MTH302: Set Theory and Mathematical Logic

LECTURE NOTES

Some history

(1874) Cantor showed that the set of real numbers is uncountable. As a corollary, he deduced the existence of transcendental numbers.

(1883) Cantor introduced ordinal and cardinal numbers, the well-ordering principle and the continuum hypothesis. But he could not prove the well-ordering principle or the continuum hypothesis.

(1879) Frege laid down the foundations of first order logic. His work remained obscure until Russell popularized it near the end of the 19th century.

(1900) Hilbert posed his 23 problems for the next century. Problems 1, 2 and 10 would motivate further research in mathematical logic.

Some history

(1901) Russell discovered "Russell's paradox" and, as a result, showed that Frege's formal system was inconsistent.

(1904) Zermelo introduced the axiom of choice and proved the well-ordering theorem.

(1920) Lowenheim (1915) and Skolem (1920) discovered that first order theories cannot capture the cardinality of their models. In particular, any first order axiomatization of set theory will have countable model (Skolem's paradox).

(1925) Von Neumann gave his definition of ordinals and proposed the axiom of foundation.

(1930) Gödel established the completeness theorem for first order logic.

Some history

(1931) Gödel proved the incompleteness of arithmetic. He also showed that there is no consistency proof of arithmetic that can be formalized in arithmetic (Hilbert's 2nd problem).

(1935) Tarski introduces his theory of truth in a first order structure.

(1936) Church, Gödel and Turing independently discover the class of computable functions. Church and Turing independently showed and that the decision problem for first order logic is undecidable.

(1938) Gödel proved that the continuum hypothesis cannot be proved in ZFC.

(1963) Cohen showed that the continuum hypothesis is independent of ZFC (Hilbert's first problem).

(1970) Matiyasevich showed that the solvability of Diophantine equations is undecidable (Hilbert's 10th problem).

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

- **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

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

- **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 $X_1, X_2, \ldots, X_n$ are sets, we define

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

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 will give a precise definition of "first-order property" later.

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 $\bigcap \mathcal{F} = \{v \in Z : (\forall X \in \mathcal{F})(v \in X)\}$ 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 empty set. 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 "first-order 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)

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

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)

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

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 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

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

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 $\text{dom}(R)$ to $\text{range}(R)$. 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 $x \in \text{dom}(F)$, there is a unique $y \in \text{range}(F)$ 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, $\text{dom}(F) = A$ and $\text{range}(F) \subseteq B$.

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 $\text{range}(F) = B$
- 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

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

Restrictions, Images/preimages, Finite/Infinite

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

- The **restriction of F to X**, denoted $F \restriction X$ is defined by: $\text{dom}(F \restriction X) = X$ and for every $x \in X$, $(F \restriction X)(x) = F(x)$.
- 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 say 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 the following.*

- **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* $a, b \in A$ *if* $a \neq b$, *then either* $a \prec b$ *or* $b \prec a$.

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 $\text{pred}(A, \prec, x) = \{y \in A : y \prec x\}$.

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* $(pred(X, \prec, x), \prec)$ *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 $f(z) \prec z$. $\square$

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 $x \prec y$. $\square$

Well-orderings

Corollary

Suppose $(X, \prec_1)$ *and* $(Y, \prec_2)$ *are well-ordering and* $f, g : X \to Y$ *are isomorphisms from* $(X, \prec_1)$ *to* $(Y, \prec_2)$*. Then* $f = g$*.*

Proof: Note that $g^{-1} \circ f$ is an isomorphism from $(X, \prec_1)$ to $(X, \prec_2)$. By the previous theorem, $g^{-1} \circ f$ is the identity function on X. It follows that $f = g$. □

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* $x \in X$, $(pred(X, \prec_1, x), \prec_1) \cong (Y, \prec_2)$.

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

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 $\mathrm{dom}(F) = \mathcal{E}$ 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: Done in lectures. □

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

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

Proof of (a):
(i) 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 $z \subseteq y$.
(ii) 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 $\in$. It follows that y is an ordinal.
(iii) $y = \text{pred}(x, \in, y)$: 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

Proof of (b): Fix an isomorphism $f : (x, \in) \to (y, \in)$. Towards a contradiction, suppose f is not the identity function on x and let $v \in x$ be $\in$-least such that $f(v) \neq v$. Note that for every $u \in v$, $u = f(u) \in f(v)$. So $v \subseteq f(v)$. Next, suppose $t \in f(v)$. Then $t \in y$ as y is transitive. Since f is surjective, $t = f(w)$ for some $w \in x$. Since f preserves $\in$ and $t = f(w) \in f(v)$, we must have $w \in v$. Since v is the least non fixed point of f, $t = f(w) = w$. So $t \in v$. It follows that $f(v) \subseteq v$. But now $f(v) = v$ which is a contradiction. So f is the identity on x and hence $y = \text{range}(f) = x$. □

Proof of (c): Clear, since $\in$ is an irreflexive relation on x. □

Ordinals

Proof of (d): By a previous theorem about well-orderings, exactly one of the following cases must occur.
$(x, \in) \cong (y, \in)$: In this case, By part (b), $x = y$.
For some $v \in x$, $(\text{pred}(x, \in, v), \in) \cong (y, \in)$: In this case, $v = \text{pred}(x, \in, v)$ (by part (a)) and hence by part (b), $v = y$. So $y \in x$.
For some $u \in y$, $(\text{pred}(y, \in, u), \in) \cong (x, \in)$: In this case, $u = \text{pred}(y, \in, u)$ (by part (a)) and hence by part (b), $u = x$. So $x \in y$. □

Ordinals

Proof of (e): Let $x \in C$. If $x \cap C = \emptyset$, then x is as required: If $y \in C$, then by part (d), either $y = x$ or $y \in x$ or $x \in y$, and $y \notin x$ because $x \cap C = \emptyset$. So assume $x \cap C \neq \emptyset$. Since $(x, \in)$ is a well-ordering, we can choose $\in$-least member $z \in x \cap C$. As x is transitive, it follows that $z \cap C = \emptyset$. So z is as required. □

Proof of (f): Parts (c),(d) imply that $(A, \in)$ is a linear ordering. That every nonempty subset of A has an $\in$-least member follows from part (e). So $(A, \in)$ is a well-ordering. If A is transitive, the definition of being an ordinal implies that 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) of the previous theorem. Let Y be the set of all $x \in X$ such that $(\text{pred}(X, \prec, x), \prec)$ 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 $(\text{pred}(X, \prec, x), \prec)$. Let $A = \text{range}(f)$. 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 $Y = \text{pred}(X, \prec, b)$. But $(\text{pred}(X, \prec, b), \prec)$ is isomorphic to the ordinal A. So $b \in Y$ which is a contradiction. $\square$

Order types, sup/min

Definition (Order type)

If $(X, \prec)$ *is a well ordering, let* **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:

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

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)* $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$

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. $\square$

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.

Restrictions of functions

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

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

Sequences indexed by ordinals, Countable/uncountable

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. □

Transfinite induction

Theorem

Suppose κ is an ordinal and $P \subseteq \kappa$. Assume $(\forall \alpha < \kappa)(\alpha \subseteq P \implies \alpha \in P)$. Then $P = \kappa$.

Proof: Suppose not and define $\alpha = \min(\kappa \setminus X)$. Note that $\beta \in P$ for every $\beta < \alpha$ and therefore $\alpha \subseteq P$. Our assumption implies that $\alpha \in P$. A contradiction. $\square$

Corollary

Suppose $P \subseteq \omega$ and $(\forall n < \omega)(n \subseteq P \implies n \in P)$. Then $P = \omega$.

Transfinite recursion

Theorem

Let $\mathbf{F} : \mathbf{V} \to \mathbf{V}$ *be a "function from sets to sets". Then for each ordinal* γ*, there is a unique function* h *such that*

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

Warning: There is no function defined on all sets otherwise its domain will be the set of all sets which does not exist. $\mathbf{F}$ is suppose to be a first order formula $\mathbf{F}(x, y)$ satisfying: For every x, there exists a unique y such that $\mathbf{F}(x, y)$ holds. Some examples are $\mathbf{F}(x) = \{x\}$, $\mathbf{F}(x) = \bigcup x$ etc.

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 $h(0) = \mathbf{F}(h \restriction 0) = \mathbf{F}(0)$. Having defined $h(\beta)$ for every $\beta < \alpha$, we feed $h \restriction \alpha = \langle h(\beta) : \beta < \alpha \rangle$ to $\mathbf{F}$ to get $h(\alpha)$.

Proof of transfinite recursion

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.

Next, we prove the existence of such h for each γ. **Towards a contradiction,** fix the least γ for which there is no function h satisfying clauses $1 + 2$. Note that $\gamma > 0$. Now for each $\eta < \gamma$, there exists a unique h_η such that $\mathrm{dom}(h_\eta) = \eta$ and for every $\alpha < \eta$, $h_\eta(\alpha) = \mathbf{F}(h_\eta \restriction \alpha)$. Fix such h_η for each $\eta < \gamma$.

Claim

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

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

Define $h = \bigcup\{h_\eta : \eta < \gamma\}$ if γ is a limit ordinal and $h = h_\beta \cup \{(\beta, \mathbf{F}(g))\}$ if $\gamma = \beta + 1$. Note that the claim implies that h is a function with domain γ. It is easy to check that h also satisfies clause 2. **A contradiction**. □

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 $\alpha < \gamma$,

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

We claim that there must be some ordinal γ such that $s_\star \in \text{range}(h_\gamma)$. Otherwise, applying replacement axiom to the formula $\phi(x, y)$ which says "y is the least ordinal such that $x \in \text{range}(h_{y+1})$" and the set $S = \{x \in X : (\exists \gamma)(x \in \text{range}(h_\gamma))\}$, we'll get the set $\{\gamma : (\exists x \in S)(\phi(x, \gamma))\}$ that contains all ordinals which is impossible. Let γ be least such that $s_\star \in \text{range}(h_\gamma)$. Then h_γ is a bijection from γ to $X \cup \{s_\star\}$. Hence X can be well-ordered. $\square$

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 $p, q \in P$, 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 $W_p = \{q \in P : q \preceq p\}$. 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 $q \in X$, $q \preceq p$.
- We say that p is a **lower bound** of X iff for every $q \in X$, $p \preceq q$.
- 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. $\square$

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 $\mathcal{F}$. $\square$

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 $\text{dom}(f) \subseteq A$, $\text{range}(f) \subseteq B$ 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 $\text{dom}(h) = A$ or $\text{range}(h) = B$. 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 $\text{dom}(h) \neq A$ and $\text{range}(h) \neq B$. Fix $x \in A \setminus \text{dom}(h)$ and $y \in B \setminus \text{range}(h)$. Define $h' = h \cup \{(x, y)\}$. Then $h' \in \mathcal{F}$. Hence h is not maximal in $\mathcal{F}$ which is a contradiction. $\square$

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. $\square$

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 A, 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 $\text{pred}(\kappa \times \kappa, \prec, (\alpha, \alpha))$ 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 $\text{type}(\kappa \times \kappa, \prec) = \kappa$. 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 lemma. □

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 $X = \bigcup\{X_\alpha : \alpha < \kappa\}$. Using the axiom of choice, fix a function h with domain κ such that for every $\alpha < \kappa$, $h(\alpha)$ is an 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 $g(x) = (\alpha, h(\alpha)(x))$. So $|X| \leq |\kappa \times \kappa| = \kappa$. □

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, $|\mathcal{P}(\kappa)| = \kappa^+$.

Finitary functions and closures

Definition

*We say that f is an n-***ary function** *on A iff $f : A^n \to A$. 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 $\mathrm{range}(f \restriction B^n) \subseteq B$.*

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

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 each $f \in \mathcal{F}$ and $D \subseteq A$, define $f \star D = \text{range}(f \restriction D^n)$ 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 also easy to see that $B \subseteq C \subseteq A$ and C is closed under every function in $\mathcal{F}$. $\square$

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 $a_1, a_2, \ldots, a_n \in \mathbb{Q}$,

$$(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0) \implies (a_1 = a_2 = \cdots = a_n = 0)$$

(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 every $0 \neq x \in \mathbb{R}$ can be uniquely written as $x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$ 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 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 $|H| = |\mathbb{R}| = \mathfrak{c}$. $\square$

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 $a_1, a_2, \ldots, a_n \in \mathbb{Q}$,

$$(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0) \implies (a_1 = a_2 = \cdots = a_n = 0)$$

(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 every $0 \neq x \in \mathbb{C}$ can be uniquely written as $x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$ 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 $x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$ where $a_1, a_2, \ldots, a_n \in \mathbb{Q}$ and $x_1, x_2, \ldots, x_n \in H_1$, then

$$f(x) = a_1 h(x_1) + a_2 h(x_2) + \cdots + a_n h(x_n)$$

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

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, 1914)

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. □

Propositional logic

The **language of propositional logic** consists of the following.

(1) A set $\mathcal{V}ar$ of propositional variables.

(2) Logical connectives: $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction), $\Longrightarrow$ (implication), $\Longleftrightarrow$ (equivalence).

(3) Parenthesis: (,).

The set of **propositional formulas**, denoted $\mathcal{PF}$, is defined to be the **smallest** set satisfying (a) and (b) below.

(a) Every propositional variable is in $\mathcal{PF}$.

(b) If ϕ and ψ are in $\mathcal{PF}$, then so are $(\neg\phi)$, $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \Longrightarrow \psi)$, $(\phi \Longleftrightarrow \psi)$.

