* Y*.*

$$X \times Y = \{(x, y) : x \in X \;\&\; y \in Y\}$$

Note that $X \times Y$ is a subset of $\mathcal{P}(\mathcal{P}(X \cup Y))$ which exists by the pairing, union and power set axioms. So the existence of

$$X \times Y = \{v \in \mathcal{P}(\mathcal{P}(X \cup Y)) : (\exists x \in X)(\exists y \in Y)(v = (x, y))\}$$

follows from comprehension.

Relations

A **relation** R is a set of ordered pairs. If R is a relation, then

- $\operatorname{dom}(R) = \{x : (\exists y)((x, y) \in R)\}$
- $\operatorname{range}(R) = \{y : (\exists x)((x, y) \in R)\}$

We say that R is a **relation from** A **to** B iff $R \subseteq A \times B$. Note that every relation R is a relation from $\operatorname{dom}(R)$ to $\operatorname{range}(R)$. We say that R is a **relation on** A iff R is a relation from A to A.
Notation: If R is a relation, we sometimes write xRy (read x is R-related to y) instead of $(x, y) \in R$.

Functions

F is a **function** iff F is a relation and for every $x \in \mathrm{dom}(F)$, there is a unique $y \in \mathrm{range}(F)$ such that $(x, y) \in F$. We write $F(x) = y$ instead of $(x, y) \in F$. We say that F **is a function from** A **to** B, and write $F : A \to B$, iff F is a function, $\mathrm{dom}(F) = A$ and $\mathrm{range}(F) \subseteq B$.

Suppose $F : A \to B$. We say that

- F is **injective** (one-one) if for every $x \neq y$ in A, $F(x) \neq F(y)$.
- F is **surjective** (onto) if $\mathrm{range}(F) = B$
- F is **bijective** iff it is both injective and surjective.

If $f : A \to B$ and $g : B \to C$, then $g \circ f : A \to C$ defined by $(g \circ f)(x) = g(f(x))$ is the **composition** of f with g.

If $f : A \to B$ is a bijection, the **inverse** of f is the function $f^{-1} : B \to A$ defined by

$$f^{-1}(y) = x \iff f(x) = y$$

Images/preimages, Finite/Infinite, countable/uncountable

Suppose $F : A \to B$, $X \subseteq A$ and $Y \subseteq B$.

- The **image of** X **under** F is $F[X] = \{F(x) : x \in X\}$.
- The **preimage of** Y **with respect to** F is

$$F^{-1}[Y] = \{x \in A : F(x) \in Y\}$$

A set X is **finite** iff for some natural number n, there exists a bijection $f : n \to X$. Otherwise, it is **infinite**.

Isomorphism

Definition (Isomorphism)

Suppose R, S are relations and A, B are sets. We sat that (A, R) is isomorphic to (B, S) *and write $(A, R) \cong (B, S)$ iff there is a bijection $f : A \to B$ such that for every $x, y \in A$, xRy iff $f(x)Sf(y)$.*

Equivalence relations and partitions

We say that R is an **equivalence relation** on A iff R is a relation from A to A which satisfies the following.

- **Reflexive** For every $a \in A$, aRa.
- **Symmetric** If aRb, then bRa.
- **Transitive** If aRb and bRc, then aRc.

We say that $\mathcal{F}$ is a **partition** of A iff $\mathcal{F}$ is a disjoint family and $\bigcup \mathcal{F} = A$.

Exercise

Suppose R is an equivalence relation on A. For each $a \in A$, define the R-equivalence class of a by $[a] = \{b \in A : aRb\}$. Then $\{[a] : a \in A\}$ is a partition of A.

Linear orderings

Definition

A linear ordering *is a pair $(A, \prec)$ such that A is a nonempty set and $\prec$ is a binary relation on A that satisfies*

- Irreflexive *For every $a \in A$, $\neg(a \prec a)$ ($\neg$ denotes negation).*
- Transitive *If $a \prec b$ and $b \prec c$, then $a \prec c$.*
- Total *For every $x, y \in X$ if $x \neq y$, then either $x \prec y$ or $y \prec x$.*

If $(A, \prec)$ is a linear ordering, we define the relation $\preceq$ on A by

$$a \preceq b \iff (a \prec b \text{ or } a = b)$$

If $(A, \prec)$ is a linear ordering and $x \in A$, we define the set if **predecessors** of x in $(A, \prec)$ by $\mathrm{pred}(A, \prec, x) = \{y \in A : y \prec x\}$.

Well-orderings

Suppose $(X, \prec)$ is a linear ordering, $A \subseteq X$ and $y \in A$. We say that y is the $\prec$-least member of A iff for every $z \in A$, $y \preceq z$.

Definition

A well-ordering *is a pair* $(X, \prec)$ *such that* $\prec$ *is a linear ordering on* X *such that for every nonempty* $A \subseteq X$*,* A *has a* $\prec$*-least member.*

Note that if $(X, \prec)$ is a well ordering then for every $x \in X$, either x is $\prec$-largest member of X or x has a $\prec$-successor y which means that $x \prec y$ and for every $z \prec y$, $z \preceq x$. So the first few members of X look like: $x_0 \prec x_1 \prec x_2 \prec \ldots$

Well-orderings

Lemma

Suppose $(X, \prec)$ *is a well-ordering. Then* $(X, \prec)$ *is not isomorphic to* $(\mathit{pred}(X, \prec, x), \prec)$ *for any* $x \in X$*.*

Proof: Suppose not and let $f : X \to \text{pred}(X, \prec, x)$ be an isomorphism. Note that $f(x) \prec x$ so the set

$$W = \{y \in X : f(y) \prec y\}$$

is nonempty. Let z be $\prec$-least member of W. So $f(z) \prec z$. Since f preserves $\prec$, we also get $f(f(z)) \prec f(z)$. Put $w = f(z)$ and note that $w \in W$. Since z is the $\prec$-least member of W, $z \preceq w = f(z)$ which is a contradiction as $f(z) \prec z$. □

Well-orderings

Lemma

Suppose $(X, \prec)$ is a well-ordering and $f : X \to X$ is an isomorphism. Then f is the identity function on X.

Proof Let $f : X \to X$ be an isomorphism and, towards a contradiction, suppose for some $v \in X$, $f(v) \neq v$. So the set $W = \{v \in X : f(v) \neq v\}$ is nonempty. Let x be the $\prec$-least member of W. Put $y = f(x)$. Then either $y \prec x$ or $x \prec y$. If $y \prec x$, then $f(y) = y$ as x was $\prec$-least non fixed point of f. But since f preserves $\prec$ and $y \prec x$, $y = f(y) \prec f(x) = y$ which is impossible. Next suppose $x \prec y$. Since f is surjective, there is some $w \in X$ such that $f(w) = x$. Clearly, $w \not\preceq x$ so $x \prec w$. But then $y = f(x) \prec f(w) = x$ which contradicts $x \prec y$. □

Well-orderings

Theorem

Suppose $(X, \prec_1)$ and $(Y, \prec_2)$ are well-orderings. Then exactly one of the following holds.

(1) $(X, \prec_1) \cong (Y, \prec_2)$.

(2) *For some* $x \in X$, $(\mathit{pred}(X, \prec_1, x), \prec_1) \cong (Y, \prec_2)$.

(3) *For some* $y \in Y$, $(\mathit{pred}(Y, \prec_2, y), \prec_2) \cong (X, \prec_1)$.

Furthermore, in each of the three cases, the isomorphism is unique.

Proof: See Homework.

Well-ordering theorem

The **axiom of choice** says the following. For every family $\mathcal{E}$ of nonempty sets, there is a function F such that $\operatorname{dom}(F) = \mathcal{E}$ and for every $A \in \mathcal{E}$, $F(A) \in A$. We say that F is a choice function on $\mathcal{E}$.

Theorem (Zermelo, 1904)

Every set can be well-ordered.

Proof: Watch video.

Ordinals

Definition (Transitive sets)

A set x is transitive *iff for every $y \in x$, $y \subseteq x$.*

Definition (Ordinals)

x is an ordinal *iff x is transitive and $(x, \in)$ is a well-ordering.*

We are slightly abusing the notation here since $\in$ is not a set. Nevertheless, for any set x, the relation $\varepsilon_x = \{(y, z) : y, z \in x \ \& \ y \in z\}$ is the restriction of the membership relation on x. So $\in$ stands for ε_x in the pair $(x, \in)$.

Examples

- $0 = \emptyset$ is an ordinal.
- $1, 2, 3, \ldots, n, n+1, \ldots$ are ordinals.
- The set of natural numbers ω is an ordinal.
- $\omega \cup \{\omega\}$ is an ordinal.
- The set of even numbers $E = \{0, 2, 4, 6, \ldots, 2n, \ldots\}$ is well-ordered by $\in$ but E is not an ordinal since it is not a transitive set.

Ordinals

Claim

If x is an ordinal and $y \in x$, then y is an ordinal and $y = pred(x, \in, y)$.

Proof: (a) y is transitive: Suppose $z \in y$. We must check that $z \subseteq y$. Fix $w \in z$. So $w \in z \in y \in x$. As x is transitive, each one of z, w, y is in x. Since x is well-ordered by $\in$, in particular, $\in$ is a transitive relation on x. As $w \in z \in y$, we get $w \in y$. Hence $z \subseteq y$.

(b) y is well-ordered by $\in$: Note that since x is transitive, $y \subseteq x$. Now if $(A, \prec)$ is a well-ordering and $B \subseteq A$, then the restriction of $\prec$ to B is also a well-order. So y is well-ordered by $\in$.

(c) $y = \text{pred}(x, \in, y)$: If $z \in y$, then $z \in x$. So $z \in \text{pred}(x, \in, y)$. Hence $y \subseteq \text{pred}(x, \in, y)$. If $z \in \text{pred}(x, \in, y)$, then $z \in y$. So $\text{pred}(x, \in, y) \subseteq y$. □

Ordinals

The proofs of the following facts are left to the reader.

Theorem

(a) *If x is an ordinal and $y \in x$, then y is an ordinal and $y = pred(x, \in, y)$.*

(b) *If x, y are ordinals and $(x, \in) \cong (y, \in)$, then $x = y$.*

(c) *If x is an ordinal, then $x \notin x$.*

(d) *If x, y are ordinals, then exactly one of the following holds: $x = y$, $x \in y$, $y \in x$.*

(e) *If C is a non empty set of ordinals, then there exists $x \in C$ such that $(\forall y \in C)(y = x \text{ or } x \in y)$.*

(f) *If A is a set of ordinals, then $(A, \in)$ is a well-ordering. Hence if A is a transitive set of ordinals, then A is an ordinal.*

Ordinals and well-orderings

Theorem

For every well-ordering $(X, \prec)$, there is a unique ordinal A such that $(X, \prec) \cong (A, \in)$.

Proof: Uniqueness follows from clause (b) above. Let Y be the set of all $x \in X$ such that $(\mathrm{pred}(X, \prec, x), \prec)$ is isomorphic to an ordinal. Using the axiom of replacement, define a function f on Y by letting $f(x)$ to be the unique ordinal which is isomorphic to $(\mathrm{pred}(X, \prec, x), \prec)$. Let $A = \mathrm{range}(f)$. Note that A is a transitive set of ordinals. Hence A is an ordinal. It is also easy to check that $f : Y \to A$ is an isomorphism from $(Y, \prec)$ to $(A, \in)$.
So we would be done if $Y = X$. Suppose $Y \neq X$. Note that Y is a $\prec$-initial segment of X. Let b be the $\prec$-least member of $X \setminus Y$. Then $Y = \mathrm{pred}(X, \prec, b)$. But $(\mathrm{pred}(X, \prec, b), \prec)$ is isomorphic to the ordinal A. So $b \in Y$ which is a contradiction. □

Order types, sup/min

Definition (Order type)

If $(X, \prec)$ *is a well ordering, let* $\mathrm{type}(X, \prec)$ *be the unique ordinal* A *such that* $(X, \prec) \cong (A, \in)$*.*

We denote ordinals by Greek letters: α, β, γ, etc. and from now on we'll write $\alpha < \beta$ instead of $\alpha \in \beta$.

Definition (sup, min)

For a set of ordinals A*, define* $\sup(A) = \bigcup A$ *and, if* $A \neq 0$*,* $\min(A) = \bigcap A$*.*

Check that $\sup(A)$ is the least ordinal which is greater than or equal to every ordinal in A and $\min(A)$ is the least ordinal in A.

Definition (Successor and limit)

The successor of α *is defined by* $S(\alpha) = \alpha \cup \{\alpha\}$*.*
An ordinal α *is called a* successor ordinal *if for some ordinal* β*,* $\alpha = S(\beta)$*. Otherwise* α *is a* limit ordinal*.*

Note that $S(\alpha)$ is the least ordinal bigger than α.
The first few ordinals are:

$$0 < 1 < 2 < \cdots < n < n+1 < \cdots < \omega < S(\omega) < S(S(\omega)) < \ldots$$

Note that ω is a limit ordinal.

Sum of linear orders

Given two linear orderings $(L_1, \prec_1)$ and $(L_2, \prec_2)$, one can define another linear ordering by putting a copy of $(L_2, \prec_2)$ after a copy of $(L_1, \prec_1)$. The following definition makes this precise.

Definition

Suppose $(L_1, \prec_1)$ *and* $(L_2, \prec_2)$ *are linear orderings. We define the sum* $(L, \prec) = (L_1, \prec_1) \oplus (L_2, \prec_2)$ *as follows.*

(1) $L = (L_1 \times \{0\}) \bigcup (L_2 \times \{1\})$.

(2) *For every* $x, y \in L$, $x \prec y$ *iff one of the following holds*
(i) $x = (a, 0)$, $y = (b, 0)$ *and* $a \prec_1 b$.
(ii) $x = (a, 1)$, $y = (b, 1)$ *and* $a \prec_2 b$.
(iii) $x = (a, 0)$ *and* $y = (b, 1)$.

Note that we defined $L = (L_1 \times \{0\}) \bigcup (L_2 \times \{1\})$ (and not $L = L_1 \bigcup L_2$) because L_1, L_2 may not be disjoint.

Sum of ordinals

Definition (Ordinal addition)

$$\alpha + \beta = \mathit{type}((\alpha, <) \oplus (\beta, <))$$

It is easy to check that $\alpha + \beta$ is an ordinal. Note that $S(\alpha) = \alpha + 1$ and if $m, n < \omega$, then $m + n$ is the usual sum. Ordinal addition is not commutative in general: For example $\omega = 1 + \omega \neq \omega + 1$. The first few ordinals are:

$$0 < 1 < \cdots < \omega < S(\omega) = \omega + 1 < \omega + 2 < \cdots < \omega + \omega < \ldots$$

Exercise Show that if $\alpha < \beta$, there is a unique ordinal γ such that $\alpha + \gamma = \beta$. (Hint: $\gamma = \text{type}(\beta \setminus \alpha, \in)$).

Lexicographic product of linear orders

Definition (Product of linear orders)

Suppose $(L_1, \prec_1)$ *and* $(L_2, \prec_2)$ *are linear orderings. We define the product* $(L, \prec) = (L_1, \prec_1) \otimes (L_2, \prec_2)$ *as follows.*

(1) $L = L_1 \times L_2$.

(2) *For every* (x_1, y_1) *and* (x_2, y_2) *in* L, $(x_1, y_1) \prec (x_2, y_2)$ *iff*
(a) Either $x_1 \prec_1 x_2$ *or*
(b) $x_1 = x_2$ *and* $y_1 \prec_2 y_2$.

Product of ordinals

Definition (Ordinal multiplication)

$$\alpha \cdot \beta = \mathit{type}((\beta, <) \otimes (\alpha, <))$$

It is easy to check that $\alpha \cdot \beta$ is an ordinal. If $m, n < \omega$, then $m \cdot n$ is the usual product. Ordinal multiplication is not commutative in general: $\omega \cdot 2 = \omega + \omega \neq 2 \cdot \omega = \omega$.

Laws of ordinal arithmetic

Fact

For any α, β and γ the following hold.

(i) *(Associativity)* $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ *and* $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$

(ii) $\alpha + 0 = \alpha$, $\alpha \cdot 0 = 0$ *and* $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$.

(iii) *(Continuity at limits) If* β *is a limit ordinal,* $\alpha + \beta = \sup\{\alpha + \eta : \eta < \beta\}$ *and* $\alpha \cdot \beta = \sup\{\alpha \cdot \eta : \eta < \beta\}$

(iv) *(Left distributivity)* $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$

Restrictions of functions

Suppose f is a function and $X \subseteq \mathrm{dom}(f)$. We define the **restriction of** f **to** X, denoted $f \upharpoonright X$, as follows.

- $\mathrm{dom}(f \upharpoonright X) = X$.
- For each $a \in X$, $(f \upharpoonright X)(a) = f(a)$.

Burali-Forti paradox

Theorem

No set contains all ordinals.