Valuations

A **valuation** is a function $val : \mathcal{V}ar \to \{0, 1\}$. We interpret 0 as "False" and 1 as "True". Given a valuation $val : \mathcal{V}ar \to \{0, 1\}$, there is a unique $H : \mathcal{PF} \to \{0, 1\}$ satisfying the following.

(A) $H(p) = val(p)$ for every $p \in \mathcal{V}ar$.

(B) For every ϕ and ψ in $\mathcal{PF}$, (i)-(v) hold.

$$\text{(i) } H((\neg\phi)) = \begin{cases} 1 & \text{if } H(\phi) = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{(ii) } H((\phi \wedge \psi)) = \begin{cases} 1 & \text{if } H(\phi) = 1 \text{ and } H(\psi) = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{(iii) } H((\phi \vee \psi)) = \begin{cases} 0 & \text{if } H(\phi) = 0 \text{ and } H(\psi) = 0 \\ 1 & \text{otherwise} \end{cases}$$

Valuations and truth

$$\text{(iv) } H((\phi \implies \psi)) = \begin{cases} 0 & \text{if } H(\phi) = 1 \text{ and } H(\psi) = 0 \\ 1 & \text{otherwise} \end{cases}$$

$$\text{(v) } H((\phi \iff \psi)) = \begin{cases} 1 & \text{if } H(\phi) = 1 \text{ and } H(\psi) = 1 \\ 1 & \text{if } H(\phi) = 0 \text{ and } H(\psi) = 0 \\ 0 & \text{otherwise} \end{cases}$$

We will denote this unique extension by **val** (boldface val). Given a valuation $val : \mathcal{V}ar \to \{0, 1\}$, and a propositional formula ϕ, we say that ϕ is **true under the valuation** *val* iff $\mathbf{val}(\phi) = 1$. Otherwise, ϕ is **false under the valuation** *val*.

Exercise: Suppose $val : \mathcal{V}ar \to \{0, 1\}$, $val' : \mathcal{V}ar \to \{0, 1\}$, ϕ is a propositional formula and for every propositional variable p occurring in ϕ, we have $val(p) = val'(p)$. Then ϕ is true under *val* iff ϕ is true under val'.

Satisfiability and tautologies

Definition

- *A propositional formula ϕ is* **satisfiable** *iff there exists a valuation $val : \mathcal{V}ar \to \{0, 1\}$ under which ϕ is true.*
- *A subset $S \subseteq \mathcal{PF}$ is* **satisfiable** *iff there exists a valuation $val : \mathcal{V}ar \to \{0, 1\}$ under which every formula in S is true.*

Exercise: Show that $((p \vee q) \iff (\neg p))$ is satisfiable and $((p \wedge q) \iff ((\neg p) \vee (\neg q)))$ is not satisfiable.

Definition

A propositional formula ϕ is a **tautology** *iff for* **every** *valuation $val : \mathcal{V}ar \to \{0, 1\}$, ϕ is true under val.*

Exercise: Show that $((p \implies q) \iff ((\neg p) \vee q))$ is a tautology.

Compactness theorem

Theorem

Let S be a set of propositional formulas. Then S is satisfiable iff every finite subset of S is satisfiable.

Proof: If S is satisfiable, then clearly every finite subset of S is also satisfiable. So assume that every finite subset of S is satisfiable and we will show that S is also satisfiable. In this proof, we will assume that the set of propositional variables $\mathcal{V}ar$ is **countable**. The general case can be proved using Zorn's lemma (see Homework). Let $\mathcal{V}ar = \{p_0, p_1, p_2, \ldots, \}$ enumerate all the propositional variables.

Compactness theorem

Define h to be a **good partial valuation** iff the following hold.

(a) h is a function, $\mathrm{dom}(h) \subseteq \mathcal{V}ar$ and $\mathrm{range}(h) \subseteq \{0,1\}$.

(b) For every finite $F \subseteq S$, there exists a valuation $val : \mathcal{V}ar \to \{0,1\}$ such that $h \subseteq val$ and every formula in F is true under val.

Lemma

Let h be a good partial valuation and $p \in \mathcal{V}ar$. Let $h_0 = h \cup \{(p,0)\}$ and $h_1 = h \cup \{(p,1)\}$. Then one of h_0, h_1 is a good partial valuation.

Proof of Lemma: Suppose h_0 is not a good partial valuation. Then we can fix a finite $F \subseteq S$ such that there is no valuation extending h_0 under which every formula in F is true. Let G be an arbitrary finite subset of S. Put $K = F \cup G$. Since h is a good partial valuation, there is a valuation $val : \mathcal{V}ar \to \{0,1\}$ such that $h \subseteq val$ and every formula in K is true under val. Since there is no valuation extending h_0 under which every formula in F is true, it follows that $val(p) = 1$. Hence $h_1 \subseteq val$ and every formula in G is true under val. This shows that h_1 is a good partial valuation.

Compactness theorem

Using the previous lemma, inductively construct $\langle f_n : n < \omega \rangle$ such that

(1) $f_n : \{p_0, p_1, \ldots, p_n\} \to \{0, 1\}$ is a good partial valuation.

(2) For every $m < n$, $f_m \subseteq f_n$.

To construct f_0, use the Lemma with $h = \emptyset$ and $p = p_0$. Having defined f_n, to obtain f_{n+1}, use the previous Lemma with $h = f_n$ and $p = p_{n+1}$. Having completed the construction, define $val = \bigcup\{f_n : n < \omega\}$.

We claim that every formula in S is true under val and therefore S is satisfiable. To see this, let $\phi \in S$. Choose n large enough so that every propositional variable occurring in ϕ is among $\{p_0, p_1, \ldots, p_n\}$. Since f_n is a good partial valuation and $\{\phi\}$ is a finite subset of S, there exists a valuation $val' : \mathcal{V}ar \to \{0, 1\}$ such that $f_n \subseteq val'$ and ϕ is true under val'. Since val and val' agree on all the propositional variables occurring in ϕ, we get that ϕ is also true under val. $\square$

Some number-theoretic statements

Number theory studies the "structure" $(\omega, +, \cdot, 0, 1)$. Some examples of number-theoretic statements are as follows.

(a) Addition is commutative.

(b) There are no positive integers x, y, z such that $x^3 + y^3 = z^3$.

These statements can be expressed using $+, \cdot, 0, 1$ and some purely logical notions as follows.

(a) $(\forall x)(\forall y)(x + y = y + x)$.

(b) $\neg(\exists x)(\exists y)(\exists z)[(x \cdot y \cdot z \neq 0) \wedge ((x \cdot x \cdot x) + (y \cdot y \cdot y) = z \cdot z \cdot z)]$.

The symbols $\wedge$, $\neg$, $=$, $\exists$, x, y, z are what are called logical symbols and are **not confined to number theory**. On the other hand, $+, \cdot, 0, 1$ are extra-logical symbols and specific to the structure being studied. First order logic is the abstract study of these purely logical notions.

First order languages

A **first order language** $\mathcal{L} = (\mathcal{C}onst, \mathcal{R}el, \mathcal{F}unct)$ consists of the following.

(1) A (possibly empty) set $\mathcal{C}onst$ of **constant symbols**.

(2) A (possibly empty) set $\mathcal{R}el$ of **relation symbols**. Each relations symbol $R \in \mathcal{R}el$ has a a finite arity $n \geq 1$ and we say that R is an n-ary relation symbol.

(3) A (possibly empty) set of **function symbols**. Each function symbol $F \in \mathcal{F}unct$ has a finite arity $n \geq 1$ and we say that F is an n-ary function symbol.

First order structures

Let $\mathcal{L} = (\mathcal{C}onst, \mathcal{R}el, \mathcal{F}unct)$ be a first order language. An **$\mathcal{L}$-structure** $\mathcal{M}$ consists of the following.

(1) A nonempty set M called the **domain** of the structure $\mathcal{M}$.

(2) For each $c \in \mathcal{C}onst$, a member $c^{\mathcal{M}} \in M$ called the **interpretation** of c in $\mathcal{M}$.

(3) For each $R \in \mathcal{R}el$ of arity $n \geq 1$, a subset $R^{\mathcal{M}} \subseteq M^n$ called the **interpretation** of R in $\mathcal{M}$.

(4) For each function symbol $F \in \mathcal{F}unct$ of arity $n \geq 1$, a function $F^{\mathcal{M}} : M^n \to M$ called the **interpretation** of F in $\mathcal{M}$.

Example: Let $\mathcal{L} = (\mathcal{C}onst, \mathcal{R}el, \mathcal{F}unct)$ where $\mathcal{C}onst = \emptyset$, $\mathcal{R}el = \{R\}$ where R is a ternary relation symbol and $\mathcal{F}unct = \emptyset$. Define an $\mathcal{L}$-structure $\mathcal{M}$ as follows. The domain of $\mathcal{M}$ is the set of all real numbers $\mathbb{R}$ and $R^{\mathcal{M}} = \{(x, y, z) \in \mathbb{R}^3 : x + y = z\}$.

Logical symbols

Given a first order language $\mathcal{L}$, we are going to define what are called $\mathcal{L}$-terms and $\mathcal{L}$-formulas. For this we need to introduce the following **logical symbols**.

(1) **Variables**: Typically denoted by $x, y, z, v, v_0, v_1, \ldots$

(2) **Logical connectives**: $\neg$, $\wedge$, $\vee$, $\Longrightarrow$, $\Longleftrightarrow$

(3) **Quantifiers**: $\exists$ (Existential), $\forall$ (Universal)

(4) **Equality**: $=$

(5) **Parenthesis**: (,)

$\mathcal{L}$-terms

Let $\mathcal{L}$ be a first order language. The set of $\mathcal{L}$-**terms** is defined as follows.

(A) Every constant symbol of $\mathcal{L}$ is an $\mathcal{L}$-term.

(B) Every variable is an $\mathcal{L}$-term.

(C) If $t_1, t_2, \ldots, t_n$ are $\mathcal{L}$-terms and F is an n-ary function symbol of $\mathcal{L}$, then $F(t_1, t_2, \ldots, t_n)$ is an $\mathcal{L}$-term.

(D) Nothing else is an $\mathcal{L}$-term.

$\mathcal{L}$-formulas

Let $\mathcal{L}$ be a first order language. The set of **atomic $\mathcal{L}$-formulas** is defined as follows.

(A) If s, t are $\mathcal{L}$-terms, then $s = t$ is an atomic $\mathcal{L}$-formula.

(B) If $t_1, t_2, \ldots, t_n$ are $\mathcal{L}$-terms and R is an n-ary relation symbol of $\mathcal{L}$, then $R(t_1, t_2, \ldots, t_n)$ is an atomic $\mathcal{L}$-formula.

(C) Nothing else is an atomic $\mathcal{L}$-formula.

The set of $\mathcal{L}$-**formulas** is defined as follows.

(1) Every atomic $\mathcal{L}$-formula is an $\mathcal{L}$-formula.

(2) If ϕ, ψ are $\mathcal{L}$-formulas, then so are $(\neg\phi)$, $(\phi \wedge \psi)$, $(\phi \vee \psi)$, $(\phi \implies \psi)$, $(\phi \iff \psi)$.

(3) If ϕ is an $\mathcal{L}$-formula and x is any variable, then $(\forall x)(\phi)$ and $(\exists x)(\psi)$ are $\mathcal{L}$-formulas.

(4) Nothing else is an $\mathcal{L}$-formula.

Free variables and sentences

Suppose $\mathcal{L}$ is a first order language, ϕ is an $\mathcal{L}$-formula and x is a variable. We say that x is **free** in ϕ iff one of the following holds.

(1) ϕ is atomic and x occurs in ϕ.

(2) $\phi \equiv (\neg\psi)$ and x is free in ψ.

(3) $\phi \equiv (\psi_1 \wedge \psi_2)$ and either x is free in ψ_1 or x is free in ψ_2.

(4) $\phi \equiv (\psi_1 \vee \psi_2)$ and either x is free in ψ_1 or x is free in ψ_2.

(5) $\phi \equiv (\psi_1 \implies \psi_2)$ and either x is free in ψ_1 or x is free in ψ_2.

(6) $\phi \equiv (\psi_1 \iff \psi_2)$ and either x is free in ψ_1 or x is free in ψ_2.

(7) $\phi \equiv (\forall y)(\psi)$ and $y \neq x$ and x is free in ψ.

(8) $\phi \equiv (\exists y)(\psi)$ and $y \neq x$ and x is free in ψ.

An $\mathcal{L}$-formula ϕ is a **sentence** iff it has no free variables.

Valuations and terms

Suppose $\mathcal{L}$ is a first order language and $\mathcal{M}$ is an $\mathcal{L}$-structure. A **valuation in** $\mathcal{M}$ is a function from the set of variables to the universe M of $\mathcal{M}$. Given a valuation val : **Variables** $\to M$, we can extend val to a function **val** defined on the set of all $\mathcal{L}$-terms as follows.

(i) If x is a variable, then $\mathbf{val}(x) = val(x)$.

(ii) If c is a constant symbol of $\mathcal{L}$, then $\mathbf{val}(c) = c^{\mathcal{M}}$.

(iii) If F is an n-ary function symbol of $\mathcal{L}$ and $t_1, t_2, \ldots, t_n$ are $\mathcal{L}$-terms, then $\mathbf{val}(F(t_1, t_2, \ldots, t_n)) = F^{\mathcal{M}}(\mathbf{val}(t_1), \mathbf{val}(t_2), \ldots, \mathbf{val}(t_n))$

Exercise: Suppose t is an $\mathcal{L}$-term, val, val' are two valuations such that for every variable x that occurs in t, $val(x) = val'(x)$. Then $\mathbf{val}(t) = \mathbf{val}'(t)$.

Truth inside a structure

Suppose $\mathcal{L}$ is a first order language, $\mathcal{M}$ is an $\mathcal{L}$-structure, ϕ is an $\mathcal{L}$-formula and *val* is a valuation in M. Define $(\mathcal{M}, val) \models \phi$ (the symbol "$\models$" is read "**models**") by induction on the length of ϕ as follows.

(a) If t_1, t_2 are $\mathcal{L}$-terms, then $(\mathcal{M}, val) \models t_1 = t_2$ iff $\mathbf{val}(t_1) = \mathbf{val}(t_2)$.

(b) If R is an n-ary relation symbol of $\mathcal{L}$ and $t_1, \ldots, t_n$ are $\mathcal{L}$-terms, then $(\mathcal{M}, val) \models R(t_1, \ldots, t_n)$ iff $(\mathbf{val}(t_1), \ldots, \mathbf{val}(t_n)) \in R^{\mathcal{M}}$.

(c) $(\mathcal{M}, val) \models (\exists x)(\phi)$ iff there exists val' : **Variables** $\to M$ such that for every variable $y \neq x$, $val(y) = val'(y)$ and $(\mathcal{M}, val') \models \phi$.

(d) $(\mathcal{M}, val) \models (\forall x)(\phi)$ iff for every val' : **Variables** $\to M$ such that for every variable $y \neq x$, $val(y) = val'(y)$ and $(\mathcal{M}, val') \models \phi$.

(e) $(\mathcal{M}, val) \models (\neg\phi)$ iff it is not the case that $(\mathcal{M}, val) \models \phi$.

(f) $(\mathcal{M}, val) \models (\phi \wedge \psi)$ iff "$(\mathcal{M}, val) \models \phi$ and $(\mathcal{M}, val) \models \psi$".

(g) $(\mathcal{M}, val) \models (\phi \vee \psi)$ iff "$(\mathcal{M}, val) \models \phi$ or $(\mathcal{M}, val) \models \psi$".

(h) $(\mathcal{M}, val) \models (\phi \implies \psi)$ iff "$(\mathcal{M}, val) \models \phi$ implies $(\mathcal{M}, val) \models \psi$".

(i) $(\mathcal{M}, val) \models (\phi \iff \psi)$ iff "$(\mathcal{M}, val) \models \phi$ iff $(\mathcal{M}, val) \models \psi$".

Only free variables matter

Lemma

Suppose $\mathcal{L}$ is a first order language, $\mathcal{M}$ is an $\mathcal{L}$-structure and ϕ is an $\mathcal{L}$-formula. Suppose val : **Variables** $\to M$ *and* val' : **Variables** $\to M$ *are such that for every free variable x of ϕ, $val(x) = val'(x)$. Then $(\mathcal{M}, val) \models \phi$ iff $(\mathcal{M}, val') \models \phi$.*

Proof: By induction on the length of ϕ. The details are left to the reader. □

If all free variables of ϕ appear in the list $(x_1, x_2, \ldots, x_n)$ and $(a_1, a_2, \ldots, a_n)$ is a n-tuple in M, then we write

$$\mathcal{M} \models \phi(a_1/x_1, a_2/x_2, \ldots, a_n/x_n)$$

iff for some (equivalently, for every) valuation val, if $val(x_k) = a_k$ for every $k \leq n$, then $(\mathcal{M}, val) \models \phi$.

Definition

*Suppose $\mathcal{L}$ is a first order language, $\mathcal{M}$ is an $\mathcal{L}$-structure and ϕ is an $\mathcal{L}$-**sentence**. Define $\mathcal{M} \models \phi$ iff for some (equivalently, for every) valuation val in $\mathcal{M}$, $(\mathcal{M}, val) \models \phi$.*

Theories, models and logical validity

Let $\mathcal{L}$ be a first order language. An $\mathcal{L}$-**theory** is a set of $\mathcal{L}$-sentences. Suppose T is an $\mathcal{L}$-theory, ϕ is an $\mathcal{L}$-sentence and $\mathcal{M}$ is an $\mathcal{L}$-structure.

(1) We say that $\mathcal{M}$ **is a model of** T and write $\mathcal{M} \models T$ iff for every sentence $\phi \in T$, $\mathcal{M} \models \phi$.

(2) Define $T \models \phi$ iff for every model $\mathcal{M}$ of T, $\mathcal{M} \models \phi$.

(3) ϕ is **logically valid** (denoted by $\models \phi$) iff for every $\mathcal{L}$-structure $\mathcal{M}$, $\mathcal{M} \models \phi$.

The languages of arithmetic and set theory

The **first order language $\mathcal{L}_{PA}$ of arithmetic** has one constant symbol: 0, one unary function symbol: S, two binary function symbols: $+$ and $\cdot$ and no relations symbols. Its "standard interpretation" is the $\mathcal{L}_{PA}$-structure $\mathcal{M} = (\omega, S, +, \cdot, 0)$. Here the domain of $\mathcal{M}$ is ω and $0, S, +$ and $\cdot$ have the usual interpretation. We will follow the traditional practice of writing $x + y$ for the term $+(x, y)$ and $x \cdot y$ for the term $\cdot(x, y)$.

The **first order language $\mathcal{L}_{ST}$ of set theory** has a binary relation symbol: $\in$ and no constant or function symbols.

Subformulas

Suppose $\mathcal{L}$ is a first order language and ϕ, χ are $\mathcal{L}$-formula. We say that χ is a **subformula** of ϕ iff one of the following holds.

(1) ϕ is atomic and $\chi \equiv \phi$.

(2) $\phi \equiv (\neg\psi)$ and either $\chi \equiv \phi$ or χ is a subformula of ψ.

(3) $\phi \equiv (\psi_1 \wedge \psi_2)$ and either $\chi \equiv \phi$ or χ is a subformula of ψ_1 or χ is a subformula of ψ_2.

(4) $\phi \equiv (\psi_1 \vee \psi_2)$ and either $\chi \equiv \phi$ or χ is a subformula of ψ_1 or χ is a subformula of ψ_2.

(5) $\phi \equiv (\psi_1 \implies \psi_2)$ and either $\chi \equiv \phi$ or χ is a subformula of ψ_1 or χ is a subformula of ψ_2.

(6) $\phi \equiv (\psi_1 \iff \psi_2)$ and either $\chi \equiv \phi$ or χ is a subformula of ψ_1 or χ is a subformula of ψ_2.

(7) $\phi \equiv (\forall y)(\psi)$ and either $\chi \equiv \phi$ or χ is a subformula of ψ.

(8) $\phi \equiv (\exists y)(\psi)$ and either $\chi \equiv \phi$ or χ is a subformula of ψ.

Free and bound occurrences, scope of a quantifier

Suppose $\mathcal{L}$ is a first order language, x is a variable and ϕ is an $\mathcal{L}$-formula. We say that an occurrence of x in ϕ is **bound** iff there is a subformula of ϕ of the form $(\forall x)(\psi)$ or of the form $(\exists x)(\psi)$ that contains this occurrence of x. An occurrence of x in ϕ is **free** iff it is not bound.

Example: Let $\phi \equiv ((\forall x)(y = x + S(0)) \wedge (x = z \cdot z))$. Then the first two occurrences of x (marked red) in ϕ are bound and the third occurrence of x (marked blue) in ϕ is free.

Suppose $(Qx)(\psi)$ is a subformula of ϕ where $Q \in \{\forall, \exists\}$. The scope of this occurence of Qx in ϕ consists of all variables that appear in ψ.

Example: Let $\phi \equiv ((\forall x)(y = x + S(0)) \wedge (x = z \cdot z))$. Then the scope of the only occurrence of $\forall x$ in ϕ is $\{x, y\}$.

Term substitutions in term

Suppose $\mathcal{L}$ is a first order language, x is a variable and s, t are $\mathcal{L}$-terms. Define $s(t/x)$ to be the term obtained by replacing all occurrences of x in s with t.

Example: Let $s = (y \cdot (x + v) + x)$ and $t = z \cdot x$. Then $s(t/x)$ is the term $(y \cdot ((z \cdot x) + v) + (z \cdot x))$.

Lemma

*Suppose $\mathcal{L}$ is a first order language, x is a variable and s, t are $\mathcal{L}$-terms. Let $\mathcal{M}$ be an $\mathcal{L}$-structure and val : **Variables** $\to M$ be a valuation in $\mathcal{M}$. Put $a =$ **val**(t). Let val$'$: **Variables** $\to M$ be defined by: val$'(y) =$ val(y) for every $y \neq x$ and val$'(x) = a$. Then* **val**$(s(t/x)) =$ **val**$'(s)$.

Proof: By induction on the length of s. □

Term substitutions in formulas

Suppose $\mathcal{L}$ is a first order language, x is a variable and ϕ is an $\mathcal{L}$-formula. Let t be an $\mathcal{L}$-term. We say that t **is free for** x **in** ϕ iff no free occurrence of x is in the scope of a quantifier of the form $\exists y$ or $\forall y$ where y is a variable that occurs in t.

Example: Let $\phi \equiv ((\forall x)(y = x + S(0)) \wedge (y = z \cdot z))$. Then the term $(x + z)$ is free for z in ϕ but not free for y in ϕ.

If t is free for x in ϕ, by $\phi(t/x)$ we denote the formula obtained by replacing every free occurrence of x in ϕ by t. Note that all occurrences of variables in t remain free in $\phi(t/x)$.

Example: Let $\phi \equiv ((\forall x)(y = x + S(0)) \wedge (y = z \cdot z))$. Then the term $(x + z)$ is free for z in ϕ and

$$\phi((x + z)/z) \equiv ((\forall x)(y = x + S(0)) \wedge (y = (x + z) \cdot (x + z)))$$

Free substitution and truth

Lemma (Free substitution lemma)

*Suppose $\mathcal{L}$ is a first order language, x is a variable and ϕ is an $\mathcal{L}$-formula. Let t be an $\mathcal{L}$-term which is free for x in ϕ. Let $\mathcal{M}$ be an $\mathcal{L}$-structure and val : **Variables** $\to M$ be a valuation in $\mathcal{M}$. Put $a = \mathbf{val}(t)$. Let val' : **Variables** $\to M$ be defined by: $val'(y) = val(y)$ for every $y \neq x$ and $val'(x) = a$. Then $(\mathcal{M}, val) \models \phi(t/x)$ iff $(\mathcal{M}, val') \models \phi$.*

Proof: We argue by induction on the length of ϕ.

Case 1: ϕ is an atomic formula. First suppose ϕ is $s_1 = s_2$, where s_1 and s_2 are $\mathcal{L}$-terms. Then by the previous lemma on term substitution in terms, $\mathbf{val}(s_k(t/x)) = \mathbf{val}'(s_k)$ for $k = 1, 2$. Hence $\mathbf{val}(s_1(t/x)) = \mathbf{val}(s_2(t/x))$ iff $\mathbf{val}'(s_1) = \mathbf{val}'(s_2)$. Since $\phi(t/x)$ is $s_1(t/x) = s_2(t/x)$, the result follows. Next suppose, ϕ is $R(s_1, \ldots, s_n)$ where R is an n-ary relation symbol of $\mathcal{L}$ and $s_1, \ldots, s_n$ are $\mathcal{L}$-terms. By the previous lemma, $\mathbf{val}(s_k(t/x)) = \mathbf{val}'(s_k)$ for each $k \leq n$. It follows that $(\mathbf{val}(s_1(t/x)), \ldots, \mathbf{val}(s_n(t/x))) \in R^{\mathcal{M}}$ iff $(\mathbf{val}'(s_1), \ldots, \mathbf{val}'(s_n)) \in R^{\mathcal{M}}$. Since $\phi(t/x)$ is $R(s_1(t/x), \ldots, s_n(t/x))$, the result follows.

Free substitution and truth

Case 2: ϕ is of the form $(\phi_1 \wedge \phi_2)$. By the inductive hypothesis applied to the shorter formulas ϕ_1 and ϕ_2, we get $(\mathcal{M}, val) \models \phi_k(t/x)$ iff $(\mathcal{M}, val') \models \phi_k$ for $k = 1, 2$. Now $(\mathcal{M}, val) \models (\phi_1 \wedge \phi_2)$ iff $(\mathcal{M}, val) \models \phi_1$ and $(\mathcal{M}, val) \models \phi_2$. Therefore, $(\mathcal{M}, val) \models \phi(t/x)$ iff $(\mathcal{M}, val') \models \phi$.

The argument for the cases when ϕ is of the form $(\neg\psi)$, $(\phi_1 \vee \phi_2)$, $(\phi_1 \implies \phi_2)$ and $(\phi_1 \iff \phi_2)$ are similar to case 2 and left to the reader.

Case 3: ϕ is of the form $(\exists y)(\psi)$ or $(\forall y)(\psi)$.
First suppose that x is not free in ϕ. Then $\phi(t/x)$ is just ϕ. Now since val and val' agree at every variable except x and x is not free in ϕ, we must have $(\mathcal{M}, val) \models \phi$ iff $(\mathcal{M}, val') \models \phi$ and the lemma is clear.
So we can assume that x has at least one free occurrence in ϕ. It follows that $y \neq x$. It also follows that y cannot occur in the term t since t is free for x in ϕ.