Proof: Suppose there is a set X such that every ordinal is a member of X. Using comprehension, define $\Gamma = \{y \in X : y \text{ is an ordinal}\}$. Then Γ is a transitive set of ordinals and hence Γ is also an ordinal. Since all ordinals are members of X, this means that $\Gamma \in \Gamma$ which is impossible. □

Sequences indexed by ordinals

A **sequence** is a function whose domain is an ordinal. If f is a sequence and $\mathrm{dom}(f) = \gamma$, we sometimes write $\langle f(\alpha) : \alpha < \gamma \rangle$ instead of f. If f is a sequence with $\mathrm{dom}(f) = \gamma$, we also say that f is a sequence of **length** γ. We say that f is a sequence in X if $\mathrm{range}(f) \subseteq X$. A set X is **countable** iff there is a sequence $\langle x_n : n < \omega \rangle$ whose range is X. If there is no such sequence, X is **uncountable**. If X is a set and α is an ordinal, define X^α to be the set of all functions from α to X. If $n < \omega$, members of X^n are called n-**tuples** in X.

Lemma

Let X be any set. Then there are an ordinal γ and an injective sequence $\langle x_\alpha : \alpha < \gamma \rangle$ whose range is X.

Proof: Let $\prec$ be a well-order on X. Put $\gamma = \mathrm{type}(X, \prec)$ and fix an order isomorphism f from $(\gamma, <)$ to $(X, \prec)$. For each $\alpha < \gamma$, define $x_\alpha = f(\alpha)$. Then $\langle x_\alpha : \alpha < \gamma \rangle$ is an injective sequence whose range is X. □

Formalizing mathematics within ZFC

We have already constructed $(\omega, +, .)$ where $+$ and $.$ denote addition and multiplication of finite ordinals (natural numbers). One can go on and construct $(\mathbb{Z}, +, .)$ (the ring of integers), $(\mathbb{Q}, +, .)$ (the field of rational numbers), $(\mathbb{R}, +, .)$ (the field of real numbers) and $(\mathbb{C}, +, .)$ (the field of complex numbers), Euclidean spaces $\mathbb{R}^n$ etc. in the usual way. Once this as been done, it is not difficult to convince oneself, that all the theorems in various fields of mathematics can be expressed and proved within ZFC. We won't pursue this path here.

Ordinary induction

The principle of mathematical induction says the following. Suppose $P(n)$ is a property of natural numbers. Assume

- $P(0)$ holds and
- for every $n < \omega$,

$$[(\forall k < n)P(k)] \implies P(n)$$

Then $P(n)$ holds for every $n < \omega$.

Transfinite induction

Theorem

Suppose $P(\alpha)$ is a property of ordinals. Assume

(1) *$P(0)$ holds and*

(2) *for every ordinal $\alpha > 0$,*

$$[(\forall\beta < \alpha)P(\beta)] \implies P(\alpha)$$

Then $P(\alpha)$ holds for every ordinal α.

Proof: Suppose not and fix the least ordinal α such that $P(\alpha)$ fails. By clause (1), $\alpha > 0$. Note that $P(\beta)$ holds for every $\beta < \alpha$. Hence clause (2) implies that $P(\alpha)$ holds. A contradiction. □

Ordinary recursion

Ordinary recursion constructs objects through finite stages. An example follows.

Theorem (Cantor, 1874)

The set of real numbers is uncountable.

Proof: It suffices to show that for every sequence $\langle a_n : n < \omega\rangle$ of real numbers, there is a real number which does not appear in this sequence. Recursively construct a sequence of closed intervals $[x_n, y_n]$ such that

1. $[x_0, y_0] = [0, 1]$,
2. $x_n < x_{n+1} < y_{n+1} < y_n$ and
3. $a_n \notin [x_{n+1}, y_{n+1}]$.

Choose $x \in \bigcap\{[x_n, y_n] : n < \omega\}$. Then x is not in the sequence. □

Ordinary recursion

Theorem

Let F *be "function" defined on the class of all sets. Suppose x is a set. Then there is a unique function h such that*

1. *$dom(h) = \omega$,*
2. *$h(0) = x$ and*
3. *for each $n \geq 1$, $h(n) = \mathsf{F}(h \restriction n)$.*

In the previous proof, F can be chosen as follows. If for some $n < \omega$, x is a sequence of length n whose last entry is a closed interval $[x, y]$, then $\mathsf{F}(x)$ is a closed subinterval of $[x, y]$ which does not contain a_n. Otherwise, define $\mathsf{F}(x) = 0$.

Transfinite recursion

Theorem

Let F *be a "function" defined on the class of all sets. Then for each ordinal* γ*, there is a unique function* h *such that*

1. $dom(h) = \gamma$
2. *For each* $\alpha < \gamma$, $h(\alpha) = \mathsf{F}(h \upharpoonright \alpha)$.

In applications of this theorem, we imagine the function h as being defined in γ stages. At stage 0, by clause 2, we must define $h(0) = \mathsf{F}(h \upharpoonright 0) = \mathsf{F}(0)$. Having defined $h(\beta)$ for every $\beta < \alpha$, we feed $h \upharpoonright \beta = \langle h(\beta) : \beta < \alpha \rangle$ to F to get $h(\alpha)$.

Well-ordering theorem revisited

Let us use transfinite recursion to give another proof of the well-ordering theorem. Let X be a set. Using the axiom of choice, fix a choice function $f : \mathcal{P}(X) \setminus \{0\} \to X$. Fix a set $s_\star \notin X$. By transfinite recursion, for each ordinal γ, define a function $h_\gamma : \gamma \to X \cup \{s_\star\}$ as follows. For every ordinal $\alpha < \gamma$,

$$h_\gamma(\alpha) = \begin{cases} f(X \setminus \mathrm{range}(h_\gamma \restriction \alpha)) & \textbf{if } \mathrm{range}(h_\gamma \restriction \alpha) \neq X \\ s_\star & \textbf{otherwise} \end{cases} \tag{1}$$

We claim that there must be some ordinal γ such that $s_\star \in \mathrm{range}(h_\gamma)$. Otherwise, applying replacement axiom to the formula $\phi(x, y)$ which says "y is the least ordinal such that $x \in \mathrm{range}(h_y)$", we'll get a set that contains all ordinals which is impossible. Let γ be least such that $s_\star \in \mathrm{range}(h_\gamma)$. Then h_γ is a bijection from γ to $X \cup \{s_\star\}$. Hence X can be well-ordered. □

Ordinal exponentiation

As another application of transfinite recursion, let us define ordinal exponentiation α^β. By transfinite recursion on β, define α^β as follows.

(i) $\alpha^0 = 1$.

(ii) $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$.

(iii) If β is a limit ordinal, then $\alpha^\beta = \sup(\{\alpha^\gamma : \gamma < \beta\})$.

Note that we have divided the construction into three cases: $\beta = 0$, β is a successor ordinal, β is a limit ordinal.

Ordinal exponentiation

Let us use transfinite **induction**, to prove the following

$$\alpha^{\beta+\gamma} = \alpha^{\beta} \cdot \alpha^{\gamma}$$

We prove this by induction on γ.

- **Case** $\gamma = 0$: $\alpha^{\beta+\gamma} = \alpha^{\beta+0} = \alpha^{\beta} = \alpha^{\beta} \cdot 1 = \alpha^{\beta} \cdot \alpha^{0} = \alpha^{\beta} \cdot \alpha^{\gamma}$
- **Case** $\gamma = \delta + 1$: $\alpha^{\beta+\gamma} = \alpha^{\beta+(\delta+1)} = \alpha^{(\beta+\delta)+1} = \alpha^{\beta+\delta} \cdot \alpha =$ $= (\alpha^{\beta} \cdot \alpha^{\delta}) \cdot \alpha = \alpha^{\beta} \cdot (\alpha^{\delta} \cdot \alpha) = \alpha^{\beta} \cdot \alpha^{\delta+1} = \alpha^{\beta} \cdot \alpha^{\gamma}$
- **Case** γ is limit: $\alpha^{\beta+\gamma} = \sup(\{\alpha^{\beta+\delta} : \delta < \gamma\}) =$ $= \sup(\{\alpha^{\beta} \cdot \alpha^{\delta} : \delta < \gamma\}) = \alpha^{\beta} \cdot \sup(\{\alpha^{\delta} : \delta < \gamma\}) = \alpha^{\beta} \cdot \alpha^{\gamma}$.

Proof of transfinite recursion

Theorem

Let F *be a "function" defined on the class of all sets. Then for each ordinal γ, there is a unique function h such that*

1. *dom*$(h) = \gamma$
2. *For each* $\alpha < \gamma$, $h(\alpha) = \mathsf{F}(h \restriction \alpha)$.

Proof: Note that the following proof will not use the axiom of choice. Let us first check uniqueness. Suppose h, h' are two distinct functions satisfying clauses 1 and 2. Let $\alpha < \gamma$ be least such that $h(\alpha) \neq h'(\alpha)$. Then

$$h(\alpha) = \mathbf{F}(h \restriction \alpha) = \mathbf{F}(h' \restriction \alpha) = h'(\alpha)$$

which is a contradiction.

Proof of transfinite recursion

Next, we prove the existence of h by transfinite induction on γ. Suppose for each $\eta < \gamma$, there exists h_η such that $\mathrm{dom}(h_\eta) = \eta$ and for every $\alpha < \eta$, $h_\eta(\alpha) = \mathsf{F}(h_\eta \upharpoonright \alpha)$. We will construct h such that $\mathrm{dom}(h) = \gamma$ and for every $\alpha < \gamma$, $h(\alpha) = \mathsf{F}(h \upharpoonright \alpha)$.

Claim

For every $\eta < \theta < \gamma$, $h_\eta = h_\theta \upharpoonright \eta$.

Proof of Claim: Just note that both h_η and $h_\theta \upharpoonright \eta$ satisfy clauses 1 + 2 for $\gamma = \eta$. Hence by uniqueness, $h_\eta = h_\theta \upharpoonright \eta$. □

Define, $h = \bigcup\{h_\eta : \eta < \gamma\}$ and note that the claim implies that h is a function with domain γ. It is clear that h is as required. □

Partial orderings

A **partial ordering** is a pair $(P, \preceq)$ where $\preceq$ is a binary relation on P that satisfies the following.

- **Reflexive** For every $p \in P$, $p \preceq p$
- **Antisymmetric** For every $p, q \in P$, if $p \preceq q$ and $q \preceq p$, then $p = q$.
- **Transitive** For every $p, q, r \in P$, if $p \preceq q$ and $q \preceq r$, then $p \preceq r$.

Note that we do not require that any two members of P be $\preceq$-comparable. If $(P, \preceq)$ is a partial ordering and $p, q \in \mathbb{P}$, we write $p \prec q$ iff $p \preceq q$ and $p \neq q$.

Examples

Examples

(1) If $(L, \prec)$ is a linear ordering, then $(L, \preceq)$ is a partial ordering.

(2) For any family of sets $\mathcal{F}$, $(\mathcal{F}, \subseteq)$ is a partial ordering.

The second example is universal in the following sense.

Proposition

Every partial ordering $(P, \preceq)$ is isomorphic to $(\mathcal{F}, \subseteq)$ for some $\mathcal{F}$.

Proof: For each $p \in P$, let $W_p = \{q \in P : q \leq p\}$. Define $\mathcal{F} = \{W_p : p \in P\}$. Then it is easy to check that $(P, \preceq) \cong (\mathcal{F}, \subseteq)$ via the function $p \mapsto W_p$. □

Upper/lower bounds, Maximal/minimal

Suppose $(P, \preceq)$ is a partial ordering, $p \in P$ and $X \subseteq P$.

- We say that p is an **upper bound** of X iff for every $q \in X$, $q \leq p$.
- We say that p is a **lower bound** of X iff for every $q \in X$, $p \leq q$.
- We say that p is a **maximal element** of P iff there is no $q \in P$ such that $p \prec q$.
- We say that p is a **minimal element** of P iff there is no $q \in P$ such that $q \prec p$.

Chains in partially ordered sets

Chains are **linearly** ordered subsets of partial orderings.

- Suppose $(P, \preceq)$ is a partial ordering and $C \subseteq P$. We say that C is a **chain** in $(P, \preceq)$ iff for every $p, q \in C$, either $p \preceq q$ or $q \preceq p$.
- If $\mathcal{F}$ is a family of sets, by a chain in $\mathcal{F}$, we mean a chain in $(\mathcal{F}, \subseteq)$.

Exercise: Show that there is an uncountable chain in $\mathcal{P}(\omega)$.

Zorn's lemma

Theorem

Let $(P, \preceq)$ be a partial ordering in which every chain has an upper bound. Then P has a maximal element.

Proof: Towards a contradiction, suppose P has no maximal element. Fix an ordinal γ and an injective sequence $\langle p_\alpha : \alpha < \gamma \rangle$ whose range is P. By transfinite recursion on $\alpha < \gamma$, construct a sequence $\langle C_\alpha : \alpha < \gamma \rangle$ such that the following hold.

- Each C_α is a chain in P and $C_0 = 0$.
- For every $\alpha < \beta < \gamma$, $C_\alpha \subseteq C_\beta$.
- If α is limit, $C_\alpha = \bigcup\{C_\beta : \beta < \alpha\}$.
- For every $\alpha < \gamma$, $C_{\alpha+1}$ is defined as follows. **If** p_α is an upper bound of C_α, **then** $C_{\alpha+1} = C_\alpha \cup \{p_\eta\}$ where η is least such that $p_\alpha \prec p_\eta$. **Otherwise**, $C_{\alpha+1} = C_\alpha$.

Put $C = \bigcup\{C_\alpha : \alpha < \gamma\}$. Then it is easy to check that C is a chain in P and C has no upper bound in P. A contradiction. □

Equivalents of AC

Let ZF be the theory ZFC without the axiom of choice. In ZF, the following are equivalent.

(1) Axiom of choice

(2) Well-ordering theorem

(3) Zorn's lemma

Proof: We already proved (1) $\implies$ (2) and (2) $\implies$ (3). So it suffices to prove (3) $\implies$ (1).
Let X be a set and $\mathcal{F} = \mathcal{P}(X) \setminus \{0\}$. Define h to be a partial choice function on $\mathcal{F}$ iff h is a function, $\mathrm{dom}(h) \subseteq \mathcal{F}$ and for every $A \in \mathrm{dom}(h)$, $h(A) \in A$. Let $\mathcal{G}$ be the family of all partial choice functions on $\mathcal{F}$. Note that every chain in $(\mathcal{G}, \subseteq)$ has an upper bound, namely its union. Using Zorn's lemma, fix a maximal element h in $\mathcal{G}$. Note that $\mathrm{dom}(h) = \mathcal{F}$, otherwise fix some $A \in \mathcal{F} \setminus \mathrm{dom}(h)$, $a \in A$ and consider $h' = h \cup \{(A, a)\}$. Clearly $h' \in \mathcal{G}$ is larger than h which contradicts the maximality of h. So $\mathrm{dom}(h) = \mathcal{F}$ and hence it is a choice function on $\mathcal{F}$. □

Applications of Zorn's Lemma: Example I

Theorem

For any two sets A and B, either there is an injection from A to B or there is an injection from B to A.

Proof: Let $\mathcal{F}$ be the family of all functions f such that $\mathrm{dom}(f) \subseteq A$, $\mathrm{range}(f) \subseteq B$ and f is injective. Then $(\mathcal{F}, \subseteq)$ is a partial ordering.

Exercise: Check that every chain in $\mathcal{F}$ has an upper bound.

By Zorn's lemma, $\mathcal{F}$ has a maximal member h. We claim that either $\mathrm{dom}(h) = A$ or $\mathrm{range}(h) = B$. This suffices since in the former case, h is an injection from A to B and in the latter case, h^{-1} is an injection from B to A. Towards a contradiction, suppose $\mathrm{dom}(h) \neq A$ and $\mathrm{range}(h) \neq B$. Fix $x \in A \setminus \mathrm{dom}(A)$ and $y \in B \setminus \mathrm{range}(h)$. Define $h' = h \cup \{(x, y)\}$. Then $h' \in \mathcal{F}$. Hence h is not maximal in $\mathcal{F}$ which is a contradiction. □

Example II

Lemma

Every partial ordering $(P, \preceq)$ contains a $\subseteq$-maximal chain C. In other words, C is a chain in P and for every chain D in P, if $C \subseteq D$, then $C = D$.

Proof: Consider the partial ordering $(\mathcal{F}, \subseteq)$ where $\mathcal{F}$ be the family of all chains in P. If $\mathcal{E}$ is a chain in $\mathcal{F}$, then $\bigcup \mathcal{E}$ is a chain in P [Why?]. Hence every chain in $(\mathcal{F}, \subseteq)$ has an upper bound. Let C be a maximal element of $(\mathcal{F}, \subseteq)$. Then C is a $\subseteq$-maximal chain in P. □

Example III

Notation: $\mathbb{Q}$ is the set of rational numbers, $\mathbb{R}$ is the set of real numbers, $\mathbb{R}^+$ is the set of positive real numbers and $\mathbb{Q}^+$ is the set of positive rational numbers.

Theorem

$\mathbb{R}^+$ is the disjoint union of two nonempty sets, each closed under addition.