Free substitution and truth

Subcase 3(a): ϕ is of the form $(\exists y)(\psi)$: By definition, $(\mathcal{M}, val) \models (\exists y)(\psi)(t/x)$ iff there exists a valuation val'' such that val and val'' agree on **Variables** $\setminus \{y\}$ and $(\mathcal{M}, val'') \models \psi(t/x)$. Fix such a valuation val''. Note that as y does not occur in t, we have $\mathbf{val}''(t) = \mathbf{val}(t) = a$. Define val''' to be the valuation such that val''' and val'' agree on **Variables** $\setminus \{x\}$ and $val'''(x) = a$. By the inductive hypothesis applied to the formula ψ, we get $(\mathcal{M}, val'') \models \psi(t/x)$ iff $(\mathcal{M}, val''') \models \psi$. So we get $(\mathcal{M}, val''') \models \psi$. Now val' and val''' agree on **Variables** $\setminus \{y\}$. So we get that $(\mathcal{M}, val') \models (\exists y)(\psi)$. Hence $(\mathcal{M}, val) \models \phi(t/x)$ implies $(\mathcal{M}, val') \models \phi$. The proof of the other implication is similar.

Subcase 3(b): ϕ is of the form $(\forall y)(\psi)$: Similar to the above. □

Universal closure, free substitution

Definition

A universal closure of ϕ is a formula of the form $(\forall x_1)(\forall x_2)\ldots(\forall x_n)(\phi)$ such that every free variable of ϕ is in $\{x_1, x_2, \ldots, x_n\}$.

For example, $(\forall x)(\forall y)(\forall z)((x + y) = (y + x))$ and $(\forall x)(\forall y)((x + y) = (y + x))$ are both universal closures of $(x + y) = (y + x)$.

The following is an immediate corollary of the free substitution lemma.

Corollary (Free substitution is logically valid)

If t is free for x in ϕ, then a universal closure of each one of $((\forall x)(\phi) \implies \phi(t/x))$ and $(\phi(t/x) \implies (\exists x)(\phi))$ is logically valid.

Propositional tautologies

Let $\mathcal{L}$ be a first order language. A **propositional tautology of** $\mathcal{L}$ is a formula obtained by replacing each propositional variable in a tautology by an $\mathcal{L}$-formula.

Example: Let $\mathcal{L} = \mathcal{L}_{PA} = (\{0\}, \emptyset, \{S, +, \cdot\})$. Since $(p \implies q) \iff (\neg p \vee q)$ is a tautology, by replacing p with $(\exists y)(x + x = y)$ and q with $\neg(x = y)$, we get that

$$((\exists y)(x + x = y) \implies \neg(x = y)) \iff (\neg(\exists y)(x + x = y) \vee \neg(x = y))$$

is propositional tautology of $\mathcal{L}$.

Exercise: Suppose $\mathcal{L}$ is a first order language and ϕ is a universal closure of a propositional tautology of $\mathcal{L}$. Show that for every $\mathcal{L}$-structure $\mathcal{M}$, we have $\mathcal{M} \models \phi$.

Logical axioms for first order logic

Let $\mathcal{L}$ be a first order language. A **logical axiom of** $\mathcal{L}$ is any sentence of $\mathcal{L}$ which is a universal closure of a formula of one of the types below. Here $x, y, z, x_1, x_2, \ldots$ denote arbitrary variables and ϕ denotes arbitrary $\mathcal{L}$-formula.

1. Propositional tautologies of $\mathcal{L}$ (See previous slide).
2. $\phi \implies (\forall x)(\phi)$, where x is not free in ϕ.
3. $(\forall x)(\phi \implies \psi) \implies ((\forall x)(\phi) \implies (\forall x)(\psi))$
4. $(\forall x)(\phi) \implies \phi(t/x)$ where t is any $\mathcal{L}$-term which is free for x in ϕ.
5. $\phi(t/x) \implies (\exists x)(\phi)$ where t is any $\mathcal{L}$-term which is free for x in ϕ.
6. $(\forall x)(\neg\phi) \iff \neg(\exists x)(\phi)$
7. $(x = x)$
8. $(x = y) \iff (y = x)$
9. $((x = y) \wedge (y = z)) \implies (x = z)$
10. $((x_1 = y_1) \wedge \cdots \wedge (x_n = y_n)) \implies (F(x_1, \ldots, x_n) = F(y_1, \ldots, y_n))$ where $n \geq 1$ and F is any n-ary function symbol of $\mathcal{L}$.
11. $((x_1 = y_1) \wedge \cdots \wedge (x_n = y_n)) \implies (R(x_1, \ldots, x_n) \iff R(y_1, \ldots, y_n))$ where $n \geq 1$ and R is any n-ary relation symbol of $\mathcal{L}$.

Logical axioms are logically valid

Theorem

Let $\mathcal{L}$ be a first order language. Every logical axiom of $\mathcal{L}$ is logically valid.

Proof: Let χ be a logical axiom of $\mathcal{L}$ and suppose $\mathcal{M}$ be an arbitrary $\mathcal{L}$-structure. We must show $\mathcal{M} \models \chi$. By definition, χ is a universal closure of one of the formulas listed in the 11 cases before. Cases 1, 7, 8, 9, 10 and 11 are easy consequences of the definition of $\mathcal{M} \models \chi$ and are left as an exercise for the reader. Cases 4 and 5 follow from the free substitution lemma. Let us do cases 2, 3 and 6.

Case 2. χ is of the form $(\forall x_1) \ldots (\forall x_n)(\phi \implies (\forall x)(\phi))$ where every free variable of ϕ is in $\{x_1, \ldots, x_n\}$. Let *val* : **Variables** $\to M$ be any valuation in $\mathcal{M}$. We must show $(\mathcal{M}, val) \models (\forall x_1) \ldots (\forall x_n)(\phi \implies (\forall x)(\phi))$. By Clause (d) in the definition of $\models$, this is equivalent to showing: For every valuation val' that agrees with *val* on **Variables** $\setminus \{x_1, \ldots, x_n\}$, we have $(\mathcal{M}, val') \models (\phi \implies (\forall x)(\phi))$. By Clause (h) in the definition of $\models$, this reduces to showing: If $(\mathcal{M}, val') \models \phi$, then $(\mathcal{M}, val') \models (\forall x)(\phi)$. So assume $(\mathcal{M}, val') \models \phi$. Since x is not free in ϕ, it follows that for every valuation val'' that agrees with val' on **Variables** $\setminus \{x\}$, we have $(\mathcal{M}, val'') \models \phi$. By Clause (d) in the definition of $\models$, it follows that $(\mathcal{M}, val') \models (\forall x)(\phi)$ and we are done.

Logical axioms are logically valid

Case 3. χ is of the form

$$(\forall x_1)\ldots(\forall x_n)[(\forall x)(\phi \implies \psi) \implies ((\forall x)(\phi) \implies (\forall x)(\psi))]$$

where all free variables of ϕ and ψ are in $\{x_1, \ldots, x_n\}$. Let val : **Variables** $\to M$ be any valuation in $\mathcal{M}$. We must show $(\mathcal{M}, val) \models \chi$. By Clause (d) in the definition of $\models$, this is equivalent to showing: For every valuation val' that agrees with val on **Variables** $\setminus \{x_1, \ldots, x_n\}$, we have $(\mathcal{M}, val') \models (\forall x)(\phi \implies \psi) \implies ((\forall x)(\phi) \implies (\forall x)(\psi))$. By Clause (h) in the definition of $\models$, this reduces to showing: If $(\mathcal{M}, val') \models (\forall x)(\phi \implies \psi)$ and $(\mathcal{M}, val') \models (\forall x)(\phi)$ then $(\mathcal{M}, val') \models (\forall x)(\psi)$. So assume $(\mathcal{M}, val') \models (\forall x)(\phi \implies \psi)$ and $(\mathcal{M}, val') \models (\forall x)(\phi)$.
Let val'' be any valuation that agrees with val' on **Variables** $\setminus \{x\}$. By Clause (d), it suffices to show $(\mathcal{M}, val'') \models \psi$. Since $(\mathcal{M}, val') \models (\forall x)(\phi \implies \psi)$, by Clause (d), we get $(\mathcal{M}, val'') \models \phi \implies \psi$ and $(\mathcal{M}, val'') \models \phi$. By Clause (h), we get $(\mathcal{M}, val'') \models \psi$ and we are done.

Logical axioms are logically valid

Case 6. χ is of the form $(\forall x_1)\ldots(\forall x_n)((\forall x)(\neg\phi) \iff \neg(\exists x)(\phi))$ where every free variable of ϕ is in $\{x_1,\ldots,x_n\}$. Let *val* : **Variables** $\to M$ be any valuation in $\mathcal{M}$. We must show $(\mathcal{M}, val) \models (\forall x_1)\ldots(\forall x_n)((\forall x)(\neg\phi) \iff \neg(\exists x)(\phi))$. By Clause (d) in the definition of $\models$, this is equivalent to showing: For every valuation val' that agrees with *val* on **Variables** $\setminus \{x_1,\ldots,x_n\}$, we have $(\mathcal{M}, val') \models ((\forall x)(\neg\phi) \iff \neg(\exists x)(\phi))$. By Clause (i), this reduces to showing $(\mathcal{M}, val') \models (\forall x)(\neg\phi)$ iff $(\mathcal{M}, val') \models \neg(\exists x)(\phi)$. By Clause (e) and (c), $(\mathcal{M}, val') \models \neg(\exists x)(\phi)$ iff there is no valuation val'' that agrees with val' on **Variables** $\setminus \{x\}$ such that $(\mathcal{M}, val'') \models \phi$. The latter statement is equivalent to: For every valuation val'' that agrees with val' on **Variables** $\setminus \{x\}$ we have $(\mathcal{M}, val'') \models \neg\phi$. This is equivalent to $(\mathcal{M}, val') \models (\forall x)(\neg\phi)$.

Proofs and theorems in a first order theory

Suppose $\mathcal{L}$ is a first order language, T is an $\mathcal{L}$-theory and ϕ is an $\mathcal{L}$-sentence. A **proof in** T is a finite sequence $\phi_1, \phi_2, \ldots, \phi_n$ of $\mathcal{L}$-sentences such that for each $i \leq n$,

(i) Either ϕ_i is a logical axiom or

(ii) $\phi_i \in T$ or

(iii) (**Modus Ponens**): There are $j, k < i$ such that ϕ_k is $(\phi_j \implies \phi_i)$.

Furthermore, if ϕ_n is ϕ, then we say that $\phi_1, \phi_2, \ldots, \phi_n$ is a **proof of** ϕ **in** T.

Definition

Let $\mathcal{L}$, T, ϕ be as above. We say that "T proves ϕ" or "ϕ is a theorem in T" and write $T \vdash_{\mathcal{L}} \phi$ iff there is a proof of ϕ in T. If the underlying language $\mathcal{L}$ is clear from the context, we drop it and just write $T \vdash \phi$.

Note that $T \vdash \phi$ iff for some finite $S \subseteq T$, $S \vdash \phi$. We sometimes write $\phi \vdash \psi$ instead of $\{\phi\} \vdash \psi$.

Soundness theorem

Theorem (Soundness theorem)

Suppose $\mathcal{L}$ is a first order language, T is an $\mathcal{L}$-theory and ϕ is an $\mathcal{L}$-sentence. If $T \vdash \phi$, then $T \models \phi$.

Proof: We need to show that every model of T is a model of ϕ. Assume that $\mathcal{M} \models T$ and we will show that $\mathcal{M} \models \phi$. Fix a proof $\phi_1, \ldots, \phi_n$ of ϕ in T. So ϕ_n is ϕ. By induction on $i \leq n$, we'll show that $\mathcal{M} \models \phi_i$.

Suppose $i = 1$. Then either $\phi_1 \in T$ or ϕ_1 is a logical axiom. If $\phi_1 \in T$, then $\mathcal{M} \models \phi_1$ since $\mathcal{M} \models T$. If ϕ_1 is a logical axiom, then by the previous theorem, ϕ_1 is logically valid and hence $\mathcal{M} \models \phi_1$.

Next suppose $i > 1$ and $\mathcal{M} \models \phi_j$ for every $j < i$. If either $\phi_i \in T$ or ϕ_i is a logical axiom, by repeating the argument in the previous paragraph, we get $\mathcal{M} \models \phi_i$. Finally, suppose there are $j, k < i$ such that ϕ_k is $\phi_j \implies \phi_i$. By the inductive hypothesis, we get $\mathcal{M} \models \phi_j$ and $\mathcal{M} \models (\phi_j \implies \phi_i)$. It easily follows that $\mathcal{M} \models \phi_i$. □

Deduction theorem

Theorem

Suppose T is an $\mathcal{L}$-theory and ϕ, ψ are $\mathcal{L}$-sentences. Then

$$T \vdash (\phi \implies \psi) \text{ iff } T \cup \{\phi\} \vdash \psi$$

Proof: First suppose $T \vdash (\phi \implies \psi)$. Let $\phi_1, \phi_2, \ldots, (\phi \implies \psi)$ be a proof of $(\phi \implies \psi)$ in T. By Modus Ponens, $\phi_1, \phi_2, \ldots, (\phi \implies \psi), \phi, \psi$ is a proof of ψ in $T \cup \{\phi\}$.

Next suppose $T \cup \{\phi\} \vdash \psi$. Let $\phi_1, \phi_2, \ldots, \phi_n$ be a proof of ψ in $T \cup \{\phi\}$. By induction on $i \leq n$, we'll show that $T \vdash (\phi \implies \phi_i)$. This suffices since ϕ_n is ψ. We have the following cases.

Case 1: ϕ_i is either a logical axiom or in $T \cup \{\phi\}$. Note that $\phi_i \implies (\phi \implies \phi_i)$ is a propositional tautology. It follows that if either ϕ_i is a logical axiom or $\phi_i \in T$, then $\phi_i, \phi_i \implies (\phi \implies \phi_i), \phi \implies \phi_i$ is a proof of $\phi \implies \phi_i$ in T. Also, if ϕ_i is ϕ, then $\phi \implies \phi$ is a propositional tautology. Hence in Case 1, $\phi \implies \phi_i$ has a proof in T. Note that Case 1 also covers the case $i = 1$ since ϕ_1 must be either a logical axiom or in $T \cup \{\phi\}$.

Deduction theorem

Case 2: For some $j, k < i$, ϕ_k is $(\phi_j \implies \phi_i)$. By the inductive hypothesis, $T \vdash (\phi \implies \phi_j)$ and $T \vdash (\phi \implies (\phi_j \implies \phi_i))$. Observe that

$$(\phi \implies \phi_j) \implies [(\phi \implies (\phi_j \implies \phi_i)) \implies (\phi \implies \phi_i)]$$

is a propositional tautology. So by applying Modus Ponens twice we get $T \vdash (\phi \implies \phi_i)$. □

Consistent theories and proofs by contradiction

A theory T in a first order language $\mathcal{L}$ is **inconsistent** iff there is an $\mathcal{L}$-sentence ϕ such that $T \vdash \phi$ and $T \vdash \neg\phi$. Otherwise T is **consistent**.

Exercise: Let T be a theory in a first order language $\mathcal{L}$. Suppose T has a model. Show that T is consistent.

Lemma

Let T be a theory in a first order language $\mathcal{L}$. Then T is inconsistent iff for every $\mathcal{L}$-sentence ϕ, $T \vdash \phi$.

Proof: The right to left implication is clear. Next assume T is inconsistent and fix an $\mathcal{L}$-sentence ψ such that $T \vdash \psi$ and $T \vdash \neg\psi$. Let ϕ be any $\mathcal{L}$-sentence. Note that $(\psi \implies (\neg\psi \implies \phi))$ is a propositional tautology. Hence $T \vdash (\psi \implies (\neg\psi \implies \phi))$. Applying Modus Ponens twice, we get $T \vdash \phi$. □

Corollary (Proof by contradiction)

$T \vdash \phi$ iff $T \cup \{\neg\phi\}$ is inconsistent.

Proof: If $T \vdash \phi$, then $T \cup \{\neg\phi\}$ proves both ϕ and $\neg\phi$ and so it is inconsistent. Next assume $T \cup \{\neg\phi\}$ is inconsistent. By the previous lemma, $T \cup \{\neg\phi\} \vdash \phi$. By the deduction theorem, $T \vdash (\neg\phi \implies \phi)$. Now $((\neg\phi \implies \phi) \implies \phi)$ is a propositional tautology. Hence by Modus Ponens, $T \vdash \phi$. □

Completeness theorems

The following theorem says that every consistent theory has a model. It was proved by Kurt Gödel in 1930.

Theorem (Completeness theorem I)

Let T be a consistent theory in a first order language $\mathcal{L}$. Then there is an $\mathcal{L}$-structure $\mathcal{M}$ such that $\mathcal{M} \models T$.

An immediate corollary is the following.

Corollary (Completeness theorem II)

Let T be a consistent theory in a first order language $\mathcal{L}$. Then for every $\mathcal{L}$-sentence ϕ, $T \vdash \phi$ iff $T \models \phi$.

Proof of Corollary: The left to right implication is the soundness theorem. For the converse, assume $T \nvdash \phi$. Then $T \cup \{\neg\phi\}$ is consistent. By the above theorem, there is an $\mathcal{L}$-structure $\mathcal{M}$ such that $\mathcal{M} \models T \cup \{\neg\phi\}$. It follows that $T \not\models \phi$. $\square$

Compactness theorem

Theorem (Compactness theorem)

Let T be a theory in a first order language $\mathcal{L}$. Suppose every finite subset of T has a model. Then T has a model.

Proof: By the completeness theorem, it suffices to show that T is consistent. But this is obvious since every finite subset of T is consistent. □

Isomorphism between first order structures

Suppose $\mathcal{L}$ is a first order language and $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures. Let $h : M \to N$ where M and N are the domains of $\mathcal{M}$ and $\mathcal{N}$ respectively. We say that h **is an isomorphism from** $\mathcal{M}$ **to** $\mathcal{N}$ iff h is a bijection and the following hold.

1. For every constant symbol c of $\mathcal{L}$, $h(c^{\mathcal{M}}) = c^{\mathcal{N}}$.
2. For every n-ary function symbol F of $\mathcal{L}$ and $a_1, \ldots, a_n, b$ in M, $F^{\mathcal{M}}(a_1, \ldots, a_n) = b$ iff $F^{\mathcal{N}}(h(a_1), \ldots, h(a_n)) = h(b)$.
3. For every n-ary relation symbol R of $\mathcal{L}$ and $a_1, \ldots, a_n$ in M, $(a_1, \ldots, a_n) \in R^{\mathcal{M}}$ iff $(h(a_1), \ldots, h(a_n)) \in R^{\mathcal{N}}$.

We say that $\mathcal{M}$ **is isomorphic to** $\mathcal{N}$ **and write** $\mathcal{M} \cong \mathcal{N}$ iff there is an isomorphism from $\mathcal{M}$ to $\mathcal{N}$.

Complete theory and Theory of a model

Suppose $\mathcal{L}$ is a first order language and $\mathcal{M}$ is an $\mathcal{L}$-structures. The **theory of $\mathcal{M}$, denoted** $Th(\mathcal{M})$ is the set of all $\mathcal{L}$-sentence ϕ such that $\mathcal{M} \models \phi$.

T is a **complete $\mathcal{L}$-theory** iff for every $\mathcal{L}$-sentence ϕ either $T \vdash \phi$ or $T \vdash \neg\phi$.

Exercise: Show that $Th(\mathcal{M})$ is a complete theory.

Lemma

Suppose $\mathcal{L}$ is a first order language and $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures. Let $h : M \to N$ be an isomorphism from $\mathcal{M}$ to $\mathcal{N}$. Let val be a valuation in $\mathcal{M}$ and ϕ be any $\mathcal{L}$-formula. Define $val' = h \circ val$ and note that val' is a valuation in $\mathcal{N}$. Then $(\mathcal{M}, val) \models \phi$ iff $(\mathcal{N}, val') \models \phi$

Proof By induction on length of ϕ. The details are left to the reader. □

The following is an easy corollary of the previous lemma.

Corollary

Suppose $\mathcal{L}$ is a first order language and $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures. Then $\mathcal{M} \cong \mathcal{N}$ implies $Th(\mathcal{M}) = Th(\mathcal{N})$.

An application of compactness

For each $n \geq 2$, let $\exists_{\geq n}$ denote the following sentence:

$$(\exists x_1)(\exists x_2)\ldots(\exists x_n)\left(\bigwedge_{i<j\leq n} \neg(x_i = x_j)\right)$$

For example, $\exists_{\geq 3}$ is $(\exists x_1)(\exists x_2)(\exists x_3)(\neg(x_1 = x_2) \wedge \neg(x_1 = x_3) \wedge \neg(x_2 = x_3))$.

If $\mathcal{M}$ is an $\mathcal{L}$-structure, then by **cardinality of** $\mathcal{M}$, we mean $|M|$ = the cardinality of the domain of $\mathcal{M}$.

Theorem

Let T be an $\mathcal{L}$-theory such that for every natural number n, T has a model of size $\geq n$. Then T has an infinite model.

Proof: Let $S = T \cup \{\exists_{\geq n} : n \geq 2\}$. Then every finite subset of S has a model. By compactness theorem, S has a model $\mathcal{M}$. Now since for every $n \geq 2$, $\mathcal{M} \models \exists_{\geq n}$ we must have $|M| \geq n$ where M is the domain of $\mathcal{M}$. It follows that M is infinite. □

Upward Löwenheim-Skolem theorem

Theorem (Upward Löwenheim-Skolem)

Let T be any $\mathcal{L}$-theory such that T has an infinite model. Then for every cardinal κ, T has a model of cardinality $\geq \kappa$.

Proof: Let $\mathcal{C} = \{c_\alpha : \alpha < \kappa\}$ be a set of new constant symbols. Let $\mathcal{L}'$ be the extension of $\mathcal{L}$ obtained by adding these constant symbols to $\mathcal{L}$. Consider the $\mathcal{L}'$ theory

$$S = T \cup \{c_\alpha \neq c_\beta : \alpha < \beta < \kappa\}$$

Note that every finite $F \subseteq S$ has a model: Just take any infinite model $\mathcal{M}$ of T and interpret the finitely many constant symbols of $\mathcal{C}$ that occur in the formulas in F as distinct members of M. The rest of the constant symbols in $\mathcal{C}$ can be interpreted arbitrarily.
By the compactness theorem, it follows that S has a model $\mathcal{N}$. Since $\mathcal{N} \models c_\alpha \neq c_\beta$, it follows that $c_\alpha^{\mathcal{N}}$'s are pairwise distinct members of N. Hence $|N| \geq \kappa$. □

Löwenheim-Skolem theorem

Theorem (Löwenheim-Skolem)

Let T be any $\mathcal{L}$-theory such that T has an infinite model. Let κ be any infinite cardinal such that $|T| \leq \kappa$. Then T has a model of cardinality κ.

We skip the proof that involves starting with a model $\mathcal{M} \models T$ with $|M| \geq \kappa$ and constructing an "elementary submodel" $\mathcal{N}$ of $\mathcal{M}$ with $|N| = \kappa$.

Peano Arithmetic

Recall that the language of arithmetic $\mathcal{L}_{PA} = (0, S, +, \cdot)$. Peano arithmetic (abbreviated PA) is the $\mathcal{L}_{PA}$-theory whose axioms 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. $(\forall x)(\forall y)(x \cdot S(y) = (x \cdot y) + x)$
7. Induction scheme: Suppose ϕ is an $\mathcal{L}_{PA}$-formula and x is a variable. Then any universal closure of the following is an axiom of PA:

$$[\phi(0/x) \wedge (\forall x)(\phi \implies \phi(S(x)/x))] \implies (\forall x)(\phi)$$

Models of PA

The **standard model of PA** is $(\omega, 0, S, +, \cdot)$ where ω is the set of natural numbers and 0, S, $+$ and $\cdot$ are interpreted in the usual way. The following is easily verified.

Theorem

$(\omega, 0, S, +, \cdot) \models PA$.

Some theorems of PA

We list some frequently used theorems in PA here.

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

True arithmetic

Definition (True arithmetic)

True arithmetic is $TA = Th(\omega, 0, S, +, \cdot)$.

Since $(\omega, 0, S, +, \cdot) \models$ PA, every theorem of PA is in TA. However, we will later see that there exist sentences $\phi \in TA$ such that $PA \nvdash \phi$. So PA is not a complete $\mathcal{L}_{PA}$-theory. By the Löwenheim-Skolem theorem, we get the following.

Theorem

For every infinite cardinal κ*, TA has a model of cardinality* κ.

But what about countable models of TA? Are all countable models of TA isomorphic to the standard model $(\omega, 0, S, +, \cdot)$? We will show that the answer is no.

A non-standard countable model of TA

For each $n < \omega$, let $S^n(0)$ be the closed $\mathcal{L}_{PA}$-term defined as follows: $S^0(0) \equiv 0$ and for each $n < \omega$, $S^{n+1}(0) \equiv S(S^n(0))$. Let $\mathcal{L}$ be the language obtained by adding a new constant symbol c to $\mathcal{L}_{PA}$. Define an $\mathcal{L}$-theory T as follows:

$$T = TA \cup \{c \neq S^n(0) : n < \omega\}$$

Note that every finite subset of TA has a model. Just take the standard model $(\omega, 0, +, \cdot)$ and interpret c to be a sufficiently large natural number. By the compactness theorem, T has a model as well. By the Löwenheim-Skolem theorem, T has a countable model $\mathcal{N} = (N, 0^{\mathcal{N}}, +^{\mathcal{N}}, \cdot^{\mathcal{N}}, c^{\mathcal{N}})$. Let $\mathcal{M} = (N, 0^{\mathcal{N}}, +^{\mathcal{N}}, \cdot^{\mathcal{N}})$. It is clear that $\mathcal{M}$ is a model of TA. It is easy to see that $\mathcal{M}$ is not isomorphic to the standard model $(\omega, 0, +, \cdot)$.

One can also show that TA has continuum many pairwise non-isomorphic countable models. See Homework.

Categoricity

Let T be an $\mathcal{L}$-theory and κ be a cardinal. We say that T **is** κ**-categorical** iff any two models of T of cardinality κ are isomorphic.

For example, TA is **not** ω-categorical.

Theorem

Let T be a consistent $\mathcal{L}$-theory where $\mathcal{L}$ is a countable language. Assume T has no finite models. Suppose for some infinite cardinal κ, T is κ-categorical. Then T is a complete $\mathcal{L}$-theory.