Proof: Let P be the set of all pairs (A, B) where

- $A, B \subseteq \mathbb{R}^+$, $A \cap B = 0$, $A \neq 0$ and $B \neq 0$,
- A and B are closed under addition and
- A and B are closed under multiplication by any positive rational number

Note that $(\mathbb{Q}^+, \sqrt{2}\mathbb{Q}^+) \in P$ so P is nonempty. Define a partial order $\prec$ on P by $(A_1, B_1) \prec (A_2, B_2)$ iff $A_1 \subseteq A_2$ and $B_1 \subseteq B_2$.

Example III

Suppose C is a chain in $(P, \preceq)$. Let A (respectively B) be the union of all the first (respectively second) coordinates of the pairs in C. Then it is easy to check that $(A, B) \in P$ and hence (A, B) is an upper bound of C. So by Zorn's lemma, P has a maximal member say $(A_\star, B_\star)$. It suffices to show that $A_\star \cup B_\star = \mathbb{R}^+$. Towards a contradiction , suppose not and fix $x \in \mathbb{R}^+ \setminus (A_\star \cup B_\star)$. Define $A_1 = \{a + rx : a \in A_\star, r \in \mathbb{Q}^+ \cup \{0\}\}$ and $B_1 = \{b + rx : b \in B_\star, r \in \mathbb{Q}^+ \cup \{0\}\}$. It is easy to see that A_1 and B_1 are both closed under addition and multiplication by a positive rational. Since $(A_\star, B_\star)$ is a maximal element of $(P, \prec)$, neither one of $(A_\star, B_1)$ and $(A_1, B_\star)$ is in P. This must mean that $A_\star \cap B_1 \neq \emptyset$ and $A_1 \cap B_\star \neq \emptyset$. Fix $y \in A_\star \cap B_1$ and $z \in A_1 \cap B_\star$. Choose $r, s \in \mathbb{Q}^+ \cup \{0\}$, $a \in A_\star$ and $b \in B_\star$ such that $y = b + rx$ and $z = a + sx$. Since $A_\star \cap B_\star = \emptyset$, both $r, s > 0$. As $a, y \in A_\star$ and $A_\star$ is closed under addition and multiplication by positive rationals, we get $sy + ra = ra + sb + srx \in A_\star$. Similarly, $rz + sb = sb + ra + rsx \in B_\star$. So $A_\star \cap B_\star \neq \emptyset$ which is a contradiction . □

Additive functions

Definition

A function $f : \mathbb{R} \to \mathbb{R}$ *is* additive *iff for every* $x, y \in \mathbb{R}$,

$$f(x + y) = f(x) + f(y)$$

Exercise: Suppose $f : \mathbb{R} \to \mathbb{R}$ is additive and $a = f(1)$.

- Show that $f(0) = 0$.
- Show that for every $x \in \mathbb{R}$, $f(-x) = -f(x)$.
- Show that for every $x \in \mathbb{Q}$, $f(x) = ax$.

Continuous additive functions

Proposition

Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous and additive. Let $f(1) = a$. Then for every $x \in \mathbb{R}$, $f(x) = ax$.

Proof: Let $\langle x_n : n < \omega \rangle$ be a sequence of rationals converging to x. By the previous exercise, $f(x_n) = ax_n$. By the continuity of f at x,

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} ax_n = a\left(\lim_{n\to\infty} x_n\right) = ax$$

□

Question

Are these the only additive functions?

$\mathbb{Q}$-linear independence

(a) $X \subseteq \mathbb{R}$ is $\mathbb{Q}$-**linearly independent** iff for every finite $\{x_1, x_2, \ldots, x_n\} \subseteq X$ and $a_1, a_2, \ldots, a_n \in \mathbb{Q}$,

$$(a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0) \implies (a_1 = a_2 = \cdots = a_n = 0)$$

(b) $H \subseteq \mathbb{R}$ is a **Hamel basis** iff H is a $\subseteq$-maximal $\mathbb{Q}$-linearly independent subset of $\mathbb{R}$.

Exercise: Suppose $H \subseteq \mathbb{R}$ is a Hamel basis. Then for every $0 \neq x \in \mathbb{R}$ can be uniquely written as $x = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ where $x_1, x_2, \ldots, x_n \in H$ and $a_1, a_2, \ldots a_n$ are nonzero rationals.
Exercise: Suppose $H \subseteq \mathbb{R}$ is a Hamel basis and $f : H \to \mathbb{R}$. Then there is a unique additive function $g : \mathbb{R} \to \mathbb{R}$ such that $f \subseteq g$.

Hamel basis

Theorem

Let $X \subseteq \mathbb{R}$ be $\mathbb{Q}$-linearly independent. Then there is a Hamel basis $H \subseteq \mathbb{R}$ such that $X \subseteq H$.

Proof: Let $\mathcal{F}$ be the family of all $\mathbb{Q}$-linearly independent sets Y such that $X \subseteq Y$. Then every $\subseteq$-chain C in $\mathcal{F}$ has an upper bound, namely $\bigcup C$ [Why?]. Hence by Zorn's lemma, $\mathcal{F}$ has a maximal member H. □

Corollary

There is a discontinuous additive function $f : \mathbb{R} \to \mathbb{R}$.

Proof: Since $\{1\}$ is $\mathbb{Q}$-linearly independent, by the previous theorem there is a Hamel basis $H \subseteq \mathbb{R}$ such that $1 \in H$. Define $f : H \to \mathbb{R}$ by $f(1) = 0$ and $f(x) = 1$ if $x \in H \setminus \{1\}$. Let $g : \mathbb{R} \to \mathbb{R}$ be the unique additive function such that $f \subseteq g$. Towards a contradiction, suppose g is continuous. Since $g(1) = 0$ and g is continuous, we must have $g(x) = x0 = 0$ for every $x \in \mathbb{R}$ – A contradiction. So g is not continuous. □

Cardinality I

Definition

1. *We say that A has* smaller cardinality *than B iff there is an injection from A to B.*
2. *We say that A and B have the* same cardinality *iff there is a bijection from A to B.*

Note that we haven't defined "cardinality of A" yet. This will be done later using the well-ordering theorem. The following are obvious.

1. A has smaller cardinality than A.
2. If A has smaller cardinality than B and B has smaller cardinality than C, then A has smaller cardinality than C.

Next, we'll prove the following: If A has smaller cardinality than B and B has smaller cardinality than B, then A and B have the same cardinality.

Schröder-Bernstein theorem

Theorem (ZF)

Suppose there is an injection from A to B and there is an injection from B to A. Then there is a bijection from A to B.

Proof: Fix injections $f : A \to B$ and $g : B \to A$. We'll construct a bijection $h : A \to B$. Recursively, define

- $A_0 = A$, $B_0 = B$ and
- for each $n < \omega$, $B_{n+1} = f[A_n]$, $A_{n+1} = g[B_n]$.

By induction on $n < \omega$, it is easy to check that for every $n < \omega$, $A_{n+1} \subseteq A_n$ and $B_{n+1} \subseteq B_n$. Define $A_\omega = \bigcap\{A_n : n < \omega\}$ and $B_\omega = \bigcap\{B_n : n < \omega\}$. Then we have the following.

(a) $A = A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots \supseteq A_n \supseteq A_{n+1} \supseteq \cdots \supseteq A_\omega$

(b) $B = B_0 \supseteq B_1 \supseteq B_2 \supseteq \cdots \supseteq B_n \supseteq B_{n+1} \supseteq \cdots \supseteq B_\omega$

Schröder-Bernstein theorem

Next, define

(i) $A_{\text{even}} = \bigcup\{A_{2n} \setminus A_{2n+1} : n < \omega\}$

(ii) $A_{\text{odd}} = \bigcup\{A_{2n+1} \setminus A_{2n+2} : n < \omega\}$

(iii) $B_{\text{even}} = \bigcup\{B_{2n} \setminus B_{2n+1} : n < \omega\}$

(iv) $B_{\text{odd}} = \bigcup\{B_{2n+1} \setminus B_{2n+2} : n < \omega\}$

Using (a) and (b) above, the following are clear.

(1) In each one of the equations (i)-(iv), the right hand side is the union of a disjoint family.

(2) $\{A_{\text{even}}, A_{\text{odd}}, A_\omega\}$ is a partition of A and $\{B_{\text{even}}, B_{\text{odd}}, B_\omega\}$ is a partition of B.

Schröder-Bernstein theorem

Claim

(3) $f \restriction A_{\text{even}}$ is a bijection from A_{even} to B_{odd}.

(4) $g \restriction B_{\text{even}}$ is a bijection from B_{even} to A_{odd}.

(5) $f \restriction A_\omega$ is a bijection from A_ω to B_ω.

Proof of Claim: Since f is injective,

$$f[A_{2n+1} \setminus A_{2n}] = f[A_{2n+1}] \setminus f[A_{2n}] = B_{2n+2} \setminus B_{2n+1}$$

Taking union over $n < \omega$, we get (3). The proof of (4) is similar. For (5), observe that

$$f[A_\omega] = f\left[\bigcap_{n<\omega} A_n\right] = \bigcap_{n<\omega} f[A_n] = B_\omega$$

where we use the fact that f is injective to interchange f and $\bigcap_{n<\omega}$.

Schröder-Bernstein theorem

Finally, define

$$h(x) = \begin{cases} f(x) & \textbf{if } x \in A_{\text{even}} \cup A_{\omega} \\ g^{-1}(x) & \textbf{if } x \in A_{\text{odd}} \end{cases} \tag{2}$$

Using (2)-(5), it is clear that $h : A \to B$ is a bijection. Note that this proof did not use the axiom of choice. □

Cardinality II

Recall that by the well-ordering theorem, every set can be well-ordered. Hence for every set X, there is an ordinal α and a bijection $f : \alpha \to X$.

Definition (Cardinality and cardinals)

1. *The* cardinality of X, *denoted* $|X|$, *is the least ordinal* α *such that there is a bijection between* X *and* α*.*
2. *A* cardinal *is an ordinal* α *such that* $|\alpha| = \alpha$*.*

We denote cardinals by higher Greek letters like κ, λ, δ, θ etc. $0, 1, 2, \ldots,$ are the finite cardinals. ω is the first infinite cardinal. $\omega + 1$ is not a cardinal since $|\omega + 1| = \omega$. Note that X is countable iff $|X| \leq \omega$.

Cardinality II

Exercise

1. *For every ordinal* α, $|\alpha| \leq \alpha$.
2. *If* κ *is a cardinal and* $\alpha < \kappa$, *then* $|\alpha| < \kappa$.
3. *There is an injection from* X *to* Y *iff* $|X| \leq |Y|$.
4. *There is a surjection from* X *to* Y *iff* $|Y| \leq |X|$.
5. *There is a bijection from* X *to* Y *iff* $|X| = |Y|$.

It follows that the previous definitions of "X has smaller cardinality than Y" and "X and Y have the same cardinality" are equivalent to "$|X| \leq |Y|$" and "$|X| = |Y|$" respectively.

There is no largest cardinal

Theorem (Cantor)

For any set X, there is no surjective function $f : X \to \mathcal{P}(X)$.

Proof.

Let $f : X \to \mathcal{P}(X)$. Define $Y = \{v \in X : v \notin f(v)\}$. We claim that $Y \notin \text{range}(f)$. Suppose not and let $a \in X$ be such that $f(a) = Y$. Then $a \in Y$ iff $a \notin f(a)$ iff $a \notin Y$ which is impossible. □

Corollary

For every cardinal κ, $|\mathcal{P}(\kappa)| > \kappa$.

Proof: Since κ injects into $\mathcal{P}(\kappa)$, $\kappa \leq |P(\kappa)|$. So either $\kappa < |\mathcal{P}(\kappa)|$ or $\kappa = |\mathcal{P}(\kappa)|$. The latter is ruled out by Cantor's theorem. □

Successor/Limit cardinals

Definition (Successor/Limit cardinals)

Suppose α is an ordinal and κ is a cardinal. Then

(a) *α^+ is the least cardinal $> \alpha$.*

(b) *κ is a* successor cardinal *iff $\kappa = \alpha^+$ for some α.*

(c) *κ is a* limit cardinal *iff κ is not a successor cardinal.*

Omega/aleph Hierarchy

Definition (Omega hierarchy)

Using transfinite recursion on α, define ω_α as follows.

(i) $\omega_0 = \omega$.

(ii) $\omega_{\alpha+1} = (\omega_\alpha)^+$.

(iii) *If α is a limit ordinal, then $\omega_\alpha = \sup(\{\omega_\beta : \beta < \alpha\})$.*

For historic reasons, sometimes people also write $\aleph_\alpha$ instead of ω_α. The first few cardinals are as follows.

$$0 < 1 < 2 \ldots \omega = \omega_0 < \omega_1 < \cdots < \omega_\omega < \omega_{\omega+1} < \cdots < \omega_{\omega+\omega} \ldots$$

Note that ω_α is a limit cardinal iff α is a limit ordinal.

Countable sets

Theorem

(a) $|\omega \times \omega| = \omega$.

(b) *For each* $1 \leq n < \omega$, $|\omega^n| = \omega$.

(c) $|\mathbb{Q}| = \omega$ *where* $\mathbb{Q}$ *is the set of rational numbers.*

(d) $|\mathbb{R}| \geq \omega_1$ *where* $\mathbb{R}$ *is the set of real numbers.*

Proof: (a) $(m, n) \mapsto 2^m 3^n$ defines an injection from $\omega \times \omega$ to ω. So $|\omega \times \omega| \leq \omega$. Clearly, $|\omega \times \omega| \geq \omega$. Hence $|\omega \times \omega| = \omega$. (b) Use induction on n. We leave the proof of (c) to the reader. (d) Since $\mathbb{R}$ is uncountable, $|\mathbb{R}| > \omega$. As ω_1 is the least cardinal $> \omega$, $|\mathbb{R}| \geq \omega_1$. □

Cardinality of products

Lemma

Suppose κ is an infinite cardinal. Then $|\kappa \times \kappa| = \kappa$.

Proof By transfinite induction on κ. If $\kappa = \omega$, then this holds. So assume $\kappa > \omega$ and for every cardinal $\theta < \kappa$, $|\theta \times \theta| = \theta$. Define an ordering $\prec$ (called the **max-lexicographic order**) on $\kappa \times \kappa$ as follows: $(\alpha_1, \beta_1) \prec (\alpha_2, \beta_2)$ iff

- either $\max(\{\alpha_1, \beta_1\}) < \max(\{\alpha_2, \beta_2\})$ or
- $\max(\{\alpha_1, \beta_1\}) = \max(\{\alpha_2, \beta_2\})$ and $\alpha_1 < \alpha_2$ or
- $\max(\{\alpha_1, \beta_1\}) = \max(\{\alpha_2, \beta_2\})$ and $\alpha_1 = \alpha_2$ and $\beta_1 < \beta_2$.

It is easy to check that $\prec$ is a well ordering on $\kappa \times \kappa$. If $\alpha < \kappa$ is infinite, then the set $\mathrm{pred}(\kappa \times \kappa, \prec, (\alpha, \alpha))$ of $\prec$-predecessors of (α, α) is contained in $(\alpha + 1) \times (\alpha + 1)$ and hence, by inductive hypothesis, has cardinality
$|(\alpha + 1) \times (\alpha + 1)| = ||\alpha + 1| \times |\alpha + 1|| = ||\alpha| \times |\alpha|| = |\alpha| \leq \alpha < \kappa$.
Since κ is a cardinal, it follows that every $\prec$-initial segment of $(\kappa \times \kappa, \prec)$ has order type $< \kappa$. So $\mathrm{type}(\kappa \times \kappa, \prec) = \kappa$. Hence $|\kappa \times \kappa| = \kappa$. □

Cardinality of products

Corollary

1. *If κ and λ are infinite cardinals, then $|\kappa \times \lambda| = \max(\{\kappa, \lambda\})$.*
2. *If X and Y are infinite sets, then $|X \cup Y| = |X \times Y| = \max(\{|X|, |Y|\})$.*
3. *If X is an infinite set and $1 \leq n < \omega$, then $|X^n| = |X|$. In particular, $|\mathbb{R}^n| = |\mathbb{R}|$.*

Proof: Use the previous theorem. □

Cardinalities of infinite unions

Lemma

Suppose κ is an infinite cardinal and $|X_\alpha| \leq \kappa$ for every $\alpha < \kappa$. Then $|\bigcup\{X_\alpha : \alpha < \kappa\}| \leq \kappa$.

Proof.

Put $X = \bigcup\{X_\alpha : \alpha < \kappa\}$. Fix a well ordering $\prec$ of $\mathcal{P}(X \times \kappa)$. Let h be a function with domain κ such that for every $\alpha < \kappa$, $h(\alpha)$ is the $\prec$-least injective function from X_α to κ. It follows that there is an injective function $g : X \to \kappa \times \kappa$ – Given $x \in X$, pick the least α such that $x \in X_\alpha$ and define $g(x) = (\alpha, h(\alpha)(x))$. It follows that $|X| \leq |\kappa \times \kappa| = \kappa$. □

Corollary

Suppose $\{X_n : n < \omega\}$ is a countable family of countable sets. Then $\bigcup\{X_n : n < \omega\}$ is countable.