Proof: Towards a contradiction, suppose T is incomplete and fix an $\mathcal{L}$-sentence ϕ such that T does not prove either one of $\phi, \neg\phi$. Then $T_1 = T \cup \{\phi\}$ and $T_2 = T \cup \{\neg\phi\}$ are both consistent $\mathcal{L}$-theories. Since T has no finite models, every model of T_1 (resp. T_2) is infinite. As $\mathcal{L}$ is countable, by the Löwenheim-Skolem theorem, we can find $\mathcal{M}, \mathcal{N}$ such that $\mathcal{M} \models T_1$, $\mathcal{N} \models T_2$ and $|M| = |N| = \kappa$. Since T is κ-categorical, we must have $\mathcal{M} \cong \mathcal{N}$. But this is impossible since $Th(\mathcal{M}) \neq Th(\mathcal{N})$. $\square$

Dense linear orderings without end-points

Let $\mathcal{L}$ consist of just one binary relation symbol: $\prec$. Define DLO to be the $\mathcal{L}$-theory whose axioms are as follows.

1. $(\forall x)(\neg(x \prec x))$
2. $(\forall x)(\forall y)(\forall z)(((x \prec y) \wedge (y \prec z)) \implies (x \prec z))$
3. $(\forall x)(\forall y)((x = y) \vee (x \prec y) \vee (y \prec x))$
4. $(\forall x)(\forall y)(\exists z)((x \prec y) \implies ((x \prec z) \wedge (z \prec y)))$
5. $(\forall x)(\exists y)(x \prec y)$
6. $(\forall x)(\exists y)(y \prec x)$

Axioms 1, 2, 3 are saying that $\prec$ is a linear ordering. Axiom 4 says that $\prec$ is a dense linear ordering. Axiom 5 is saying that there is no $\prec$-largest element and axiom 6 is saying that there is no $\prec$-least element.

DLO is ω-categorical

Theorem

DLO is ω-categorical.

Proof: Let $(L_1, \prec_1)$ and $(L_2, \prec_2)$ be two models of DLO where $|L_1| = |L_2| = \omega$. Let $L_1 = \{a_0, a_1, \dots\}$ and $L_2 = \{b_0, b_1, \dots, \}$. Recursively, define $\langle f_n : n < \omega \rangle$ such that the following hold.

1. Each f_n is a finite function, $\text{dom}(f_n) \subseteq L_1$ and $\text{range}(f_n) \subseteq L_2$.
2. $f_0 = \{(a_0, b_0)\}$ and for every $m < n < \omega$, $f_m \subseteq f_n$.
3. For every $n < \omega$, $a_n \in \text{dom}(f_{2n})$ and $b_n \in \text{range}(f_{2n+1})$.
4. For every $a, a' \in \text{dom}(f_n)$, $a \prec_1 a'$ iff $f_n(a) \prec_2 f_n(a')$.

Note that Clause 3 can be satisfied using the fact that L_1 and L_2 are dense linear orders without endpoints (See video). Having constructed $\langle f_n : n < \omega \rangle$, define $f = \bigcup\{f_n : n < \omega\}$. By Clause 3, $\text{dom}(f) = L_1$ and $\text{range}(f) = L_2$. It follows that f is an isomorphism from $(L_1, \prec_1)$ to $(L_2, \prec_2)$. □

Corollary

DLO is a complete theory.

Proof: Every model of DLO is infinite and DLO is ω-categorical. Hence DLO is complete. □

Torsion free divisible abelian groups

Let $\mathcal{L} = \{0, +\}$ where 0 is a constant symbol and $+$ is a binary function symbol. For every $n < \omega$, let nx be the $\mathcal{L}$-term defined as follows: $0x \equiv 0$ and $(n+1)x \equiv (nx + x)$. Define TFDAG to be the $\mathcal{L}$-theory whose axioms are as follows.

1. $(\forall x)(\forall y)(\forall z)((x + y) + z = x + (y + z))$.
2. $(\forall x)(x + 0 = 0 + x = x)$
3. $(\forall x)(\exists y)((x + y = y + x = 0))$
4. $(\forall x)(\forall y)(x + y = y + x)$
5. $(\forall x)(x \neq 0 \implies nx \neq 0)$ for each $n \geq 1$.
6. $(\forall x)(\exists y)(ny = x)$ for each $n \geq 1$.

Axioms 1, 2 and 3 are axioms for group theory. Axiom 4 says that the group is abelian/commutative. Axiom scheme 5 says that the group is torsion free. Axiom scheme 6 says that the group is divisible.

Torsion free divisible abelian groups

Let $(V, 0, +)$ be a model of TFDAG. The following are easy to check.

1. For each $x \in V$, there is a unique $y \in V$ such that $x + y = 0$. We denote this unique member by $-x$.
2. For each $x \in V$ and $n \geq 1$, there is a unique $y \in V$ such that $ny = x$. We denote this unique member by x/n.

For each $n \geq 1$ and $m \geq 0$, define $(m/n)(x) = m(x/n)$ and $(-m/n)(x) = -((m/n)(x))$. Then $(V, 0, +)$ is a vector space over the field $\mathbb{Q}$ with the scalar product defined above. Let $dim(V)$ denote the cardinality of any basis of V over $\mathbb{Q}$.

Exercise: If V is uncountable, then $dim(V) = |V|$.

Recall that any two vector spaces over the same field are isomorphic iff they have the same dimension. Therefore we get the following.

Theorem

TFDAG is κ-categorical for every uncountable cardinal κ.

Alphabets and Strings/Words

1. An **alphabet** Σ is a set of symbols.
2. A **string/word** over Σ is a finite sequence of symbols from Σ.
3. The **empty string** is denoted by $\langle\rangle$.
4. $\Sigma^\star$ is the set of all strings over Σ.

Suppose σ, τ are strings over an alphabet Σ. The **concatenation of σ and τ**, denoted $\sigma^\frown\tau$, is the string obtained by writing τ after σ.

Example: Let $\Sigma = \{0, 1\}$. Then

$$\Sigma^\star = \{\langle\rangle, 0, 1, 00, 01, 11, 10, 000, 001, 010, 011, 100, \dots\}$$

Let $\sigma = 001$ and $\tau = 10$. Then $\sigma^\frown\tau = 00110$.

Language over an Alphabet

Definition (Language)

Let Σ be an alphabet. We say that L is a **language over** Σ *iff* $L \subseteq \Sigma^\star$.

Example: Let Σ be an alphabet. Define

$$Palindrome(\Sigma) = \{\sigma \in \Sigma^\star : \sigma = r(\sigma)\}$$

where $r(\sigma)$ is the string obtained by reversing the symbols in σ. For example, $r(abc) = cba$.

Decidable Languages: Informal Definition

Suppose Σ is a **finite** alphabet and L is a language over Σ. We say that L is **decidable** iff there is a computer program P that on input $\sigma \in \Sigma^\star$ does the following.

1. If $\sigma \in L$, the program P **halts** and outputs 1.
2. If $\sigma \notin L$, the program P **halts** and outputs 0.

If there is no such program P, we say that L is **undecidable**.

Example: Let L be the set of palindromes over $\Sigma = \{0, 1\}$. Then L is decidable.

This definition of a computer program can be made precise via **Turing machines**. But we will not do it here.

Computable Functions: Informal Definition

Suppose Σ and Π are **finite** alphabets and $F : \Sigma^\star \to \Pi^\star$. We say that F is **computable** iff there is a computer program P that on each input $\sigma \in \Sigma^\star$, halts and outputs $F(\sigma)$.

Example: Let $\Sigma = \{0, 1, 2, \ldots, 9\}$, $\Pi = \{0, 1\}$ and $F : \Sigma^\star \to \Pi^\star$ be defined by $F(\sigma)$ is the binary representation of σ. For example, $F(13) = 1101$ and $F(008) = 1000$. Then F is computable.

Definition (Computable function on natural numbers)

We say that $f : \omega^n \to \omega$ is computable iff there is a computer program P that on each input $(x_1, \ldots, x_n) \in \omega^n$, halts and outputs $f(x_1, \ldots, x_n)$.

Primitive recursive functions

Recall that a finitary function on ω is an n-ary function $f : \omega^n \to \omega$ for some $1 \leq n < \omega$. The set of **primitive recursive functions**, denoted PRec, is defined to be the smallest set of finitary functions on ω satisfying the following.

1. (Identically Zero) Every $f : \omega^n \to \omega$ defined by $f \equiv 0$ is in PRec.
2. (Projections) For each $1 \leq k \leq n$, the function $f : \omega^n \to \omega$ defined by $f(x_1, \ldots, x_n) = x_k$ is in PRec.
3. (Successor function) $f : \omega \to \omega$ defined by $f(x) = x + 1$ is in PRec.
4. (Compositions) If $f : \omega^n \to \omega$ is in PRec and for each $1 \leq k \leq n$, $g_k : \omega^m \to \omega$ is in PRec, then $h : \omega^m \to \omega$ is in PRec where h is defined by

$$h(x_1, \ldots, x_m) = f(g_1(x_1, \ldots, x_m), g_2(x_1, \ldots, x_m), \ldots, g_n(x_1, \ldots, x_m))$$

5. (Recursion) If $g : \omega^{n+1} \to \omega$ and $h : \omega^{n-1} \to \omega$ are both in PRec, then $f : \omega^n \to \omega$ is in PRec where

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} h(x_2, \ldots, x_n) & \text{if } x_1 = 0 \\ g(f(x_1 - 1, x_2, \ldots, x_n), x_1 \ldots, x_n) & \text{if } x_1 \geq 1 \end{cases}$$

Primitive recursive functions

Most elementary functions that arise in arithmetic are primitive recursive. For example, addition, multiplication, exponentiation, factorial, the function $f(n) = n^{th}$ prime etc.

Observe that the set of primitive recursive functions is countable so most functions $f : \omega \to \omega$ are not primitive recursive.

It is not difficult to convince oneself that **every primitive recursive function is computable in the sense that one can write a computer program that computes it.**

Diagonalization

Is every computable function also primitive recursive? The answer is no. Let us see why.

To every primitive recursive function f, one can associate a "certificate" C which shows how f was built from the basic functions (identically zero, projections and successor) using a finite number of applications of compositions and recursion. We can enumerate all of these certificates in a computable way as $C_1, C_2, C_3, \ldots$.

Now define a function $f : \omega \to \omega$ as follows. If C_x is a certificate of a unary primitive recursive function $g : \omega \to \omega$, then $f(x) = g(x) + 1$. Otherwise, $f(x) = 0$. It should be intuitively clear that f is computable in the sense that one could write a computer program to compute it. We claim that f is not primitive recursive. Suppose it is. Then f has a certificate C. Since every certificate appears in the list $C_1, C_2, \ldots$, we can find an x such that $C = C_x$. Now by definition, $f(x) = f(x) + 1$: A contradiction. So f is not primitive recursive.

Unbounded search and halting

```
#include<stdio.h>
/*Finding rational square roots.*/
int main(){
    int n1, n2, x, s = 0;
    printf("Numerator:"); scanf("%d", &n1);
    printf("Denominator:"); scanf("%d", &n2);
    if(n1 > n2){x = n2; n2 = n1; n1 = x; s = 1;}
    int t = 0, a, b;
    for(a = 1; t==0; a++){
      for(b = 1; b <=a; b++){
        if(b*b*n2 == n1*a*a){
            if(s==0){
            printf("Success:");
            printf("(%d/%d)^2=%d/%d", b, a, n1, n2);
            t = 1;break;}
            else{printf("(%d/%d)^2=%d/%d", a, b, n2, n1);
            t = 1;break;}}
        else{printf("Fail:%d,%d\n", b, a);}
} } }
```

Partial computable functions: Informal Definition

We say that f is a **partial function from A to B** iff f is a function, $\mathrm{dom}(f) \subseteq A$ and $\mathrm{range}(f) \subseteq B$. A **partial finitary function on ω** is a partial function from ω^n to ω for some $n \geq 1$.

Definition (Partial computable function on natural numbers)

We say that f is an n-ary partial computable function iff it is a partial function from ω^n to ω and there is a computer program P that on each input $(x_1, \ldots, x_n) \in \omega^n$ does the following.

1. *If $(x_1, \ldots, x_n) \in dom(f)$, then P halts and outputs $f(x_1, \ldots, x_n)$.*
2. *If $(x_1, \ldots, x_n) \notin dom(F)$, then P does not halt.*

Definition (Total computable functions on natural numbers)

If f is an n-ary partial computable function on ω and $dom(f) = \omega^n$, then we say that f is a total computable function. It is customary to drop the "total" and just write computable function. So "computable function" will mean a "total computable function".

General recursive functions

The set of **general recursive functions**, denoted GRec, is defined to be the smallest set of partial finitary functions on ω that satisfies the following.

1. Every primitive recursive function is in GRec.
2. (Compositions) If f is an n-ary function in GRec and for each $1 \leq k \leq n$, g_k is an m-ary function in GRec, then h is in GRec where h is defined by

$$h(x_1, \ldots, x_m) = f(g_1(x_1, \ldots, x_m), g_2(x_1, \ldots, x_m), \ldots, g_n(x_1, \ldots, x_m))$$

3. (Primitive recursion) If $g, h \in GRec$ where g is $(n+1)$-ary and h is $(n-1)$-ary, then f is in GRec where

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} h(x_2, \ldots, x_n) & \text{if } x_1 = 0 \\ g(f(x_1 - 1, x_2, \ldots, x_n), x_1 \ldots, x_n) & \text{if } x_1 \geq 1 \end{cases}$$

4. (Unbounded search) If $g \in GRec$ is an $(n+1)$-ary, then $f \in$ GRec where f is an n-ary partial function on ω defined by: $f(x_1, \ldots, x_n) = z$ iff $g(z, x_1, \ldots, x_n) = 0$ and for every $y < z$, $g(y, x_1, \ldots, x_n)$ is defined and is nonzero.

Is every partial computable function general recursive?

We saw that there are total computable functions on ω which are not primitive recursive. The proof of this used diagonalization to produce a computable function which disagreed with every primitive recursive function on some input. Let us try to produce such a proof for the class of general recursive functions.

As before, we can associate to every general recursive function f, a certificate C which describes how f was built from the basic functions using a finite number of applications of compositions, primitive recursion and unbounded search. Let $C_1, C_2, \ldots$ be a computable listing of all such certificates. As before, define a partial unary function f on ω as follows: If C_x is the certificate of a unary general recursive function g, then $f(x) = 1 + g(x)$. Clearly, f is a partial computable function. Let us assume that f is general recursive and try to get a contradiction. Fix x such that C_x is a certificate of f. Now **if** $x \in \mathrm{dom}(f)$, **then** $f(x) = 1 + f(x)$ which is impossible. So the only thing we can conclude here is that $x \notin \mathrm{dom}(f)$ which is not a contradiction.

One could try to modify this argument by insisting that $C_1, C_2, \ldots$ be a list of certificates of only total computable functions. But it is not clear at all if we can list them in a computable way.

General recursive = Partial computable

Theorem

Every partial computable function is general recursive and every general recursive function is partial computable.

That every general recursive function is partial computable is not hard to see. One has to check that the basic functions are computable and the set of partial computable functions are closed under compositions, primitive recursion and unbounded search. The proof of the fact that every partial computable function is general recursive would require a precise analysis of the notion of a "computer program" and will not be covered in this course.

C.e. sets

Let $W \subseteq \omega$. We say that W is **c.e.** (computably enumerable) iff there is there is a unary partial computable function f on ω such that $W = \mathrm{dom}(f)$. We say that W is **computable** iff its characteristic function $1_W : \omega \to \omega$ is computable. Here, $1_W(x) = 1$ if $x \in W$ and $1_W(x) = 0$ if $x \notin W$.

Theorem

Let $W \subseteq \omega$. Then W is computable iff both W and $\omega \setminus W$ are c.e.

Proof: First suppose W is computable and fix a computer program P which computes 1_W. Define another program Q which does the following: On input x, Q runs the program P with input x. If P halts and outputs 1, then Q halts and outputs 1. If P halts and outputs 0, then Q enters an infinite loop. It is clear that Q computes a partial computable function whose domain is W. Hence W is c.e. A similar argument shows that $\omega \setminus W$ is also c.e.
Next suppose both W and $\omega \setminus W$ are c.e. Let f and g be unary partial computable functions on ω such that $W = \mathrm{dom}(f)$ and $\omega \setminus W = \mathrm{dom}(g)$. Fix programs P and Q such that P computes f and Q computes g. Define a program R as follows: On input x, R starts running both P and Q with the same input x. If P halts, then R halts and outputs 1. If Q halts, then R halts and outputs 0. Note that on each input x, exactly one of P, Q halts. It is clear that R computes 1_W. Hence W is computable. $\square$

Universal Turing Machines

Recall that every C-code is a string of keyboard characters. One can therefore write a C-program U which takes two inputs S and n where S is a finite string of keyboard characters and n is a natural number. The program U on input (S, n) first checks if S is a valid C-code (using a compiler). If S is not a C-code, U enters an infinite loop. Otherwise it runs the program S on input n and outputs whatever S does. This program U is an example of a universal computing machine in the sense that it can simulate every program. An easy corollary is the following.

Theorem

There exists a partial computable binary function ϕ on ω such that for every partial computable unary function h on ω, there exists an $e < \omega$ such that

1. *For every $k < \omega$, we have $k \in dom(h)$ iff $(e, k) \in dom(\phi)$.*
2. *For every $k \in dom(h)$, $\phi(e, k) = h(k)$.*

Halting problem is c.e. but not computable

Suppose ϕ is as in the previous theorem. For each $e < \omega$, we write φ_e to denote the unary partial computable function $k \mapsto \phi(e, k)$. So $\{\varphi_e : e < \omega\}$ is the set of all unary partial computable functions.

Define the **halting problem** $H \subseteq \omega$ by

$$H = \{e < \omega : e \in \mathrm{dom}(\varphi_e)\}$$

Theorem (Halting problem)

H is c.e. but not computable.

Proof: Let U be a program that computes ϕ. Consider the program P which on input e runs U with input (e, e) and halts whenever U does. Then P halts on input e iff U halts on input (e, e) iff $(e, e) \in \mathrm{dom}(\phi)$ iff $e \in \mathrm{dom}(\varphi_e)$ iff $e \in H$. Hence H is c.e.
Next, towards a contradiction, suppose H is computable. Then $\omega \setminus H$ is c.e. and hence there exists a partial computable function h such that $\mathrm{dom}(h) = \omega \setminus H$. Since $\{\varphi_e : e < \omega\}$ has every partial computable function, we can fix an $n < \omega$ such that $\varphi_n = h$. Now observe that $n \in H$ iff $n \in \mathrm{dom}(\varphi_n)$ iff $n \in \mathrm{dom}(h)$ iff $n \in \omega \setminus H$: A contradiction. So H cannot be computable. $\square$

C.e. languages

Suppose Σ is a finite alphabet and $L \subseteq \Sigma^\star$.

1. We say that L is **computable/recursive** iff L is decidable.
2. We say that L is **c.e** (computably enumerable) iff there is a computer program P which on each input $\sigma \in \Sigma^\star$, halts iff $\sigma \in L$.

The following can be proved just like the fact that $W \subseteq \omega$ is computable iff both W and $\omega \setminus W$ are c.e.

Theorem

Suppose Σ is a finite alphabet and $L \subseteq \Sigma^\star$. Then L is computable iff both L and $\Sigma^\star \setminus L$ are c.e.

First order logic with finite alphabet

Suppose $\mathcal{L}$ is a finite first order language. Let $\Sigma_{\mathcal{L}}$ be the alphabet which consists of the following.

1. All non-logical symbols of $\mathcal{L}$.
2. x and $'$.
3. $\neg, \vee, \wedge, \implies, \iff, (,), \forall, \exists, =$.

Define the set of variables to be $\{x, x', x'', x''', \dots\}$. $\mathcal{L}$-terms and $\mathcal{L}$-formulas are defined in the usual way. Recall that $L \subseteq \Sigma_{\mathcal{L}}^{\star}$ is decidable iff there is a computer program that decides whether a given input $\sigma \in \Sigma_{\mathcal{L}}^{\star}$ is in L. The following should be clear.

Theorem

The set of all $\mathcal{L}$-terms, the set of all $\mathcal{L}$-formulas and the set of all $\mathcal{L}$-sentences are all decidable subsets of $\Sigma_{\mathcal{L}}^{\star}$.

Logical validity is c.e.

Theorem

Suppose $\mathcal{L}$ is a finite first order language.

(1) *$\{\psi : \psi$ is a logical axiom of $\mathcal{L}\}$ is computable.*

(2) *Let T be a computable set of $\mathcal{L}$-sentences. Then $\{\psi : \psi$ is an $\mathcal{L}$-sentence and $T \vdash \psi\}$ is c.e.*

Proof: (1) Recall that an $\mathcal{L}$-sentence ψ is a logical axiom of $\mathcal{L}$ iff ψ is a universal closure of an $\mathcal{L}$-formula of one of the 11 types defined on slide 116. It should be clear that each one of these types of logical axiom is computable. Hence $\{\psi : \psi$ is a logical axiom of $\mathcal{L}\}$ is also computable.

(2) First observe that checking whether a finite string $\psi_1, \psi_2, \ldots, \psi_n$ is a proof in T is computable. This is because for each $i \leq n$, checking each one of "$\psi_i \in T$", "ψ_i is a logical axiom of $\mathcal{L}$" and "there are $j, k < i$ such that ψ_k is $(\psi_j \implies \psi_i)$" is computable. Now consider a computer program P which on input ψ starts listing all possible proofs in T and halts as soon as it finds a proof of ψ. Note that P halts on input ψ iff $T \vdash \psi$. So $\{\psi : \psi$ is an $\mathcal{L}$-sentence and $T \vdash \psi\}$ is c.e. $\square$

PA and ZFC

(A) Let $\mathcal{L}$ be a finite first order language. Then the set of all logically valid $\mathcal{L}$-sentences is c.e.

(B) Let $\mathcal{L}_{PA} = \{0, S, +, \cdot\}$ be the language of Peano arithmetic. It is easy to see the set of axioms of PA is computable. It follows the the set Thm_{PA} consisting of all $\mathcal{L}_{PA}$-sentences ψ which are theorems of PA is c.e.

(C) Let $\mathcal{L} = \{\in\}$ be the language of ZFC. It is easy to see the set of axioms of ZFC is computable. It follows the the set Thm_{ZFC} consisting of all $\mathcal{L}$-sentences ψ which are theorems of ZFC is c.e.

We'll later show that neither one of the sets Thm_{PA} and Thm_{ZFC} is computable.

Computable axiomatizations

Suppose $\mathcal{L}$ is a finite first order language and T is an $\mathcal{L}$-theory.

1. Let A be a set of $\mathcal{L}$-sentences. We say that A **axiomatizes** T iff for every $\mathcal{L}$-sentence ψ, $T \vdash \psi$ iff $A \vdash \psi$.
2. We say that T is **computably axiomatizable** iff there exists a computable set A of $\mathcal{L}$-sentences such that A axiomatizes T.
3. We say that T is **decidable** iff $\{\psi : \psi \text{ is an } \mathcal{L}\text{-sentence and } T \vdash \psi\}$ is computable.

Theorem

Suppose $\mathcal{L}$ is a finite first order language and T is a consistent complete $\mathcal{L}$-theory. Assume T is computably axiomatizable. Then T is decidable.

Proof: Let A be a computable set of $\mathcal{L}$-sentences that axiomatizes T. Since T is complete, for every $\mathcal{L}$-sentence ψ, either $A \vdash \psi$ or $A \vdash \neg\psi$. Put $W = \{\psi : \psi \text{ is an } \mathcal{L}\text{-sentence and } A \vdash \psi\}$. Then by the previous theorem, W is c.e. so we can fix a program P such that P halts on input ψ iff $\psi \in W$. Now consider the program Q which on input ψ runs P simultaneously on inputs ψ and $\neg\psi$. If P halts on input ψ, then Q outputs 1. If P halts on input $\neg\psi$, then Q outputs 0. It is clear that Q computes W hence W is computable. So T is decidable. $\square$

Definability in $(\omega, 0, S, +, \cdot)$

Notation: In what follows, whenever we write "$\psi(x_1, \ldots, x_k)$ is a formula" we mean "ψ is a formula whose free variables are among $x_1, \ldots, x_k$". Furthermore, we will write $\psi(a_1, \ldots, a_k)$ to denote $\psi(a_1/x_1, \ldots, a_k/x_k)$.

Let $\mathcal{N} = (\omega, 0, S, +, \cdot)$. Suppose $k \geq 1$ and $f : \omega^k \to \omega$. We say that f is **definable in** $\mathcal{N}$ iff there is an $\mathcal{L}_{PA}$-formula $\psi(y, x_1, x_2, \ldots, x_k)$ such that the following holds: For every $(m, n_1, n_2, \ldots, n_k) \in \omega^{k+1}$, we have

$$f(n_1, n_2, \ldots, n_k) = m \text{ iff } \mathcal{N} \models \psi(m, n_1, \ldots, n_k)$$

Suppose $X \subseteq \omega^k$. We say that X is **definable in** $\mathcal{N}$ iff 1_X is definable in $\mathcal{N}$ where $1_X : \omega^k \to \omega$ is the characteristic function of X.

Exercise: Suppose $X \subseteq \omega^k$. Then X is definable in $\mathcal{N}$ iff there exists a $\mathcal{L}_{PA}$-formula $\psi(x_1, \ldots, x_k)$ such that

$$X = \{(n_1, \ldots, n_k) \in \omega^k : \mathcal{N} \models \psi(n_1, \ldots, n_k)\}$$

Definability of computable functions

Our next goal is to show that every **total computable function is definable in** $\mathcal{N}$.

Theorem

Every total computable function is definable in $\mathcal{N} = (\omega, 0, S, +, \cdot)$.

Proof: It suffices to show that every total general recursive function is definable in $\mathcal{N}$. The proof will be broken down into a series of lemmas. The first lemma says that the basic functions are definable in $\mathcal{N}$.

Lemma

Let $1 \leq j \leq k$. *Then the following are definable in* $\mathcal{N}$.

(1) $f : \omega^k \to \omega$ *and* $f(x_1, \ldots, x_k) \equiv 0$.

(2) $f : \omega^k \to \omega$ *and* $f(x_1, \ldots, x_k) = x_j$.

(3) $f : \omega \to \omega$ *and* $f(x) = x + 1$.

Proof: (1) Take ψ to be the formula $y = 0$.
(2) Take ψ to be the formula $y = x_j$.
(3) Take ψ to be the formula $y = S(x)$. □

Definability of computable functions

The next lemma says that definable functions are closed under composition.