Cardinality of $\mathbb{R}$

Definition

$\mathfrak{c} = |\mathbb{R}|$ *is the* continuum*.*

Recall that 2^ω is the set of all functions from ω to $2 = \{0, 1\}$.

Exercise

Show that $|2^\omega| = |\mathcal{P}(\omega)| = \mathfrak{c}$*.*

CH (Continuum hypothesis) is the statement $\mathfrak{c} = \omega_1$ and **GCH** (Generalized continuum hypothesis) is the statement: For every infinite cardinal κ, $|\mathcal{P}(\kappa)| = \kappa^+$.

Closures

Definition

We say that f *is a* finitary function *on* A *iff for some* $n < \omega$, $f : A^n \to A$.

Definition (Closure)

Suppose $f : A^n \to A$ *is a finitary function on* A *and* $B \subseteq A$.

(a) *We say that* B *is* closed under f *iff* $\mathit{range}(f \upharpoonright B^n) \subseteq B$.

(b) *We define the* closure of B under A *to be the set* $\bigcap\{C \subseteq A : B \subseteq C \;\&\; C \textit{ is closed under } f\}$

Cardinality of closures

Theorem

Let κ be an infinite cardinal. Suppose $B \subseteq A$, $|B| \leq \kappa$ and $\mathcal{F}$ is a set of $\leq \kappa$ finitary functions on A. Then there exists $C \subseteq A$ such that

(a) $B \subseteq C \subseteq A$,

(b) $|C| \leq \kappa$ *and*

(c) *for every* $f \in \mathcal{F}$, C *is closed under* f.

Proof: For $f \in \mathcal{F}$ and $D \subseteq A$, define $f \star D = \mathrm{range}(f \restriction D^n)$ where $f : A^n \to A$. Inductively, define $C_0 = B$ and $C_{n+1} = C_n \cup \bigcup\{f \star C_n : f \in \mathcal{F}\}$. Then, for every $n < \omega$, $|C_n| \leq \kappa$. Put $C = \bigcup\{C_n : n < \omega\}$ and note that $|C| \leq \kappa$. It is easy to see that $B \subseteq C \subseteq A$ and C is closed under every function in $\mathcal{F}$. □

Cardinality of Hamel basis

Lemma

Let $H \subseteq \mathbb{R}$ be a Hamel basis. Then $|H| = \mathfrak{c}$.

Proof: For each $1 \leq n < \omega$ and $\overline{a} \in \mathbb{Q}^n$, define $f_{\overline{a}} : \mathbb{R}^n \to \mathbb{R}$ by

$$f_{\overline{a}}(\overline{x}) = \sum_{k<n} a_k x_k$$

where $\overline{a} = \langle a_k : k < n \rangle$ and $\overline{x} = \langle x_k : k < n \rangle$. Let $\mathcal{F} = \{f_{\overline{a}} : 1 \leq n < \omega, \overline{a} \in \mathbb{Q}^k\}$. Then $|\mathcal{F}| = \omega$. By the previous theorem, there exists $C \subseteq \mathbb{R}$ such that $H \subseteq C$, $|C| \leq \max(\{|H|, \omega\})$ and C is closed under every function in $\mathcal{F}$. Since every real is a finite linear combination of members of H using coefficients in $\mathbb{Q}$, C must be $\mathbb{R}$. Hence $|\mathbb{R}| \leq \max(\{|H|, \omega\})$. As $\mathbb{R}$ is uncountable, it follows that $|H| = |\mathbb{R}| = \mathfrak{c}$. □

Hamel basis for $\mathbb{C}$

(a) $X \subseteq \mathbb{C}$ is $\mathbb{Q}$-**linearly independent** iff for every finite $\{x_1, x_2, \ldots, x_n\} \subseteq X$ and $a_1, a_2, \ldots, a_n \in \mathbb{Q}$,

$$(a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0) \implies (a_1 = a_2 = \cdots = a_n = 0)$$

(b) $H \subseteq \mathbb{C}$ is a **Hamel basis for** $\mathbb{C}$ iff H is a $\subseteq$-maximal $\mathbb{Q}$-linearly independent subset of $\mathbb{C}$.

Exercise: Suppose $H \subseteq \mathbb{C}$ is a Hamel basis for $\mathbb{C}$. Then for every $0 \neq x \in \mathbb{C}$ can be uniquely written as $x = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ where $x_1, x_2, \ldots, x_n \in H$ and $a_1, a_2, \ldots a_n$ are nonzero rationals.
Exercise: Show that a Hamel basis for $\mathbb{C}$ exists and every Hamel basis for $\mathbb{C}$ has cardinality $\mathfrak{c}$.

Proof of $(\mathbb{C}, +) \cong (\mathbb{R}, +)$

Proposition

There exists a bijection $f : \mathbb{R} \to \mathbb{C}$ *such that for every* $x, y \in \mathbb{R}$, $f(x + y) = f(x) + f(y)$.

Proof. Fix Hamel bases H_1 and H_2 for $\mathbb{R}$ and $\mathbb{C}$ respectively. Since $|H_1| = |H_2| = \mathfrak{c}$, there is a bijection $h : H_1 \to H_2$. Extend h to $f : \mathbb{R} \to \mathbb{C}$ as follows: If $x = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ where $a_1, a_2, \ldots, a_n \in \mathbb{Q}$ and $x_1, x_2, \ldots, x_n \in H_1$, then

$$f(x) = a_1h(x_1) + a_2h(x_2) + \cdots + a_nh(x_n)$$

It is easy to check that f is a bijection and for every $x, y \in \mathbb{R}$, $f(x + y) = f(x) + f(y)$. □

Avoiding mid-points

Call a subset $X \subseteq \mathbb{R}$ **mid-point free** iff whenever $a < b < c$ are in X, $a + c \neq 2b$.

Theorem (Rado)

There is a partition $\{A_n : n < \omega\}$ of $\mathbb{R}$ such that each A_n is mid-point free.

Proof: Let H be a Hamel basis for $\mathbb{R}$. Define a function F with $\mathrm{dom}(F) = \mathbb{R} \setminus \{0\}$ as follows. For each $x \in \mathbb{R} \setminus 0$, choose $x_1 < x_2 < \cdots < x_n$ in H and $a_1, a_2, \ldots, a_n$ nonzero rationals such that

$$x = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

and define $F(x) = (a_1, a_2, \ldots, a_n)$. It is clear that the range of F is countable.

Avoiding mid-points

So it suffices to show that if $x < y < z$, and if $F(x) = F(y) = F(z)$, then $x + z \neq 2y$. Suppose not. Let $(a_1, a_2, \dots, a_n)$ be the common value of F at x, y, z. Fix $x_1 < x_2 < \cdots < x_n$, $y_1 < y_2 < \cdots < y_n$ and $z_1 < z_2 < \cdots < z_n$ all in H such that

$$x = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

$$z = a_1z_1 + a_2z_2 + \cdots + a_nz_n$$

$$y = a_1y_1 + a_2y_2 + \cdots + a_ny_n$$

Comparing the coefficients of $min(x_1, y_1, z_1)$ on the two sides of $x + z = 2y$, we get $x_1 = y_1 = z_1$. Subtracting $2a_1x_1$ from both sides and repeating the argument with the coefficient of $min(x_2, y_2, z_2)$, we get $x_2 = y_2 = z_2$ and so on. It follows that $x = y = z$: Contradiction. □

Two-point sets

Definition

We say that $X \subseteq \mathbb{R}^2$ *is a* 2-point set *iff for every line* $\ell \subseteq \mathbb{R}^2$, $|X \cap \ell| = 2$*.*

Theorem (Mazurkiewicz)

2*-point sets exist.*

Exercise: Show that there is a subset X of plane such that for every line $\ell \subseteq \mathbb{R}^2$, $|X \cap \ell| = 10$.

Zorn's lemma?

Call a subset of the plane a **partial 2-point set** iff it meets every line at ≤ 2 points. Let $\mathcal{F}$ be the family of all partial 2-point sets ordered by inclusion. Every chain in $(\mathcal{F}, \subseteq)$ has an upper bound (its union). So we can find a $\subseteq$-maximal set $S \in \mathcal{F}$. Must S be a 2-point set? No, S could be a circle.

Constructing two-point sets

Let $\mathcal{L}$ be the family of all lines in plane. Note that $|\mathcal{L}| = |\mathbb{R}^2 \times \mathbb{R}^2| = |\mathbb{R}^2| = \mathfrak{c}$. Let $\langle \ell_\alpha : \alpha < \mathfrak{c} \rangle$ be an injective sequence with range $\mathcal{L}$. Using transfinite recursion, construct a sequence $\langle S_\alpha : \alpha < \mathfrak{c} \rangle$ of subsets of $\mathbb{R}^2$ such that the following hold.

1. $S_0 = 0$ and if γ is limit, then $S_\gamma = \bigcup_{\alpha<\gamma} S_\alpha$.
2. $|S_\alpha| \leq |\alpha + \omega| < \mathfrak{c}$.
3. No 3 points in S_α are collinear.
4. $\beta < \alpha \implies |S_\alpha \cap \ell_\beta| = 2$.

Having constructed S_α, $S_{\alpha+1}$ is obtained as follows. Let $\mathcal{T}$ be the set of lines that pass through 2 points in S_α. Let B be the set of points of intersection of ℓ_α with the lines in $\mathcal{T}$. Note that $|B| \leq |\alpha + \omega| < \mathfrak{c}$. By clause 3, $|S_\alpha \cap \ell_\alpha| \leq 2$ so we can add ≤ 2 points from $\ell_\alpha \setminus B$ to S_α to get $S_{\alpha+1}$. Having completed the construction, put $S = \bigcup_{\alpha<\mathfrak{c}} S_\alpha$. Then S is a 2-point set. □

Meeting every circle at 3 points

Theorem

There exists $X \subseteq \mathbb{R}^2$ such that for every circle $C \subseteq \mathbb{R}^2$, $|X \cap C| = 3$.

Proof: Let $\mathcal{E}$ be the family of all circles in plane. Note that $|\mathcal{E}| = |\mathbb{R}^2 \times \mathbb{R}^2| = |\mathbb{R}^2| = \mathfrak{c}$. Let $\langle C_\alpha : \alpha < \mathfrak{c} \rangle$ be an injective sequence with range $\mathcal{C}$. Using transfinite recursion, construct a sequence $\langle S_\alpha : \alpha < \mathfrak{c} \rangle$ of subsets of $\mathbb{R}^2$ such that the following hold.

1. $S_0 = 0$ and if γ is limit, then $S_\gamma = \bigcup_{\alpha < \gamma} S_\alpha$.
2. $|S_\alpha| \leq |\alpha + \omega| < \mathfrak{c}$.
3. No 4 points in S_α are concyclic.
4. $\beta < \alpha \implies |S_\alpha \cap C_\beta| = 3$.

Having constructed S_α, $S_{\alpha+1}$ is obtained as follows. Let $\mathcal{T}$ be the set of circles that pass through 3 points in S_α. Let B be the set of points of intersection of C_α with the circles in $\mathcal{T}$. Note that $|B| \leq |\alpha + \omega| < \mathfrak{c}$. By clause 3, $|S_\alpha \cap C_\alpha| \leq 3$ so we can add ≤ 3 points from $C_\alpha \setminus B$ to S_α to get $S_{\alpha+1}$. Having completed the construction, put $S = \bigcup_{\alpha < \mathfrak{c}} S_\alpha$. It is easy to check that S meets every circle at exactly 3 points. □

An equivalent form of CH

Definition

For $A \subseteq \mathbb{R}^2$ and $x \in \mathbb{R}$, define the vertical section *of A at x by $A_x = \{y \in \mathbb{R} : (x, y) \in A\}$. The* horizontal section *of A at y is defined by $A^y = \{x \in \mathbb{R} : (x, y) \in A\}$.*

Theorem (Sierpinski)

CH is equivalent to the following statement. There exists $A \subseteq \mathbb{R}^2$ such that every vertical section of A is countable and every horizontal section of $\mathbb{R}^2 \setminus A$ is countable.

Exercise

Show that there is no $A \subseteq \mathbb{R}^2$ such that every vertical section of A is finite and every horizontal section of $\mathbb{R}^2 \setminus A$ is finite.

Proof of Sierpinski's theorem

Proof: First assume $\mathfrak{c} = \omega_1$. Let $\prec$ be a well-ordering of $\mathbb{R}$ such that $\text{type}(\mathbb{R}, \prec) = \omega_1$. Note that for every $x \in \mathbb{R}$, $\text{pred}(\mathbb{R}, \prec, x)$ is countable. Define $A = \{(x, y) \in \mathbb{R}^2 : y \prec x\}$. Fix $a \in \mathbb{R}$. Then the vertical section of A at a, $A_a = \{y \in \mathbb{R} : y \prec a\}$ is countable. Next fix $b \in \mathbb{R}$ and note that the horizontal section of $\mathbb{R}^2 \setminus A$ at b, $(\mathbb{R}^2 \setminus A)^b = \{x \in \mathbb{R} : (x, b) \notin A\} = \{x \in \mathbb{R} : x \preceq b\}$ is also countable.

Proof of Sierpinski's theorem

Now assume *CH* fails. So $\mathfrak{c} > \omega_1$. It suffices to show the following.

Claim

Assume $\mathfrak{c} > \omega_1$. Let $A \subseteq \mathbb{R}^2$. Suppose for every $x \in \mathbb{R}$, A_x is countable. Then there exists $y \in \mathbb{R}$ such that $(\mathbb{R}^2 \setminus A)^y$ is uncountable.

Proof: Fix $X \subseteq \mathbb{R}$ such that $|X| = \omega_1$. Let $W = \bigcup\{A_x : x \in X\}$. Then $|W| \leq \omega_1$. Since $\mathfrak{c} > \omega_1$, we can fix $y \in \mathbb{R} \setminus W$. We claim that

$$X \subseteq (\mathbb{R}^2 \setminus A)^y = \{x : (x, y) \notin A\}$$

Suppose not and fix $x \in X$ such that $(x, y) \in A$. Then $y \in A_x$. Hence $y \in W$: Contradiction. As X is uncountable, it follows that $(\mathbb{R}^2 \setminus A)^y$ is also uncountable. □

Filters

Definition (Filters)

For an infinite set X, a filter on X *is a subfamily $\mathcal{F} \subseteq \mathcal{P}(X)$ satisfying the following conditions.*

(i) $0 \notin \mathcal{F}$ *and* $X \in \mathcal{F}$.

(ii) *For every* $A, B \in \mathcal{F}$, $A \cap B \in \mathcal{F}$.

(iii) *For every* $A \subseteq B \subseteq X$, *if* $A \in \mathcal{F}$, *then* $B \in \mathcal{F}$.

Examples

(a) Let $\mathcal{F} = \{A \subseteq \omega : |\omega \setminus A| < \omega\}$. $\mathcal{F}$ is called the cofinite filter on ω.

(b) Let $a \in X$ and $\mathcal{F} = \{A \subseteq X : a \in A\}$. Such filters are called **principal filters**.

Ultrafilters

Definition (Ultrafilters)

$\mathcal{F}$ is an ultrafilter on *X iff X is a filter on X and for every $A \subseteq X$, either $A \in \mathcal{F}$ or $(X \setminus A) \in \mathcal{F}$.*

The next lemma says that ultrafilters on X are precisely the $\subseteq$-maximal filters on X.

Lemma

Let $\mathcal{F}$ be a filter on X. Then $\mathcal{F}$ is an ultrafilter iff $\mathcal{F}$ is a maximal filter on X.

Proof: First assume $\mathcal{F}$ is an ultrafilter on X. Let $\mathcal{G}$ be a filter on X such that $\mathcal{F} \subseteq \mathcal{G}$. Towards a contradiction, suppose $\mathcal{G} \neq \mathcal{F}$. Fix $A \in \mathcal{G} \setminus \mathcal{F}$. Since $A \notin \mathcal{F}$, $(X \setminus A) \in \mathcal{F}$. As $\mathcal{F} \subseteq \mathcal{G}$, $(X \setminus A) \in \mathcal{G}$. Hence $\emptyset = A \cap (X \setminus A) \in \mathcal{G}$. A contradiction. Next suppose $\mathcal{F}$ is a maximal filter on X. Suppose $A \subseteq X$ and towards a contradiction, suppose both $A, X \setminus A$ are not in $\mathcal{F}$. Define $\mathcal{G} = \{C \subseteq X : (\exists B \in \mathcal{F})(A \cap B \subseteq C)\}$. Then it is easy to check that $\mathcal{G}$ is a filter on X, $\mathcal{F} \subseteq \mathcal{G}$ and $A \in \mathcal{G} \setminus F$. So $\mathcal{F}$ is not maximal: Contradiction. □

Ultrafilters

Theorem

Every filter $\mathcal{F}$ on an infinite set X can be extended to an ultrafilter on X.