Lemma

Suppose $f : \omega^n \to \omega$ *and for each* $1 \leq k \leq n$, $g_k : \omega^m \to \omega$. *Assume that each one of* $f, g_1, g_2, \ldots, g_n$ *is definable in* $\mathcal{N}$. *Then* $h : \omega^m \to \omega$ *is definable in* $\mathcal{N}$ *where*

$$h(x_1, \ldots, x_m) = f(g_1(x_1, \ldots, x_m), \ldots, g_n(x_1, \ldots, x_m))$$

Proof: Let $\psi(y, x_1, \ldots, x_n)$ be the formula witnessing that f is definable in $\mathcal{N}$. For each $1 \leq k \leq n$, let $\psi_k(y, x_1, x_2, \ldots, x_m)$ be a formula witnessing that g_k is definable in $\mathcal{N}$. Let $\eta(y, x_1, \ldots, x_m)$ be the formula

$$(\exists z_1, z_2, \ldots, z_n)(\bigwedge_{k \leq n} \psi_k(z_k, x_1, \ldots, x_m) \wedge \psi(y, z_1, \ldots, z_n))$$

Then η witnesses that h is definable in $\mathcal{N}$. $\square$

Definability of computable functions

For $n < \omega$ and $d \geq 1$, define $rem(n, d)$ to be the remainder when n is divided by d. For example, $rem(7, 1) = 0$ and $rem(14, 5) = 4$. Note that $0 \leq rem(n, d) < d$.

Definition (β-function)

Define $\beta(x_1, x_2, x_3) = rem(x_1, x_2 x_3 + x_2 + 1)$.

It is easy to see that $rem(x, y)$ and $\beta(x_1, x_2, x_3)$ are definable in $\mathcal{N}$.

Lemma (β-function lemma)

For any finite sequence of natural numbers $\langle r_0, r_1, \ldots, r_n \rangle$*, there exist natural numbers* a *and* b *such that for every* $0 \leq i \leq n$*,*

$$\beta(b, a, i) = r_i$$

Definability of computable functions

Proof of the β-function lemma: Let $\langle r_0, r_1, \ldots, r_n \rangle$ be a finite sequence of natural numbers. Choose $m > n$ large enough such that $a = m!$ is greater than each r_i.

We claim that the integers $1 + a(i+1)$ are pairwise relatively prime for $0 \leq i \leq n$ (this means that the GCD of any two of them is 1). Suppose not and fix a prime p and $0 \leq i < j \leq n$ such that p divides both $1 + a(i+1)$ and $1 + a(j+1)$. Then p divides their difference $a(j-i)$. Since p is prime, either p divides $a = m!$ or p divides $j - i$. In either case, $p \leq m$. So p divides a which means that p does not divide $1 + a(i+1)$: A contradiction.

By the Chinese remainder theorem (see homework), there exists a natural number b such that for each $0 \leq i \leq n$, $rem(b, 1 + a(i+1)) = r_i$. Hence a and b are as required. □

Definability of computable functions

Using the β-function lemma, we can now show that the set of definable functions in $\mathcal{N}$ are closed under primitive recursion.

Theorem

Suppose $g : \omega^{n+1} \to \omega$ and $h : \omega^{n-1} \to \omega$ are both definable in $\mathcal{N}$. Then $f : \omega^n \to \omega$ is definable in $\mathcal{N}$ where

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} h(x_2, \ldots, x_n) & \text{if } x_1 = 0 \\ g(f(x_1 - 1, x_2, \ldots, x_n), x_1 \ldots, x_n) & \text{if } x_1 \geq 1 \end{cases}$$

Proof: Let $\psi(y, v_1, \ldots, v_{n+1})$ witness that g is definable in $\mathcal{N}$ and $\eta(y, w_1, \ldots, w_{n-1})$ witness that h is definable in $\mathcal{N}$. Let $\chi(y, x_1, \ldots, x_n)$ be the formula which says the following: There exist a and b such that

(1) $\eta(\beta(b, a, 0), x_2, \ldots, x_n))$ and

(2) for every $0 \leq i < x_1$, $\psi(\beta(b, a, i+1), \beta(b, a, i), x_1, \ldots, x_n)$ and

(3) $y = \beta(b, a, x_1)$.

Then χ witnesses that f is definable in $\mathcal{N}$. □

Definability of computable functions

By the previous theorem and the fact that the zero functions, projections and successor functions are definable in $\mathcal{N}$, it follows that **every primitive recursive function is definable in** $\mathcal{N}$. To show that every total general recursive function is computable, we will use the following theorem. It says that every general recursive function can be obtained by at most one application of "unbounded search" to a **primitive recursive** function. Its proof will not be covered in this class.

Theorem (Kleene's normal form)

For each $k \geq 1$, there exist primitive recursive functions $T : \omega^{k+2} \to \omega$ and $U : \omega \to \omega$ such that for every k-ary partial computable function f, there exists $e < \omega$ such that for every $(x_1, \ldots, x_k) \in \omega^k$, the following hold.

(A) $(x_1, \ldots, x_k) \in dom(f)$ *iff there exists* z *such that* $T(z, e, x_1, \ldots, x_k) = 0$.

(B) *If* $(x_1, \ldots, x_k) \in dom(f)$, *then* $f(x_1, \ldots, x_k) = U(z)$ *where* z *satisfies:* $T(z, e, x_1, \ldots, x_k) = 0$ *and for every* $y < z$, $T(y, e, x_1, \ldots, x_k) \neq 0$.

Definability of computable functions

The next theorem says that definable functions in $\mathcal{N}$ are closed under "unbounded searches that always terminate".

Theorem

Suppose $g : \omega^{n+1} \to \omega$ is definable in $\mathcal{N}$. Assume that $f : \omega^n \to \omega$ is total where f is defined by $f(x_1, \ldots, x_n) = z$ iff $g(z, x_1, \ldots, x_n) = 0$ and for every $y < z$, $g(y, x_1, \ldots, x_n) \neq 0$. Then f is definable in $\mathcal{N}$.

Proof: Let $\psi(y, v_1, \ldots, v_{n+1})$ witness that g is definable in $\mathcal{N}$. Let $\eta(z, x_1, \ldots, x_n)$ be the formula that says $\psi(0, z, x_1, \ldots, x_n)$ and for every $t < z$, $\neg(\psi(0, t, x_1, \ldots, x_n))$. Then η witnesses that f is definable in $\mathcal{N}$. □

Theorem (Computable functions are definable)

Every total computable function is definable in $\mathcal{N}$.

Proof: As noted before, it suffices to show that every total general recursive function is definable in $\mathcal{N}$. But this easily follows from the above theorem and Kleene's normal form theorem. □

Gödel numbering

Let Σ_{PA} be the alphabet that has the following 16 symbols:

$$0, S, +, \cdot, x, ', \neg, \vee, \wedge, \implies, \iff, (,), \forall, \exists, =$$

For each symbol s, define $pos(s) = m$ iff s is the mth symbol listed above. For example, $pos(0) = 1$, $pos(S) = 2$ and $pos(=) = 16$. Let p_k denote the kth prime number. So $p_0 = 2$, $p_1 = 3$, $p_5 = 11$ and so on.

For a string σ over Σ_{PA}, define the **Gödel number of** σ, denoted $\ulcorner \sigma \urcorner$, as follows. If $\sigma = s_0 s_1 s_2 \ldots s_k$, then $\ulcorner \sigma \urcorner = p_0^{a_0} p_1^{a_1} \ldots p_k^{a_k}$ where $a_i = pos(s_i)$.

Example: $\ulcorner (x' = S(x)) \urcorner = 2^{12} 3^5 5^6 7^{16} 11^2 13^{12} 17^5 19^{13} 23^{13}$.

A fixed point theorem

Theorem (Fixed point theorem)

For every $\mathcal{L}_{PA}$-formula $\eta(x)$, there exists an $\mathcal{L}_{PA}$-sentence ψ such that

$$\mathcal{N} \models \psi \text{ iff } \mathcal{N} \models \eta(\ulcorner \psi \urcorner)$$

Proof: In what follows, for $n < \omega$ we will write $\overline{n}$ for the closed term $S^n(0)$. Define $d : \omega \to \omega$ as follows.

(1) If n is not the Gödel number of an $\mathcal{L}_{PA}$-formula with exactly one free variable, then $d(n) = 0$.

(2) Suppose n is the Gödel number of an $\mathcal{L}_{PA}$-formula $\phi(x)$ with exactly one free variable x. Let θ be the sentence $\phi(\overline{n})$. So θ is obtained by replacing every free occurrence of x in $\phi(x)$ by the closed term $\overline{n}$. Define $d(n) = \ulcorner \theta \urcorner$.

Note that d is computable and hence definable in $\mathcal{N}$. It follows that there is an $\mathcal{L}_{PA}$-formula $\phi(x)$ such that for every $n < \omega$, $\mathcal{N} \models \phi(n)$ iff $\mathcal{N} \models \eta(d(n))$. Let $m = \ulcorner \phi(x) \urcorner$. Let ψ be the sentence $\phi(\overline{m})$. Then $\mathcal{N} \models \psi$ iff $\mathcal{N} \models \phi(m)$ iff $\mathcal{N} \models \eta(d(m))$. By definition, $d(m)$ is the Gödel number of ψ. Therefore $\mathcal{N} \models \psi$ iff $\mathcal{N} \models \eta(\ulcorner \psi \urcorner)$. So ψ is as required. □

Tarski's undefinability of truth

Recall that TA (true arithmetic) is the set of all $\mathcal{L}_{PA}$-sentences ψ such that $\mathcal{N} \models \psi$.

Definition

Define $True_{\mathcal{N}} = \{\ulcorner \psi \urcorner : \psi \in TA\}$.

Theorem (Undefinability of truth in arithmetic)

$True_{\mathcal{N}}$ is not definable in $\mathcal{N}$.

Proof: Suppose not and towards a contradiction, fix an $\mathcal{L}_{PA}$-formula $\phi(x)$ such that for every $n < \omega$,

$$n \in \text{True}_{\mathcal{N}} \text{ iff } \mathcal{N} \models \phi(n)$$

Applying the fixed point theorem to $\eta \equiv \neg\phi$, it follows that there is an $\mathcal{L}_{PA}$-sentence ψ such that $\mathcal{N} \models \psi$ iff $\mathcal{N} \models \neg\phi(\ulcorner \psi \urcorner)$. Put $m = \ulcorner \psi \urcorner$. Then $m \in \text{True}_{\mathcal{N}}$ iff $\mathcal{N} \models \psi$ iff $\mathcal{N} \models \neg\phi(m)$: A contradiction. $\square$

Axiomatizing true arithmetic

Theorem (Arithmetical truth is undecidable)

TA is not computable.

Proof: Suppose TA is computable. Then $\text{True}_{\mathcal{N}}$ is also computable and hence definable in $\mathcal{N}$. But we just showed that this is impossible. □

Theorem (Gödel's incompleteness theorem)

TA is not computably axiomatizable. In particular, PA is incomplete.

Proof: Assume TA is computably axiomatizable and fix a computable set A of $\mathcal{L}_{PA}$-sentences such that A axiomatizes TA. Since TA is complete, A is a complete $\mathcal{L}_{PA}$-theory. By a previous theorem, this implies that the set of theorems in A must be computable. Since A axiomatizes TA, the set of theorems in A is TA. So TA is computable: A contradiction.

Since PA is a computable subset of TA, it follows that PA does not axiomatize TA. So there exists a sentence $\phi \in TA$ such that $PA \nvdash \phi$. Since every theorem in PA is in TA, it also follows that $PA \nvdash \neg\phi$. Therefore PA is incomplete. □

Numeralwise representability in PA

For $n < \omega$, we will denote the closed $\mathcal{L}_{PA}$-term $S^n(0)$ by $\overline{n}$.

Definition

Let $R \subseteq \omega^k$. We say that R is numeralwise representable in PA iff there exists an $\mathcal{L}_{PA}$-formula $\psi(x_1, \ldots, x_k)$ such that for every $(n_1, \ldots, n_k) \in \omega^k$,

(1) *If $(n_1, \ldots, n_k) \in R$, then $PA \vdash \psi(\overline{n_1}, \ldots, \overline{n_k})$*

(2) *If $(n_1, \ldots, n_k) \notin R$, then $PA \vdash \neg\psi(\overline{n_1}, \ldots, \overline{n_k})$*

A **function $f : \omega^k \to \omega$ is numeralwise representable in PA** iff the its graph $R = \{(y, x_1, \ldots, x_k) \in \omega^{k+1} : f(x_1, \ldots, x_k) = y\}$ is numeralwise representable in PA.

Theorem

Every computable function/relation is numeralwise representable in PA.

Proof: The proof of this fact is essentially the same as the one that showed that every computable function/relation is definable in $\mathcal{N}$. □

Undecidability of PA

Theorem (PA is undecidable)

$\{\phi : \phi$ is an $\mathcal{L}_{PA}$-sentence and $PA \vdash \phi\}$ is not computable.

Proof: Suppose not and towards a contradiction, fix a program P such that for every $\mathcal{L}_{PA}$-sentence ϕ, the program P on input ϕ returns 1 if $PA \vdash \phi$ and returns 0 otherwise.
Let $H \subseteq \omega$ be a non-computable c.e. set (for example, the halting problem). Since every c.e. set is the range of some total computable function (see Homework), there exists a computable $f : \omega \to \omega$ such that range$(f) = H$. Being computable, f is numeralwise representable in PA so we can fix a formula $\psi(y, x)$ such that for every $(m, n) \in \omega^2$,

1. If $f(n) = m$, then $PA \vdash \psi(\overline{m}, \overline{n})$ and
2. If $f(n) \neq m$, then $PA \vdash \neg\psi(\overline{m}, \overline{n})$.

Exercise: Let Q be a program that on input $m < \omega$ runs P with input $(\exists x)(\psi(\overline{m}, x))$ and outputs whatever P does. Show that Q computes H.

But H is not computable so we get a contradiction. Hence $\{\phi : \phi$ is an $\mathcal{L}_{PA}$-sentence and $PA \vdash \phi\}$ is not computable. □