Proof: Let $\mathfrak{F}$ be the family of all filters $\mathcal{G}$ on X such that $\mathcal{F} \subseteq \mathcal{G}$. It is easy to see that every chain $\mathfrak{C}$ in $(\mathfrak{F}, \subseteq)$ has an upper bound, namely $\bigcup C$. By Zorn's lemma, we can find a maximal member $\mathcal{U} \in \mathfrak{F}$. Then $\mathcal{F} \subseteq \mathcal{U}$ and $\mathcal{U}$ is a maximal filter on X. Hence by the previous lemma, $\mathcal{U}$ is an ultrafilter on X. □

Corollary

There exists a non-principal ultrafilter on ω.

Proof Let $\mathcal{F}$ be the cofinite filter on ω. Using the previous theorem, get an ultrafilter $\mathcal{U}$ on ω such that $\mathcal{F} \subseteq \mathcal{U}$. It is easy to check that $\mathcal{U}$ is not principal. □

Non-principal ultrafilters

Exercise

Let $\mathcal{U}$ be a non-principal ultrafilter on ω. Show the following.

1. *Every cofinite subset of ω is in $\mathcal{U}$.*
2. *If $X \in \mathcal{U}$ and $F \subseteq \omega$ is finite, then $X \setminus F \in \mathcal{U}$.*
3. *If $A \in \mathcal{U}$, and $B \cup C = A$, then either $B \in \mathcal{U}$ or $C \in \mathcal{U}$.*

Infinite Ramsey's theorem

Definition

1. *$[X]^n = \{A \subseteq X : |A| = n\}$ is the set of all n-element subsets of X. Members of $[X]^2$ are called* pairs in X*.*
2. *Suppose $f : [X]^n \to T$. We say that $Y \subseteq X$ is* f-homogeneous *iff $f \restriction [Y]^n$ is constant.*

Theorem (Ramsey)

For every $f : [\omega]^2 \to 2$, there exists an infinite $Y \subseteq \omega$ such that Y is f-homogeneous.

Corollary

For every $1 \leq N < \omega$ and $f : [\omega]^2 \to N$, there exists an infinite $Y \subseteq \omega$ such that Y is f-homogeneous.

Proof of infinite Ramsey's theorem

Fix $f : [\omega]^2 \to 2$. Let $\mathcal{U}$ be a non-principal ultrafilter on ω. For each $n < \omega$, define

$$W_n^0 = \{m < \omega : m > n \ \& \ f(\{m, n\}) = 0\}$$
$$W_n^1 = \{m < \omega : m > n \ \& \ f(\{m, n\}) = 1\}$$

Since $\{W_n^0, W_n^1\}$ is a partition of a cofinite subset in ω and $\mathcal{U}$ is a non-principal ultrafilter, exactly one of the sets W_n^0, W_n^1 belongs to $\mathcal{U}$. Let $k(n) < 2$ be such that $W_n^{k(n)} \in \mathcal{U}$. Fix $k < 2$ such that $T = \{n < \omega : k(n) = k\} \in \mathcal{U}$.
Inductively construct
$Y = \{j(0) < j(1) < \cdots < j(n) < j(n+1) < \dots\} \subseteq \omega$ such that

$$j(0) \in T \text{ and } j(n+1) \in T \cap \bigcap\{W_{j(r)}^k : r \leq n\}$$

This can be done because $T \cap \bigcap\{W_r^k : r \leq j_n\} \in \mathcal{U}$ (as $\mathcal{U}$ is closed under finite intersections). Now it is easy to check that $f \restriction [Y]^2$ takes the constant value k. □

Schur's theorem

Theorem (Schur)

Suppose $1 \leq N < \omega$ *and* $h : \omega \to N$*. Then there exist* $a < b < c < \omega$ *such that* $h(a) = h(b) = h(c)$ *and* $a + b = c$*.*

Proof: Define $f : [\omega]^2 \to N$ by $f(\{m < n\}) = h(n - m)$. By Ramsey's theorem, there are an infinite $X \subseteq \omega$ and $r < N$ such that range$(f \restriction [X]^2) = \{r\}$. Fix $\{k < m < n\} \subseteq X$ such that $n - m > m - k$. Let $a = m - k$, $b = n - m$ and $c = n - k$. Then $a < b < c$, $a + b = c$ and $h(a) = h(b) = h(c) = r$. □

Exercise

Suppose $1 \leq N < \omega$ *and* $h : \omega \to N$*. Then there exist* $a < b < c < d < e < \omega$ *such that* $h(a) = h(b) = h(c) = h(d) = h(e)$ *and* $a + b + c + d = e$*.*

Schur's theorem for product

Theorem

Suppose $1 \leq N < \omega$ *and* $h : \omega \to N$. *Then there exist* $a < b < c < \omega$ *such that* $h(a) = h(b) = h(c)$ *and* $ab = c$.

Proof: Let $P = \{2^k : k < \omega\}$. Define $f : [P]^2 \to N$ by $f(\{2^m < 2^n\}) = h(2^{n-m})$. By Ramsey's theorem, there are an infinite $X \subseteq P$ and $r < N$ such that range$(f \restriction [X]^2) = \{r\}$. Choose $\{2^k < 2^m < 2^n\} \subseteq X$ such that $n - m > m - k$. Then it is easily checked that $a = 2^{m-k}$, $b = 2^{n-m}$ and $c = 2^{n-k}$ are as required. □

Exercise

Suppose $1 \leq N < \omega$ *and* $h : \omega \to N$. *Then there exist* $a < b < c < d < e < \omega$ *such that* $h(a) = h(b) = h(c) = h(d) = h(e)$ *and* $abcd = e$.

Hindman's theorem

Theorem (Hindman)

Suppose $1 \leq N < \omega$ *and* $h : \omega \to N$. *Then there exist an infinite* $X \subseteq \omega$ *and* $r < N$ *such that for every* $\{m_1 < m_2 < \cdots < m_n\} \subseteq X$, $h(m_1 + m_2 + \cdots + m_n) = r$.

Note that Hindman's theorem generalizes Schur's theorem. There is an elegant proof of Hindman's theorem using ultrafilters. To describe it, we need to define idempotent ultrafilters.

Idempotent ultrafilters

Definition

For $A \subseteq \omega$ and an ultrafilter $\mathcal{U}$ on ω, define

$$A^{\mathcal{U}} = \{k < \omega : A - k \in \mathcal{U}\}$$

where $A - k = \{n - k : n \in A \;\&\; n \geq k\}$.

Definition

A non-principal ultrafilter $\mathcal{U}$ on ω is called an idempotent ultrafilter *iff for every $A \in \mathcal{U}$, $A^{\mathcal{U}} \in \mathcal{U}$.*

Theorem

Idempotent ultrafilters exist.

We won't prove this theorem here since, in addition to Zorn's lemma, its proof also requires some topological notions. Our next goal is to assume that idempotent ultrafilters exist and use them to prove Hindman's theorem.

Hindman's theorem

Proof of Hindman's theorem using idempotent ultrafilters: Let $h : \omega \to N$. Fix an idempotent ultrafilter $\mathcal{U}$ on ω. Choose $r < N$ and $A \in \mathcal{U}$ such that $h[A] = \{r\}$.
Inductively, define $\langle (k(n), A_n) : n < \omega \rangle$ as follows.

(a) $A_0 = A$ and $k(0) \in A_0 \cap A_0^{\mathcal{U}}$.

(b) $A_{n+1} = (A_n - k(n)) \cap A_n$, $k(n+1) \in A_{n+1} \cap A_{n+1}^{\mathcal{U}}$ and $k(n+1) > k(n)$.

Put $X = \{k(n) : n < \omega\}$.

Claim

For every $n_1 < n_2 < \cdots < n_j$, $k(n_1) + k(n_2) + \cdots + k(n_j) \in A_{n_1} \subseteq A$

Proof: Easily checked by induction on $j \geq 1$. □

It follows that for every $\{m_1 < m_2 < \cdots < m_n\} \subseteq X$, $m_1 + m_2 + \cdots + m_n \in A$ and hence $h(m_1 + m_2 + \cdots + m_n) = r$. □

Peano Arithmetic

Classical number theory studies the "structure" $(\omega, S, +, \cdot, 0)$. Almost all true statements about $(\omega, S, +, \cdot, 0, 1)$ can be derived from a simple list of axioms about natural numbers. The (first order) theory of these axioms is called Peano arithmetic (abbreviated PA). The axioms of PA are as follows.

1. $(\forall x)(S(x) \neq 0)$
2. $(\forall x, y)(S(x) = S(y) \implies x = y)$
3. $(\forall x)(x + 0 = x)$
4. $(\forall x, y)(x + S(y) = S(x + y))$
5. $(\forall x)(x \cdot 0 = 0)$
6. $(\forall x, y)(x \cdot S(y) = (x \cdot y) + x)$
7. Induction scheme: Suppose $\phi(x)$ is a property that can be expressed using only $0, S, +, \cdot$. Then the following is an axiom:

$$[\phi(0) \text{ and } (\forall x)[\phi(x) \implies \phi(S(x))]] \implies (\forall x)(\phi(x))$$

Models of PA

Definition

We say that $(X, S_\star, +_\star, \cdot_\star, 0_\star)$ is a model of PA iff the following hold.

(a) *X is a nonempty set and $0_\star \in X$.*

(b) *$S_\star : X \to X$.*

(c) *$+_\star : X \times X \to X$. For $x, y \in X$, we'll write $x +_\star y$ instead of $+_\star(x, y)$.*

(d) *$\cdot_\star : X \times X \to X$. For $x, y \in X$, we'll write $x \cdot_\star y$ instead of $\cdot_\star(x, y)$.*

(e) *Let ψ be any axiom of PA. Let $\psi_\star$ be obtained from ψ by replacing 0 by $0_\star$, S by $S_\star$, $+$ by $+_\star$, $\cdot$ by $\cdot_\star$ and by restricting the quantifiers of ψ to X. Then $\psi_\star$ holds.*

For example, if $\psi \equiv (\forall x, y)(x \cdot S(y) = (x \cdot y) + x)$, then
$\psi_\star \equiv (\forall x, y \in X)(x \cdot_\star S_\star(y) = (x \cdot_\star y) +_\star x)$.

Standard model

Theorem

$(\omega, S, +, \cdot, 0)$ *is a model of PA.*

Proof: Note that $0 = \emptyset$, $S(n) = n + 1$ is the ordinal successor of n, $+$ is the ordinal addition restricted to ω and $\cdot$ is the ordinal multiplication restricted to ω.

(1) $(\forall x \in \omega)(S(x) \neq 0)$: This is clear since 0 is not a successor ordinal.

(2) $(\forall x, y \in \omega)(S(x) = S(y) \implies x = y)$: Suppose $x, y \in \omega$ and $S(x) = S(y)$. We must have either $x = y$ or $x < y$ or $y < x$. If $x < y$, then $x + 1 \leq y < y + 1$ which is impossible. Similarly $y < x$ is impossible. So $x = y$.

Standard model

(3) $(\forall x \in \omega)(x + 0 = x)$: Clear from the definition of ordinal addition.

(4) $(\forall x, y \in \omega)(x + S(y) = S(x + y))$: Suppose $x, y \in \omega$. Since ordinal addition is associative, we have
$x + S(y) = x + (y + 1) = = (x + y) + 1 = S(x + y)$.

(5) $(\forall x \in \omega)(x \cdot 0 = 0)$: Clear from the definition of ordinal multiplication.

(6) $(\forall x, y \in \omega)(x \cdot S(y) = (x \cdot y) + x)$: Suppose $x, y \in \omega$. Then
$x \cdot S(y) = x \cdot (y + 1) = (x \cdot y) + x$.

Standard model

(7) Induction scheme : It suffices to check the following: If $W \subseteq \omega$ such that

$$0 \in W \text{ and } (\forall n)[n \in W \implies S(n) \in W]$$

then $W = \omega$. Fix $W \subseteq \omega$ such that $0 \in W$ and $(\forall n)[n \in W \implies S(n) \in W]$. Then W is an inductive set (see Slide 10). As ω is the intersection of all inductive sets, we must have $\omega \subseteq W$. Hence $W = \omega$.

It follows that $(\omega, S, +\cdot, 0)$ is a model of PA. ☕

Consequences of PA

Claim

The following is a theorem of PA.

$$(\forall x)[x = 0 \textit{ or } (\exists y)(x = S(y))]$$

Proof: Let $\phi(x)$ be the following property

$$x = 0 \text{ or } (\exists y)(x = S(y))$$

First note that $\phi(0)$ holds. Next observe that if $\phi(x)$ holds, then $\phi(S(x))$ also holds. Hence by the induction scheme for $\phi(x)$, we get $(\forall x)[\phi(x)]$. ☕

Consequences of PA

Theorem

The following hold in PA.

(a) *Addition is associative:* $(\forall x, y, z)[x + (y + z) = (x + y) + z]$.

(b) *Addition is commutative:* $(\forall x, y)[x + y = y + x]$.

Proof: (a) Let $\phi(z)$ be the formula $(\forall x, y)[x + (y + z) = (x + y) + z]$. We want to show $(\forall z)[\phi(z)]$. By the induction scheme, it suffices to show $\phi(0)$ and $(\forall z)(\phi(z) \implies \phi(S(z)))$. By axiom 3, we get $x + (y + 0) = x + y$ and $(x + y) + 0 = x + y$. Hence $x + (y + 0) = (x + y) + 0$. So $\phi(0)$ holds. Next assume $\phi(z)$ and we'll show $\phi(S(z))$. By axiom 4, we get $x + (y + S(z)) = x + S(y + z) =$ $= S(x + (y + z))$. Using $\phi(z)$, we get $S(x + (y + z)) = S((x + y) + z)$. Applying axiom 4 again, we have $S((x + y) + z) = (x + y) + S(z)$. Hence $x + (y + S(z)) = (x + y) + S(z)$. So $\phi(S(z))$ holds. Hence $(\forall z)(\forall x, y)[x + (y + z) = (x + y) + z]$.

Consequences of PA

(b) Let $\phi(y)$ be the formula $(\forall x)(x + y = y + x)$. We want to show $(\forall y)[\phi(y)]$. We use induction on y.
Step 1: We first show $\phi(0)$. By axiom 3, $x + 0 = x$ so it suffices to show $(\forall x)(0 + x = x)$. We show this by induction on x. If $x = 0$, then $0 + x = 0 + 0 = 0$ by axiom 3. Now assume $0 + x = x$ and we'll show $0 + S(x) = S(x)$. By axiom 4, $0 + S(x) = S(0 + x)$ and $S(0 + x) = S(x)$ by inductive hypothesis. So $0 + S(x) = S(x)$. Hence $(\forall x)(0 + x = x)$.
Step 2: Next, we show $(\forall x)(S(x) = 1 + x)$ where $1 = S(0)$. We use induction on x. If $x = 0$, then this is clear since $1 + 0 = 1$ (by axiom 3). So assume $S(x) = 1 + x$ and we'll show $S(S(x)) = 1 + S(x)$. By axioms 3 and 4, $S(S(x)) = S(S(x) + 0) = S(x) + S(0) = S(x) + 1$. By inductive hypothesis, $S(x) + 1 = (1 + x) + 1$. By part (a), $(1 + x) + 1 = 1 + (x + 1)$. By axioms 4 and 3, $1 + (x + S(0)) = 1 + S(x + 0) = 1 + S(x)$. Hence $S(S(x)) = 1 + S(x)$. It follows that $(\forall x)(S(x) = 1 + x)$

Consequences of PA

Step 3: Now assume $\phi(y)$ and we'll show $\phi(S(y))$. By axiom 4, $x + S(y) = S(x + y)$. By $\phi(y)$, $S(x + y) = S(y + x)$. By axiom 4, $S(y + x) = y + S(x)$. By Step 2, $y + S(x) = y + (1 + x)$. By part (a), $y + (1 + x) = (y + 1) + x$. By axioms 4 and 3, $y + 1 = y + S(0) = S(y + 0) = S(y)$. Hence $(y + 1) + x = S(y) + x$. Therefore $S(y + x) = S(y) + x$. So $\phi(S(y)$ holds.

Since $\phi(0)$ holds (Step 1) and $(\forall x)(\phi(x) \implies \phi(S(x)))$ holds (Step 3), it follows that $(\forall y)[\phi(y)]$ holds. Hence $(\forall x, y)(x + y = y + x)$. ☕

Consequences of PA

We list some frequently used consequences of PA here. They can all be proved from the axioms of PA.

(1) $(\forall x, y, z)[(x + y) + z = x + (y + z)]$

(2) $(\forall x, y)(x + y = y + x)$

(3) $(\forall x, y, z)[(x \cdot y) \cdot z = x \cdot (y \cdot z)]$

(4) $(\forall x, y)(x \cdot y = y \cdot x)$

(5) $(\forall x, y, z)[x \cdot (y + z) = (x \cdot y) + (x \cdot z)]$

(6) $(\forall x)(x + 0 = 0 + x = x)$

(7) $(\forall x)(x \cdot 1 = 1 \cdot x = x)$

(8) $(\forall x, y, z)[(x + y = x + z) \implies y = z]$

(9) $(\forall x, y, z)[(x \neq 0 \text{ and } (x \cdot y = x \cdot z)) \implies y = z]$

Non-standard models of PA

Fact (Skolem, 1933)

There are countable models of PA which are not isomorphic to the $(\omega, S, +, \cdot, 0)$*.*

We skip the proof which requires some background in first order logic.

The ring of integers

Let $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ denote the set of integers. Recall the basic properties of $(\mathbb{Z}, +, \cdot, 0, 1)$.

(1) $(\forall x, y, z \in \mathbb{Z})[(x + y) + z = x + (y + z)]$

(2) $(\forall x, y \in \mathbb{Z})(x + y = y + x)$

(3) $(\forall x \in \mathbb{Z})(x + 0 = 0 + x = x)$

(4) For every $x \in \mathbb{Z}$, there is a unique $-x \in \mathbb{Z}$ such that $x + (-x) = -x + x = 0$. We define $x - y = x + (-y)$.

(5) $(\forall x, y, z \in \mathbb{Z})[(x \cdot y) \cdot z = x \cdot (y \cdot z)]$

(6) $(\forall x, y \in \mathbb{Z})(x \cdot y = y \cdot x)$

(7) $(\forall x \in \mathbb{Z})(x \cdot 1 = 1 \cdot x = x)$

(8) $(\forall x, y, z \in \mathbb{Z})[x \cdot (y + z) = (x \cdot y) + (x \cdot z)]$

(9) $(\forall x, y, z \in \mathbb{Z})[(x + y = x + z) \implies y = z]$

(10) $(\forall x, y, z \in \mathbb{Z})[(x \neq 0 \text{ and } (x \cdot y = x \cdot z)) \implies y = z]$

The field of rationals

Let $\mathbb{Q}$ denote the set of rationals. Recall the basic properties of $(\mathbb{Q}, +, -, \cdot, 0, 1)$.

(1) $(\forall x, y, z \in \mathbb{Q})[(x+y)+z = x+(y+z)]$

(2) $(\forall x, y \in \mathbb{Q})(x+y = y+x)$

(3) $(\forall x \in \mathbb{Q})(x+0 = 0+x = x)$

(4) For every $x \in \mathbb{Q}$, there is a unique $-x \in \mathbb{Q}$ such that $x+(-x) = -x+x = 0$. We define $x - y = x + (-y)$.

(5) $(\forall x, y, z \in \mathbb{Q})[(x \cdot y) \cdot z = x \cdot (y \cdot z)]$

(6) $(\forall x, y \in \mathbb{Q})(x \cdot y = y \cdot x)$

(7) $(\forall x \in \mathbb{Q})(x \cdot 1 = 1 \cdot x = x)$

(8) For every nonzero $x \in \mathbb{Q}$, there is a unique $x^{-1} \in \mathbb{Q}$ such that $x \cdot x^{-1} = x^{-1} \cdot x = 1$. We define $x/y = x \cdot y^{-1}$.

(9) $(\forall x, y, z \in \mathbb{Q})[x \cdot (y+z) = (x \cdot y) + (x \cdot z)]$

(10) $(\forall x, y, z \in \mathbb{Q})[(x+y = x+z) \implies y = z]$

(11) $(\forall x, y, z \in \mathbb{Q})[(x \neq 0 \text{ and } (x \cdot y = x \cdot z)) \implies y = z]$

Order

The usual ordering on $\mathbb{Z}$ is the following:

$$\cdots < -(n+1) < -n < \cdots < -2 < -1 < 0 < 1 < 2 < \cdots < n < n+1 < \ldots$$

The usual ordering on $\mathbb{Q}$ can be defined as follows. For $a, b, m, n \in \mathbb{Z}$ where $m, n \geq 1$ we have

$$\frac{a}{n} < \frac{b}{m} \iff am < bn$$

From now on, we'll write xy instead of $x \cdot y$ for the product of $x, y \in \mathbb{Q}$.

Divisors, quotient and remainder

Suppose n, d are integers. We say that n is a **multiple** of d (equivalently, d is a **divisor** or **factor** of n) iff for some $k \in \mathbb{Z}$, $n = dk$. We write $d \mid n$ (read d divides n) iff d is a divisor of n.

Theorem

Suppose n, d are integers, $n \neq 0$ and $d \geq 1$. Then there are unique integers q (quotient) and r (remainder) such that $0 \leq r < d$ and $n = qd + r$.

Proof: Put $W = \{n - md : m \in \mathbb{Z} \text{ and } n - md \geq 0\}$. $W \neq \phi$ since $n + |n|d \in W$. Let $r = \min(W)$. Fix $q \in \mathbb{Z}$ such that $n - qd = r$. Note that $r < d$ since otherwise $n - (q+1)d = r - d \geq 0$ and so $r - d \in W$ which contradicts the minimality of r. This proves the existence of q, r. To see uniqueness, suppose $n = q_1 d + r_1 = q_2 d + r_2$ where $0 \leq r_1 \leq r_2 < d$. Then $r_2 - r_1 = d(q_1 - q_2)$. So d divides $r_2 - r_1$. But $0 \leq r_2 - r_1 < d$. Hence $r_2 - r_1 = 0$ and so $r_1 = r_2$. It also follows that $q_1 = q_2$. ☕

GCD

If F is a nonempty, finite set of nonzero integers, then the $\text{gcd}(F)$ is the **greatest common divisor** of all members of F. Since 1 is a divisor of every integer, it follows that $\text{gcd}(F) \geq 1$. We sometimes also write $\text{gcd}(n_1, n_2, \ldots, n_k)$ instead of $\text{gcd}(\{n_1, n_2, \ldots, n_k\})$

Theorem

Suppose a, b are nonzero integers. Then $gcd(a, b) = \min\{ma + nb : m, n \in \mathbb{Z} \text{ and } ma + nb > 0\}$. Furthermore, every common divisor of a, b is also a divisor of $gcd(a, b)$.

Proof: Let $W = \{ma + nb : m, n \in \mathbb{Z} \text{ and } ma + nb > 0\}$ and note that $W \neq \emptyset$. Put $d = \text{gcd}(a, b)$ and $d' = \min(W)$. Fix $m, n \in \mathbb{Z}$ such that $d' = ma + nb$. Since $d \mid a$ and $d \mid b$, it follows that $d \mid (ma + nb)$. So d divides d' and hence $d \leq d'$.

GCD

Using the previous theorem, write $a = qd' + r$ where $q \in \mathbb{Z}$ and $0 \leq r < d'$. Now $a = qd' + r = qma + qnb + r$ and so $r = (1 - qm)a - qnb$. If $r > 0$, then $r \in W$ and $r < d'$ which is impossible. So $r = 0$ and therefore d' divides a. A similar argument shows that d' divides b. Hence d' is a common divisor of a and b. Since d is the greatest common divisor of a and b, it follows that $d' \leq d$. As $d \leq d'$, it follows that $d = d'$.

Finally, suppose e is a common divisor of a, b. Then, e is also a divisor of $ma + nb = d' = d = \gcd(a, b)$. ☕

Coprimes

Definition

Suppose a and b are nonzero integers. We say that a, b are coprime iff $gcd(a, b) = 1$.

Corollary

a, b are coprime iff there exist $m, n \in \mathbb{Z}$ such that $am + bn = 1$.

Proof: Easily follows from the fact that $\gcd(a, b) = \min(W)$ where

$$W = \{ma + nb : m, n \in \mathbb{Z} \text{ and } ma + nb > 0\}$$

Primes and composites

An integer $n > 1$ is a **prime number** iff its only positive divisors are 1 and n. Otherwise, n is **composite**.

Lemma

Suppose p is prime and $p \mid ab$. Then either $p \mid a$ or $p \mid b$.

Proof: Suppose p does not divide a. Since $\gcd(a, p)$ divides p and p is prime, it follows that $\gcd(a, p) = 1$. Choose $m, n \in \mathbb{Z}$ such that $1 = ma + np$. Then $b = mab + nbp$. Since $p \mid ab$, it follows that $p \mid b$. ☕

Lemma

Let $n > 1$ be an integer. Then n has a prime factor.

Proof: If n is prime, we are done. So assume n is composite. Then $W = \{k : 1 < k < n \text{ and } k|n\}$ is a nonempty set. Let $p = \min(W)$. We claim that p is a prime. Suppose not. Then p has a divisor q such that $1 < q < p$. Now $q \mid p$ and $p \mid n$ implies $q \mid n$. Hence $q \in W$. But $q < p = \min(W)$ so this is impossible. Hence p is a prime. ☕

The set of primes is infinite

Theorem

There are infinitely many primes.

Proof: It suffices to show that for every prime p, there is a prime $p' > p$. Fix any prime p. Let

$$2 = p_1 < p_2 < \cdots < p_k = p$$

list all primes below p. Define $n = (p_1 p_2 \dots p_k) + 1$. Using the previous lemma, fix p' such that p' is a prime divisor of n. Since dividing n by any of the primes below p leaves the remainder 1, it follows that $p' \notin \{p_1, p_2, \dots, p_k\}$. Hence $p' > p$. ☕

Prime factorization

Theorem

Every integer $n \geq 2$ can be uniquely written as

$$n = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k}$$

where $p_1 < p_2 < \dots < p_k$ are primes and $a_1, a_2, \dots, a_k \geq 1$.

Proof: We first show existence of prime factorization by induction on n.
If $n = 1$, then $n = 2^1$ so this is clear.
So suppose $n > 2$ and assume that the result holds for all numbers $< n$.
Let p_1 be the least prime factor of n. Then $n = p_1 m$ where $1 \leq m < n$.
If $m = 1$, then $n = p_1^1$ and we are done. Otherwise $m \geq 2$ and by inductive hypothesis, $m = q_1^{b_1} q_2^{b_2} \dots q_k^{b_k}$ where $q_1 < q_2 < \dots < q_k$ are primes and $b_1, b_2, \dots, b_k \geq 1$. Since p_1 is the least prime factor of n, either $p_1 = q_1$ or $p_1 < q_1$. If $p_1 = q_1$, we have $n = q_1^{1+b_1} q_2^{b_2} \dots q_k^{b_k}$. If $p_1 < q_1$, we can write $n = p_1^1 q_1^{b_1} q_2^{b_2} \dots q_k^{b_k}$.

Prime factorization

Next, we show uniqueness. Suppose

$$n = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k} = q_1^{b_1} q_2^{b_2} \dots q_l^{b_l}$$

where where $p_1 < p_2 < \dots < p_k$ and $q_1 < q_2 < \dots < q_l$ are primes and $a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_l \geq 1$.
We first claim that $p_1 = q_1$. Suppose not. Suppose $p_1 < q_1$. Since $p_1 \mid (q_1^{b_1} q_2^{b_2} \dots q_l^{b_l})$ and p_1 is a prime, by a previous lemma, we must have $p_1 \mid q_j$ for some $1 \leq j \leq l$. But $p_1 < q_1 < q_2 \dots < q_l$ so this is impossible. Similarly $q_1 < p_1$ is impossible. So $p_1 = q_1$.
Next we claim that $a_1 = b_1$. Say $a_1 < b_1$. Then dividing n by $p_1^{a_1}$, we get $p_2^{a_2} \dots p_k^{a_k} = p_1^{b_1 - a_1} q_2^{b_2} \dots q_l^{b_l}$. But this is impossible since p_1 is not a factor of the left hand side. Similarly, $b_1 < a_1$ is impossible. Hence $a_1 = b_1$. Now we can cancel the factor $p_1^{a_1} = q_1^{b_1}$ from both sides and repeat the above argument k times to get $k = l$, $p_1 = q_1 \dots p_k = q_k$ and $a_1 = b_1, \dots a_k = b_k$. ☕

Modular arithmetic

Suppose a, b are integers and $n \geq 1$. We say that a **is congruent to** b **modulo** n and write $a \equiv b \pmod{n}$ iff $n \mid (a - b)$. The following are left as an exercise.

1. Let $E_n = \{(a, b) : a \equiv b \pmod{n}\}$. Then E_n is an equivalence relation on $\mathbb{Z}$ whose equivalence classes are $\{[r] : 0 \leq r < n\}$.
2. If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $a + c \equiv b + d \pmod{n}$ and $a - c \equiv b - d \pmod{n}$.
3. If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $ac \equiv bd \pmod{n}$.
4. If $a \equiv b \pmod{n}$ and $k \geq 1$, then $a^k \equiv b^k \pmod{n}$.
5. **Cancellation**: If $\gcd(c, n) = 1$ and $ac \equiv bc \pmod{n}$, then $a \equiv b \pmod{n}$.

$(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$

Definition

Fix $n \geq 1$ and define $E_n = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : a \equiv b \pmod n\}$. Then E_n is an equivalence relation on $\mathbb{Z}$ whose equivalence classes are $\{[r] : 0 \leq r < n\}$ where $[r] = \{a \in \mathbb{Z} : a \equiv r \pmod n\}$. We define

$$\mathbb{Z}/n\mathbb{Z} = \{[r] : 0 \leq r < n\}$$

Note that if r, r', s, s' are integers, $[r] = [r']$ and $[s] = [s']$, then $[r + s] = [r' + s']$ and $[rs] = [r's']$. So we can define addition and multiplication in $\mathbb{Z}/n\mathbb{Z}$ as follows: For $[r], [s] \in \mathbb{Z}/n\mathbb{Z}$, define $[r] + [s] = [r + s]$ and $[r] \cdot [s] = [rs]$.

Properties of $\mathbb{Z}/n\mathbb{Z}$

Let $n \geq 1$. Then the following hold.

(1) $(\forall x, y, z \in \mathbb{Z}/n\mathbb{Z})[(x + y) + z = x + (y + z)]$

(2) $(\forall x, y \in \mathbb{Z}/n\mathbb{Z})(x + y = y + x)$

(3) $(\forall x \in \mathbb{Z}/n\mathbb{Z})(x + [0] = [0] + x = x)$

(4) For every $x \in \mathbb{Z}/n\mathbb{Z}$, there is a unique $-x \in \mathbb{Z}$ such that $x + (-x) = -x + x = [0]$. We define $x - y = x + (-y)$.

(5) $(\forall x, y, z \in \mathbb{Z}/n\mathbb{Z})[(x \cdot y) \cdot z = x \cdot (y \cdot z)]$

(6) $(\forall x, y \in \mathbb{Z}/n\mathbb{Z})(x \cdot y = y \cdot x)$

(7) If $n \geq 2$, then $(\forall x \in \mathbb{Z}/n\mathbb{Z})(x \cdot [1] = [1] \cdot x = x)$

(8) $(\forall x, y, z \in \mathbb{Z}/n\mathbb{Z})[x \cdot (y + z) = (x \cdot y) + (x \cdot z)]$

(9) $(\forall x, y, z \in \mathbb{Z}/n\mathbb{Z})[(x + y = x + z) \implies y = z]$

Multiplicative inverse in $\mathbb{Z}/n\mathbb{Z}$

Lemma

Suppose $n \geq 2$ and $1 \leq a < n$, then $(\exists b \in \mathbb{Z})(ab \equiv 1 \pmod{n})$ iff a, n are coprime.

Proof: First suppose $\gcd(a, n) = 1$. Fix $b, c \in \mathbb{Z}$ such that $ba + cn = 1$. Modding out by n gives $ba \equiv 1 \pmod{n}$. Hence $(\exists b \in \mathbb{Z})(ab \equiv 1 \pmod{n})$.
Next suppose there is some $b \in \mathbb{Z}$ such that $ab \equiv 1 \pmod{n}$. Then $ab - 1$ is a multiple of n. Fix $c \in \mathbb{Z}$ such that $ab - 1 = cn$. So $ab - cn = 1$. Hence $\gcd(a, n) = 1$. ☕

Corollary

Suppose $n \geq 2$ and $r \in \{1, 2, \ldots, n-1\}$. Then

$$(\exists x \in \mathbb{Z}/n\mathbb{Z})([r] \cdot x = x \cdot [r] = [1]) \iff gcd(r, n) = 1$$

Units in $\mathbb{Z}/n\mathbb{Z}$

Definition (Multiplicative inverse and units)

Suppose $n \geq 2$ and $x, y \in \mathbb{Z}/n\mathbb{Z}$. We say that y is a multiplicative inverse *of x iff $x \cdot y = [1]$. We say that x is a unit in $\mathbb{Z}/n\mathbb{Z}$ iff x has a multiplicative inverse in $\mathbb{Z}/n\mathbb{Z}$.*

Suppose $x \in \mathbb{Z}/n\mathbb{Z}$ and y, z are multiplicative inverses of x. Then $x \cdot y = y \cdot x = [1]$ and $x \cdot z = z \cdot x = [1]$. It follows that $y = [1] \cdot y = (z \cdot x) \cdot y = z \cdot (x \cdot y) = z \cdot [1] = z$. So if x has a multiplicative inverse in $\mathbb{Z}/n\mathbb{Z}$, then it is unique. Note that the multiplicative inverse of a unit is also a unit.

Lemma

Suppose $n \geq 2$ and x, y are units in $\mathbb{Z}/n\mathbb{Z}$. Then $x \cdot y$ is also a unit in $\mathbb{Z}/n\mathbb{Z}$.

Proof: Fix x' and y' in $\mathbb{Z}/n\mathbb{Z}$ such that $x \cdot x' = y \cdot y' = [1]$. Then $x \cdot y \cdot x' \cdot y' = [1]$. Hence $x' \cdot y'$ is the multiplicative inverse of $x \cdot y$. So $x \cdot y$ is also a unit in $\mathbb{Z}/n\mathbb{Z}$. ☕

Prime fields

Theorem

Suppose $n \geq 2$. The following are equivalent.

(1) *n is prime.*

(2) *Every nonzero member of $\mathbb{Z}/n\mathbb{Z}$ is a unit.*

Proof: First suppose n is prime. Then for every $1 \leq r \leq n-1$, $\gcd(r,n) = 1$. By the previous corollary, $[r]$ has a multiplicative inverse in $\mathbb{Z}/n\mathbb{Z}$. So $[r]$ is a unit in $\mathbb{Z}/n\mathbb{Z}$.
Next suppose for every $1 \leq r \leq n-1$, $[r]$ is a unit in $\mathbb{Z}/n\mathbb{Z}$. Once again by the previous corollary, for every $1 \leq r \leq n-1$, $\gcd(r,n) = 1$. So n has no divisors between 2 and $n-1$. Hence n is prime. ☕

Fermat's little theorem

Lemma (Fermat's little theorem)

Let $p \geq 2$ be a prime. Suppose p does not divide a. Then $a^{p-1} \equiv 1$ $(mod\ p)$.

Proof: By modding out a with p, we can assume that $1 \leq a \leq p-1$. Work in $\mathbb{Z}/p\mathbb{Z}$. Let $M = \{[a], [2a], [3a], \ldots, [(p-1)a]\}$. We claim that $|M| = p-1$. Suppose not and fix $1 \leq m < n \leq p-1$ such that $[ma] = [na]$. Then $[(n-m)a] = [0]$ and so p divides $(n-m)a$. Since p is prime and p does not divide a, we get $p \mid (n-m)$. But $1 \leq n-m < p$ so this is impossible. It follows that $|M| = p-1$.
Since $[0] \notin M$, we get $M = \{[1], [2], [3], \ldots, [p-1]\}$. Taking products over all the members of M, we get

$$(a)(2a)(3a)\ldots((p-1)a) \equiv (1)(2)(3)\ldots(p-1) \text{ (mod p)}$$

So $a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$. Since p is prime, we have $\gcd(p, (p-1)!) = 1$. So we can cancel $(p-1)!$ from both sides to get $a^{p-1} \equiv 1 \pmod{p}$. ☕

Euler's totient function

Definition (Euler's totient function)

For each $n \geq 1$, define

$$\phi(n) = |\{k : 1 \leq k \leq n \text{ and } gcd(k, n) = 1\}|$$

So $\phi(n)$ is the number of positive integers below n which are coprime with n. Observe that $\phi(1) = 1$ and for every prime p, $\phi(p) = p - 1$. We leave the following as homework for the reader.

(1) $\phi(1) = 1$ and for every prime p, $\phi(p) = p - 1$.

(2) If p is prime and $k \geq 1$, then $\phi(p^k) = p^k - p^{k-1}$.

(3) If $\gcd(a, b) = 1$, then $\phi(ab) = \phi(a)\phi(b)$.

(4) For each $n \geq 2$, $\phi(n)$ is the number of units in $\mathbb{Z}/n\mathbb{Z}$.

Euler's theorem

Theorem

Suppose $gcd(a, n) = 1$*. Then* $a^{\phi(n)} \equiv 1 \ (mod\ n)$*.*

Proof: As before, we can assume $1 \leq a \leq n - 1$. Define $W = \{k : 1 \leq k \leq n \text{ and } \gcd(k, n) = 1\}$. Note that $|W| = \phi(n)$ and $\{[k] : k \in W\}$ is the set of all units in $\mathbb{Z}/n\mathbb{Z}$.

Let $M = \{[ka] : k \in W\}$. We will show that M is the set of all units in $\mathbb{Z}/n\mathbb{Z}$. First we show that every member of M is a unit: Let $k \in W$. Then $[ka] = [k] \cdot [a]$ is the product of two units in $\mathbb{Z}/n\mathbb{Z}$. Hence $[ka]$ is also a unit. So every member of M is a unit.

Next we show that $|M| = \phi(n)$. As $\mathbb{Z}/n\mathbb{Z}$ has $\phi(n)$ units, this would imply that every unit is in M. Suppose k_1, k_2 are in W. We will show that $[k_1 a] = [k_2 a]$ implies $k_1 = k_2$. Since $\gcd(a, n) = q$, we can find $x \in \mathbb{Z}/n\mathbb{Z}$ such that $[a] \cdot x = [1]$. Now $[k_1] = [k_1] \cdot ([a] \cdot x) = [k_1 a] \cdot x = [k_2 a] \cdot x = [k_2] \cdot ([a] \cdot x) = [k_2]$. As $1 \leq k_1, k_2 \leq n$, we get $k_1 = k_2$. Hence $|M| = \phi(n)$ and it follows that M has all the units in $\mathbb{Z}/n\mathbb{Z}$.

Euler's theorem

Let $1 = k_1 < k_2 < \cdots < k_{\phi(n)}$ list the members of W. Then

$$M = \{[k_j a] : j \leq \phi(n)\} = \{[k_j] : j \leq \phi(n)\}$$

By taking products we get

$$(k_1 a)(k_2 a) \ldots (k_{\phi(n)} a) \equiv (k_1)(k_2) \ldots (k_{\phi(n)}) \text{ (mod n)}$$

Put $K = k_1 k_2 \ldots k_{\phi(n)}$ and note that $\gcd(K, n) = 1$. Then $Ka^{\phi(n)} \equiv K \pmod{n}$. As $\gcd(K, n) = 1$, we can cancel K to get $a^{\phi(n)} \equiv 1 \pmod{n}$. ☕

Chinese remainder theorem

Lemma

Suppose n_1, n_2 are positive integers and $gcd(n_1, n_2) = 1$. Then for every integer x, if $n_1 \mid x$ and $n_2 \mid x$, then $n_1 n_2 \mid x$.

Proof: Since $\gcd(n_1, n_2) = 1$, there are integers m_1, m_2 such that $m_1 n_1 + m_2 n_2 = 1$. Then $x = x m_1 n_1 + x m_2 n_2$ and each summand on the right side is a multiple of $n_1 n_2$. So $n_1 n_2 \mid x$. □

Theorem

Suppose $\langle n_i : 1 \leq i \leq k \rangle$ is a sequence of positive integers which are pairwise coprime. Let $N = n_1 n_2 \ldots n_k$. Let $\langle a_i : 1 \leq i \leq k \rangle$ be a sequence of arbitrary integers. Then there exists an integer x such that for every $1 \leq i \leq k$, $x \equiv a_i \pmod{n_i}$. Moreover, if x, x' are any two integers satisfying these modular equations, then $x \equiv x' \pmod{N}$.

Proof: Suppose $x \equiv a_i \pmod{n_i}$ and $x' \equiv a_i \pmod{n_i}$ for every $1 \leq i \leq N$. Then $x - x'$ is divisible by each n_i. As x_i's are pairwise coprime, it follows that $N = n_1 n_2 \ldots n_k$ also divides $x - x'$. Hence $x \equiv x'$ $\pmod{N}$.

Chinese remainder theorem

Next we show existence of such an x. Let $N_i = N/n_i$. Then $\gcd(N_i, n_i) = 1$. Hence, for each $1 \leq i \leq k$, there are integers M_i, m_i such that $M_i N_i + m_i n_i = 1$. Define

$$x = \sum_{i=1}^{k} a_i M_i N_i$$

We claim that x is as required. To see this, fix $1 \leq j \leq N$ and we'll show that $n_j \mid (x - a_j)$. Now

$$x - a_j = a_j(M_j N_j - 1) + \sum_{1 \leq i \leq k}^{i \neq j} a_i M_i N_i = -a_j m_j n_j + \sum_{1 \leq i \leq k}^{i \neq j} a_i M_i N_i$$

Note that all the summands on the right side are divisible by n_j. Hence $n_j \mid x$. ☕

A puzzle

A mouse is trying to eat his way through a $3 \times 3 \times 3$ cube of cheese by tunneling through all of the 27 subcubes. He starts at a subcube on some face of the cube and always moves on to an uneaten neighboring subcube. Can he finish at the center of the subcube?

Graphs

A **graph** G is a pair $G = (V, E)$ where $E \subseteq [V]^2$ (Recall that $[V]^2$ is the set of all subsets of E of size 2). Members of V are called **vertices** and members of E are called **edges**. If $e = \{x, y\}$ is an edge then we say that e is **incident** on x and y or e **joins** x and y. We say that x, y are **adjacent** or **neighbors** iff $\{x, y\} \in E$.

Let $G = (V, E)$ be a graph and $v \in V$. Define the set of **neighbors** of v in G by $N_G(v) = \{w \in V : \{v, w\} \in E\}$. The **degree** of v is the cardinality of the set of neigbors of v in G:

$$\deg_G(v) = |N_G(v)|$$

G is κ-**regular** iff for every $v \in V$, $\deg_G(v) = \kappa$. G is **regular** iff it is κ-regular for some κ.

Examples

1. V is any set and $E = [V]^2$. This is the **complete graph on** V. If $V = n = \{0, 1, 2, \ldots, n-1\}$, we write K_n to denote this graph.
2. $V = n$ and $E = \{\{k, k+1 : k < n\}\} \cup \{\{0, n-1\}\}$. This is called the **cycle** on n vertices.
3. $V = X \cup Y$ where $X \cap Y = \emptyset$ and $E = \{\{x, y\} : x \in X, y \in Y\}$. This the **complete bipartite graph on** (X, Y).
4. $V = \mathbb{R}^2$ and $E = \{\{x, y\} : ||x - y|| = 1\}$. This is the **unit distance graph in** $\mathbb{R}^2$.
5. $V = \omega$ and $\{k < n\} \in E$ iff the kth bit of the binary representation of n is 1. For example $\{2, 6\} \in E$ and $\{0, 6\} \notin E$ because the binary representation of 6 is 110 in which the 2nd bit is 1 and the 0th bit is 0. This is the **Ackermann-Rado graph**.

Walks, paths and cycles

Let $G = (V, E)$ be a graph.

(a) A **walk in** G is a finite sequence $W = \langle v_k : k < n \rangle$ of vertices such that for every $k < n$, $\{v_k, v_{k+1}\} \in E$. n is the **length of the walk** W. v_0 is the **initial vertex** of W and v_{n-1} is the **terminal vertex** of W. We also say that W is a **walk from** v_0 **to** v_n.

(b) $P = \langle v_k : k < n \rangle$ is a **path** in G iff P is a walk in G and for every $j < k < n$, $v_j \neq v_k$.

(b) $C = \langle v_k : k < n \rangle$ is a **cycle** in G iff C is a path in G and $\{v_0, v_{k-1}\} \in E$. An n-**cycle** is a cycle of length $n \geq 3$. G is **acyclic** iff it has no cycles of length ≥ 3.

(c) G is **connected** iff for every $x, y \in V$, there is walk from x to y.

(d) G is a **tree** iff it is connected and acyclic.

Hamiltonian graphs

Let $G = (V, E)$ be a finite graph (This means that $|V| < \omega$). We say that $C = \langle v_k : k < n \rangle$ is a **Hamiltonian cycle** in G iff C is a cycle in G and $V = \{v_k : k < n\}$. G is a **Hamiltonian graph** iff there is a Hamiltonian cycle in G.

Theorem (Ore, 1960)

Let $G = (V, E)$ be a finite graph on $n \geq 3$ vertices. Suppose for $v, w \in V$, if $\{v, w\} \notin E$, then $deg_G(v) + deg_G(w) \geq n$. Then G is Hamiltonian.

Proof: Let m be the largest possible length of a path in G. Let $P = \langle v_k : 1 \leq k \leq m \rangle$ be a path in G of length m.

(a) If u, w are any two non-adjacent vertices, then there is a vertex that is adjacent to both of them. Otherwise, each of the $n - 2$ vertices in $V \setminus \{u, w\}$ can be adjacent to at most one of u, w which implies that $\deg_G(u) + \deg_G(w) \leq n - 2$: Contradiction.

(b) Every vertex which is adjacent to either one of v_1, v_m is in the set $\{v_1, v_2, \ldots v_m\}$. Otherwise, P can be extended to a longer path.

Ore's theorem

(c) We next show that there is a path $P' = \langle v'_k : 1 \leq k \leq m \rangle$ in G such that v'_1, v'_m are adjacent. If P is such a path, we are done. So suppose v_1, v_m are not adjacent.

We first claim that for some $2 \leq k \leq m-2$, $\{v_1, v_{k+1}\} \in E$ and $\{v_k, v_m\} \in E$. Suppose there is no such k. By (b), there are at least $\deg_G(v_m) - 1$ vertices among $\{v_2, \ldots, v_{m-2}\}$ which are adjacent to v_m. We have assumed that for every $2 \leq k \leq m-2$, if $\{v_k, v_m\} \in E$, then v_{k+1} is not adjacent to v_1. Also, by (b), every vertex adjacent to v_1 is in the set $\{v_2, \ldots, v_{m-1}\}$. Hence $\deg_G(v_1) \leq (m-2) - (\deg_G(v_m) - 1)$ which implies that $\deg_G(v_1) + \deg_G(v_m) \leq m - 1 \leq n - 1$ which is impossible. So the claim holds and we can fix some $2 \leq k \leq m-2$ such that $\{v_1, v_{k+1}\} \in E$ and $\{v_k, v_m\} \in E$.

It now follows that $P' = \langle v_1, \ldots, v_k, v_m, v_{m-1}, \ldots, v_{k+1} \rangle$ is as required. By replacing P' by P, we can assume that v_1, v_m are adjacent.

Ore's theorem

So it suffices to show that $m = n$ since then P would be a Hamiltonian cycle in G. Suppose $m < n$. Fix $u \in V \setminus \{v_k : 1 \leq k \leq n\}$. By (b), u is not adjacent to either one of v_1, v_m. Using (a), fix a vertex w which is adjacent to both v_1 and u. Since w is adjacent to v_1, by (b) $w \in \{v_k : 2 \leq k \leq m\}$. Since w is adjacent to u and u is not adjacent to v_m, we must have $w \neq v_m$. So we can fix $2 \leq k \leq m - 1$ such that $w = w_k$. It now follows that $\langle u, v_k, v_{k-1}, \ldots, v_1, v_m, v_{m-1}, \ldots, v_{k+1}\rangle$ is a path in G of length $m + 1 > m$ – A contradiction. So $m = n$ and the proof is complete. ☕

Subgraphs and induced subgraphs

Let $G = (V_1, E_1)$ and $H = (V_2, E_2)$ be graphs. We say that H is a **subgraph** of G iff $V_2 \subseteq V_1$ and $E_2 \subseteq E_1$.

Let $G = (V, E)$ be a graph and $V' \subseteq V$. The **induced subgraph** of G on V' is the graph (V', E') where $E' = E \cap [V']^2$.

Bipartite graphs and perfect matching

We say that a graph $G = (V, E)$ is **bipartite with bipartition** (A, B) iff the following hold.

(a) $V = A \cup B$ and $A \cap B = \emptyset$.

(b) For every edge $\{x, y\} \in E$, $|\{x, y\} \cap A| = |\{x, y\} \cap B| = 1$.

Suppose G is a finite bipartite graph with bipartition (A, B) where $|A| = |B|$. A subgraph M of G is a **perfect matching** for G iff the following hold.

(i) The vertex set of M is $A \cup B$.

(ii) For every $a \in A$, there is a **unique** $b \in B$ such that $\{a, b\}$ is an edge in M.

So the edges in M give a one-to-one correspondence between A and B.

Hall's marriage theorem

Let $G = (V, E)$ be a finite bipartite graph with bipartition (A, B) where $|A| = |B|$. We say that G satisfies the **marriage condition** iff for every $S \subseteq A$, $|N_G(S)| \geq |S|$ where

$$N_G(S) = \{b \in B : (\exists a \in S)(\{a, b\} \in E)\}$$

It should be clear that if G has a perfect matching, then it satisfies the marriage condition. Hall's theorem says that the converse is also true.

Theorem (Hall, 1935)

Suppose $G = (V, E)$ is a bipartite graph with bipartition (A, B) and $|A| = |B|$. Assume G satisfies the marriage condition. Then G has a perfect matching.

Hall's marriage theorem

Proof: Let $H = (V, E')$ be a subgraph of G (so $E' \subseteq E$) such that H satisfies the marriage condition and is **edge-minimal** among all such subgraphs of G. This means that for every $e \in E'$, the graph $(V, E' \setminus \{e\})$ does not satisfy the marriage condition.

Claim

For every $a \in A$, $deg_H(a) = 1$.

Proof of Claim: Suppose not and we'll produce a contradiction. Choose $a \in A$ with $\deg_H(a) \neq 1$. Since H satisfies the marriage condition, $\deg_H(a) > 0$. So $\deg_H(a) \geq 2$. Fix $b_1 \neq b_2$ such that $\{b_1, a\} \in E'$ and $\{b_2, a\} \in E'$.
For each $i \in \{1, 2\}$, put $H_i = (V, E' \setminus \{a, b_i\})$. Since H_i has one fewer edge than H, it does not satisfy the marriage condition. So we can fix $S_i \subseteq A$ such that $|N_{H_i}(S_i)| < |S_i|$. Observe that $a \in S_1 \cap S_2$ since otherwise H would violate the marriage condition.

Hall's marriage theorem

Put $T_i = N_{H_i}(S_i)$. Note that $N_H((S_1 \cap S_2) \setminus \{a\}) \subseteq T_1 \cap T_2$ and $T_1 \cup T_2 = N_H(S_1 \cup S_2)$. It follows that $|N_H((S_1 \cap S_2) \setminus \{a\})| \leq |T_1 \cap T_2| = |T_1| + |T_2| - |T_1 \cup T_2| \leq$ $\leq |S_1| - 1 + |S_2| - 1 - |T_1 \cup T_2|$. Since H satisfies the marriage condition, $|T_1 \cup T_2| \geq |S_1 \cup S_2|$. Hence $|N_H((S_1 \cap S_2) \setminus \{a\})| \leq |S_1| + |S_2| - |S_1 \cup S_2| - 2 = |S_1 \cap S_2| - 2$. But now the marriage condition for H fails for $S = (S_1 \cap S_2) \setminus \{a\}$ which is a contradiction. So the claim is true.

Using the claim, we can now show that H is a perfect matching for G. Since every $a \in A$ has exactly one neighbour in B, it is enough to show that these neighbours are pairwise distinct. If not, choose $a_1 \neq a_2$ from A with common neighbour $b \in B$. But this means that H violates the marriage condition via $S = \{a_1, a_2\}$ which is impossible. ☕

Hall's marriage theorem

By an almost identical proof, we also have the following version of Hall's theorem where we do not assume $|A| = |B|$.

Theorem

Let G be a finite bipartite graph with bipartition (A, B). Suppose G satisfies the marriage condition (so $|A| \leq |B|$). Then there exists $B' \subseteq B$ with $|B'| = |A|$ such that the induced subgraph G' of G on $A \cup B'$ has a perfect matching.

Transversals

Let S be a nonempty finite set. Suppose $\mathcal{F} = \langle X_1, X_2, \ldots, X_n \rangle$ is a finite sequence (with possible repetition) of nonempty subsets of S. We say that a function $T : \mathcal{F} \to S$ is a **transversal** for $\mathcal{F}$ iff T is one-one and for every $1 \leq i \leq n$, $T(X_i) \in X_i$.

Example: Let $S = \{1, 2, 3, 4, 5, 6, 7\}$ and $\mathcal{F} = \langle X_1, X_2, X_3, X_4, X_5 \rangle$ where $X_1 = \{1, 5, 6\}, X_2 = \{5, 6\}, X_3 = \{5, 6\}, X_4 = \{2, 3, 7\}, X_5 = \{2, 3, 4\}$. Define $T : \mathcal{F} \to \{1, 2, 3, 4, 5, 6, 7\}$ by setting $T(X_1) = 1$, $T(X_2) = 5$, $T(X_3) = 6$, $T(X_4) = 2$ and $T(X_5) = 3$. Then T is a transversal for $\mathcal{F}$.

Let $\mathcal{F}$ be as above. We say that $\mathcal{F}$ satisfies the **marriage condition** iff for every $1 \leq k \leq n$, for every $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, the following holds: $|X_{i_1} \cup X_{i_2} \cup \cdots \cup X_{i_k}| \geq k$.

Observe that if $\mathcal{F}$ has a transversal, then it satisfies the marriage condition. The next theorem says that this is also sufficient.

Transversals

Theorem

Let S be a nonempty finite set. Suppose $\mathcal{F} = \langle X_1, X_2, \ldots, X_n \rangle$ is a finite sequence (with possible repetition) of nonempty subsets of S. Suppose $\mathcal{F}$ satisfies the marriage condition. Then $\mathcal{F}$ has a transversal.

Proof: Define a bipartite graph G with bipartition $(\{1, 2, \ldots, n\}, S)$ as follows: For $1 \leq i \leq n$ and $y \in S$, $\{i, y\}$ is an edge in G iff $y \in X_i$. Note that G satisfies the marriage condition. Hence by Hall's marriage theorem, there exists $S' \subseteq S$ with $|S'| = n$ and a perfect matching M of the induced subgraph G' of G on $\{1, 2, \ldots, n\} \cup S'$. Define $T : \mathcal{F} \to S$ by setting $T(X_i) = y$ where $y \in S'$ and $\{i, y\}$ is an edge in M. Then T is a transversal for $\mathcal{F}$. ☕

Connected components

Let $G = (V, E)$ be a graph. Define a relation R on V by $(x, y) \in R$ iff there is a walk/path from x to y in G. It is clear that R is an equivalence relation on V. A **component** of G is the induced subgraph of G on an R-equivalence class. We leave the following as exercise for the reader.

Exercise: H is a component of G iff H is a maximal connected induced subgraph of G. A graph is connected iff it has exactly one component.

Chromatic number

Let $G = (V, E)$ be a graph. A **vertex coloring** of G is a function f such that $\mathrm{dom}(f) = V$ and for every edge $\{x, y\} \in E$, $f(x) \neq f(y)$. A vertex coloring f is called a κ-coloring iff $\mathrm{range}(f) \subseteq \kappa$. The **chromatic number** of G, denoted $\chi(G)$, is the least cardinal κ such that there is a vertex coloring $f : V \to \kappa$ of G.

Examples

1. $\chi(K_n) = n$ where K_n is the complete graph on n-vertices.
2. $\chi(G) \leq 2$ for every bipartite graph G.
3. $\chi(C_n) = 2$ where C_n is the n-cycle and n is even.
4. $\chi(C_n) = 3$ where C_n is the n-cycle and n is odd.

Brooks' theorem

For a finite graph $G = (V, E)$, define

$$\Delta(G) = \max(\{\deg_G(x) : x \in V\})$$

Brooks' theorem

For a finite graph $G = (V, E)$, define

$$\Delta(G) = \max(\{\deg_G(x) : x \in V\})$$

Theorem (Brooks, 1941)

Suppose $G = (V, E)$ is a finite connected graph that is neither a complete graph nor an odd cycle. Then $\chi(G) \leq \Delta(G)$.

Proof: See video. ☕

De Bruijn-Erdős theorem

Theorem

Suppose $G = (V, E)$ is an infinite graph and $1 \leq n < \omega$. Suppose for every finite $A \subseteq V$, the chromatic number of the induced subgraph of G on A is at most n. Then $\chi(G) \leq n$.

Proof sketch: Let $\mathcal{F}$ be the family of all pairs (W, f) such that $W \subseteq V$, $f : W \to n$ is a vertex-coloring of the induced subgraph of G on W **and** for every finite $A \subseteq V$, f extends to a vertex coloring $g : W \cup A \to n$ of the induced subgraph of G on $W \cup A$. Define a partial order $\preceq$ on $\mathcal{F}$ by $(W_1, f_1) \preceq (W_2, f_2)$ iff $W_1 \subseteq W_2$ and $f_1 = f_2 \restriction W_1$. Note that $(\emptyset, \emptyset) \in \mathcal{F}$ so $\mathcal{F}$ is nonempty. We leave the following as an exercise.

(1) Every chain in $(\mathcal{F}, \preceq)$ has an upper bound.

(2) If (W, f) is a $\preceq$-maximal member of $\mathcal{F}$, then $W = V$ and f is a vertex coloring of G.

It follows that $\chi(G) \leq n$. ☕

Unit distance graph

Let $G = (\mathbb{R}^2, E)$ be the unit distance graph in plane. Recall that $E = \{\{x, y\} : ||x - y|| = 1\}$. The **Hadwiger-Nelson problem** asks for the value of $\chi(G)$. It was known for a long time that $4 \leq \chi(G) \leq 7$. In 2018, Aubrey de Grey showed that $\chi(G) \geq 5$.

Unit distance graph

Lemma

Let $G = (\mathbb{R}^2, E)$ be the unit distance graph in plane. Then $\chi(G) \geq 4$.

Proof:

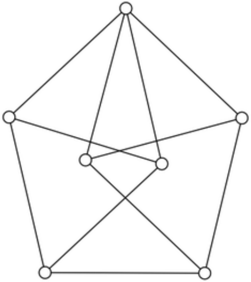

Figure: Moser's spindle

Unit distance graph

Lemma

Let $G = (\mathbb{R}^2, E)$ be the unit distance graph in plane. Then $\chi(G) \leq 7$.

Proof:

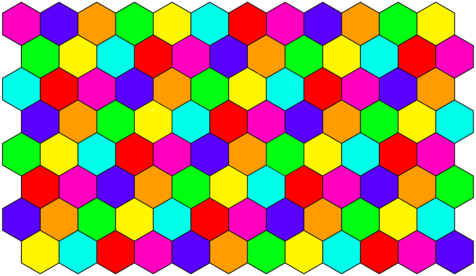

Figure: Hexagonal tiling

Finite Ramsey theorem

Definition (Arrow notation)

Suppose κ, λ, θ are nonzero cardinals. We write $\kappa \to (\lambda)_\theta$ (read "κ arrows λ with θ colors") iff for every function $f : [\kappa]^2 \to \theta$, there exists $X \subseteq \kappa$ such that $|X| = \lambda$ and $f \restriction [X]^2$ is constant.

Note that the infinite Ramsey theorem says that $\omega \to (\omega)_N$ for every $N < \omega$.

Theorem (Finite Ramsey theorem)

Given positive integers k and m, there exists an integer $N \geq m$ such that

$$N \to (m)_k$$

Proof: Towards a contradiction, fix $k, m \geq 1$ such that for every $N \geq m$,

$$N \nrightarrow (m)_k$$

Finite Ramsey theorem

For $n \geq m$ and $f : [n]^2 \to k$, we say that f is **bad** iff there is no $X \subseteq n$ such that $|X| = m$ and $f \upharpoonright [X]^2$ is constant. Let

$$T = \{f : (\exists n \geq m)(f : [n]^2 \to k \text{ is bad})\}$$

T must be infinite otherwise, letting

$$N = \max(\{n + 1 : (\exists f \in T)(\mathrm{dom}(f) = [n]^2)\})$$

we get $N \to (m)_k$ contradicting our assumption that there is no such N. Observe that if $n_2 \geq n_1 \geq m$ and $f : [n_2]^2 \to k$ is bad, then $f \upharpoonright [n_1]^2$ is also bad.

Finite Ramsey theorem

Inductively construct $\langle f_n : n \geq m \rangle$ as follows.

(1) Since T is infinite and since there are only finitely many functions in T from $[m]^2$ to k, we can choose $f_m : [m]^2 \to k$ such that $f_m \in T$ and $\{g \in T : g \text{ extends } f_m\}$ is infinite.

(2) Suppose $f_n : [n]^2 \to k$ has been defined such that $f_n \in T$ and $\{g \in T : g \text{ extends } f_n\}$ is infinite. Note that there are only finitely many functions from $[n+1]^2$ to k extending f_{n+1}. So we can choose $f_{n+1} : [n+1]^2 \to k$ in T such that $\{g \in T : g \text{ extends } f_{n+1}\}$ is infinite.

Put $f = \bigcup\{f_n : n \geq m\}$. Then $f : [\omega]^2 \to k$. By infinite Ramsey theorem, there exists an infinite $X \subseteq \omega$ so that $f \restriction [X]^2$ is constant. Choose $N \geq m$ large enough such that $|X \cap N| > m$. But since $f_N = f \restriction [N]^2$, it follows that $f_N : [N]^2 \to k$ is not bad: A contradiction. ☕

Finite Schur's theorem

Theorem (Schur)

For every $k \geq 1$, there exists N such that for every $h : N \to k$, there are $1 \leq x, y, z < N$ such that $x + y = z$ and $h(x) = h(y) = h(z)$.

Proof: Using finite Ramsey theorem, choose N large enough so that $N \to (3)_k$. Define $f : [N]^2 \to k$ by $f(\{m < n\}) = h(n - m)$. Choose $X \subseteq N$ and $r < k$ such that $|X| = 3$ and $\mathrm{range}(f \restriction [X]^2) = \{r\}$. Let $X = \{n_1 < n_2 < n_3\}$. Put $a = n_2 - n_1$, $b = n_3 - n_2$ and $c = n_3 - n_1$. Then $a + b = c$ and $h(a) = h(b) = h(c) = r$. ☕

Modular version of Fermat's last theorem

Fermat's last theorem says that for every $k \geq 3$, the equation $x^k + y^k = z^k$ has no solutions in positive integers x, y, z. A possible way of proving this could have been to show that for every $k \geq 3$, there are arbitrarily large primes p such that the modular equation $x^k + y^k \equiv z^k \pmod{p}$ has no nonzero solutions x, y, z [Why?]. Schur showed that this approach would not work.

Theorem (Schur, 1917)

For each $k \geq 1$, there exists N such that for every prime $p > N$, the modular equation $x^k + y^k \equiv z^k \pmod{p}$ has nonzero solutions x, y, z.

For the proof of Schur's theorem, we will make use of the following fact which says that the multiplicative group of $\mathbb{Z}/p\mathbb{Z}$ is cyclic.

Fact

For every prime p, there exists $1 \leq a \leq p-1$ such that in $\mathbb{Z}/p\mathbb{Z}$,

$$\{[a^j] : 1 \leq j \leq p-1\} = \{[r] : 1 \leq r \leq p-1\}$$

Modular version of Fermat's last theorem

Proof of Schur's theorem: Let $k \geq 1$. Using the previous theorem of Schur, fix $N > k$ such that for every $h : N \to k$, there are positive integers $x, y, z < N$ such that $x + y = z$ and $h(x) = h(y) = h(z)$. Let $p > N$ be prime. We'll show that $x^k + y^k \equiv z^k \pmod{p}$ has nonzero solution x, y, z.

Define $h : \{1, 2, \ldots, p-1\} \to k$ as follows. Fix $1 \leq a \leq p-1$ such that $\{[a^j] : 1 \leq j \leq p-1\}$ has every nonzero member of $\mathbb{Z}/p\mathbb{Z}$. For each $1 \leq j \leq p-1$, fix unique $1 \leq t_j \leq p-1$ such that $[a^{t_j}] = [j]$. Let q_j and r_j be unique integers such that $t_j = kq_j + r_j$ and $0 \leq r_j < k$. Define $h(j) = r_j$.

Choose $1 \leq u, v, w \leq p-1$ such that $u + v = w$ and $h(u) = h(v) = h(w) = r$. Let $x = a^{q_u}$, $y = a^{q_v}$ and $z = a^{q_w}$. Then

$$a^r x^k \equiv a^r a^{kq_u} \equiv a^{kq_u + r} \equiv a^{t_u} \equiv u \pmod{p}$$

Similarly, $a^r y^k \equiv v \pmod{p}$ and $a^r z^k \equiv w \pmod{p}$. Since $u + v = w$, we get $a^r(x^k + y^k) \equiv a^r z^k \pmod{p}$. Since $\gcd(a^r, p) = 1$, it follows that $x^k + y^k \equiv z^k \pmod{p}$. ☕

Dense linear orderings

$(L, \prec)$ is a **dense linear ordering** iff $(L, \prec)$ is a linear ordering such that for every $x \prec y$ in L, there exists $z \in L$ such that $x \prec z \prec y$. A linear ordering $(L, \prec)$ is **without end-points** iff L does not have a $\prec$-largest member and L does not have a $\prec$-least member.

Theorem (Cantor)

Suppose $(L_1, \prec_1)$ *and* $(L_2, \prec_2)$ *are dense linear orderings without end-points. Assume* $|L_1| = |L_2| = \omega$. *Then* $(L_1, \prec_1) \cong (L_2, \prec_2)$.

Proof: See video. ☕

Rado graph

Let $G = (V, E)$ be a graph. We say that G is **Rado** iff $|V| = \omega$ and for every pair A, B of disjoint finite subsets of V, there exists $x \in V$ such that x is adjacent to every vertex in A and x is not adjacent to any vertex in B.

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs. We say that $f : V_1 \to V_2$ is an isomorphism from G_1 to G_2 iff f is a bijection from V_1 to V_2 and for every $x, y \in V_1$, $\{x, y\} \in E_1 \iff \{f(x), f(y)\} \in E_2$.

Theorem (Rado)

There exists a Rado graph and any two Rado graphs are isomorphic.

Proof: Let $G = (V, E)$ be the the **Ackermann-Rado graph** defined previously. So $V = \omega$ and $\{k < n\} \in E$ iff the kth bit of the binary representation of n is 1. It is easy to check that G is Rado. The proof of the fact that any two Rado graphs are isomorphic is similar to the proof of Cantor's theorem from previous slide. ☕