Mathematical Logic and Set Theory

1 Basic set theory

Iterative concept of set.

(a) Sets are formed in stages $0, 1, \ldots, s, \ldots$.

(b) For each stage s, there is a next stage $s + 1$.

(c) There is an "absolute infinity" of stages.

(d) V_s = the collection of all sets formed before stage s.

(e) $V_0 = \emptyset$ = the empty collection.

(f) V_{s+1} = the collection of (a) all sets belonging to V_s and (b) all subcollections of V_s not previously formed into sets.

Remarks. (1) A set is formed after its members. (2) V_s itself is formed as a set at stage s.

Formal language for talking about sets.

Symbols:

$v_0, v_1, v_2, \ldots$	variables
$=$	meaning "is identical with"
$\in$	meaning "is a member of"
$\neg$	meaning "not"
$\wedge$	meaning "and"
$\exists$	meaning "there is a"
(	
)	

Formulas (inductive definition):

(i) If x and y are variables, then $x = y$ and $x \in y$ are (atomic) formulas.

(ii) If x is a variable and φ and ψ are formulas, then $\neg\varphi$, $(\varphi \wedge \psi)$, and $(\exists x)\, \varphi$ are formulas.

(iii) Nothing is a formula unless (i) and (ii) require it to be.

Free occurences of a variable in a formula:

(i) All occurrences of variables in atomic formulas $x \in y$ and $x = y$ are free.

(ii) An occurrence of x in $\neg\varphi$ is free just in case the corresponding occurrence of x in φ is free.

(iii) An occurrence of x in $(\varphi \wedge \psi)$ is free just in case the corresponding occurrence of x in φ or in ψ is free.

(iv) An occurrence of x in $(\exists y)\,\varphi$ is free just in case x is not y and the corresponding occurrence of x in φ is free.

Non-free occurrences of a variable in a formula are called *bound* occurrences. We write "$\varphi(x_1, \ldots, x_n)$" for "$\varphi$" to indicate that all variables occurring free in φ are among the (distinct, in the default case) variables $x_1, \ldots, x_n$.

Abbreviations:

$$\begin{array}{rcl} (\varphi \vee \psi) & \text{for} & \neg(\neg\varphi \wedge \neg\psi) \\ (\varphi \rightarrow \psi) & \text{for} & (\neg\varphi \vee \psi) \\ (\varphi \leftrightarrow \psi) & \text{for} & ((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)) \\ (\forall x) & \text{for} & \neg(\exists x)\neg \\ x \neq y & \text{for} & \neg\, x = y \\ x \notin y & \text{for} & \neg\, x \in y \end{array}$$

We often omit parentheses, and we often write "x," "y," etc., when we should be writing "v" with subscripts.

The Zermelo–Fraenkel (ZFC) Axioms. Below we list the formal ZFC axioms. Following each axiom, we give in parentheses an informal version of it. Our official axioms are the formal ones.

For all the axioms other than those of the Comprehension and Replacement Schema, let us use the following scheme of "abbreviation":

$$\begin{array}{llllllllllll} x & \text{for} & v_1 & z & \text{for} & v_3 & w & \text{for} & v_5 & y_2 & \text{for} & v_7 \\ y & \text{for} & v_2 & u & \text{for} & v_4 & y_1 & \text{for} & v_6 & & & \end{array}$$

For the two schemata, the variables are arbitrary. E.g., there is an instance of Comprehension for each formula φ and sequence $x, y, z, w_1, \ldots, w_n$ of distinct variables that contains all variables occurring free in φ plus the variable y that does not so occur.

Axiom of Set Existence:

$$(\exists x)\, x = x.$$

(There is a set.)

Axiom of Extensionality:

$$(\forall x)(\forall y)\left((\forall z)(z \in x \leftrightarrow z \in y) \rightarrow x = y\right).$$

(Sets that have the same members are identical.)

Axiom of Foundation:

$$(\forall x)\left((\exists y)\, y \in x \rightarrow (\exists y)(y \in x \,\wedge\, (\forall z)(z \notin x \,\vee\, z \notin y))\right).$$

(Every non-empty set x has a member that has no members in common with x.)

Axiom Schema of Comprehension: For each formula $\varphi(x, z, w_1, \ldots, w_n)$,

$$(\forall w_1)\cdots(\forall w_n)(\forall z)(\exists y)(\forall x)\left(x \in y \leftrightarrow (x \in z \,\wedge\, \varphi)\right).$$

(For any set z and any property P, there is a set whose members are those members of z that have property P.)

Axiom of Pairing:

$$(\forall x)(\forall y)(\exists z)(x \in z \,\wedge\, y \in z).$$

(For any sets x and y, there is a set to which both x and y belong, i.e., of which they are both members.)

Axiom of Union:

$$(\forall x)(\exists y)(\forall z)(\forall w)\left((w \in z \,\wedge\, z \in x) \rightarrow w \in y\right).$$

(For any set x, there is a set to which all members of members of x belong.)

The axioms of Pairing, Union, and Comprehension give us some operations on sets. For any x and y, $\{x, y\}$ is the set whose members are exactly x and y. (It exists by Pairing and Comprehension.) Let $\{x \mid \varphi(x, \ldots)\}$ be the set of all x such that $\varphi(x, \ldots)$ holds, if this is a set. For any set x,

$$\mathcal{U}(x) = \{z \mid (\exists y)(z \in y \,\wedge\, y \in x)\}.$$

($\mathcal{U}(x)$ exists by Union and Comprehension.) For any sets x and y, $x \cup y$ is the set $\mathcal{U}(\{x, y\})$. For any sets $x_1, \ldots, x_n$, $\{x_1, \ldots, x_n\}$ is the set whose members are exactly $x_1, \ldots, x_n$. (To see that this set exists, note that $\{x\} = \{x, x\}$ for any set x and that $\{x_1, \ldots, x_{m+1}\} = \{x_1, \ldots, x_m\} \cup \{x_{m+1}\}$ for $0 \leq m < n$.)

In the statement of the next axiom, "$(\exists! y)$" is short for the obvious way of expressing "there is exactly one y."

Axiom Schema of Replacement: For each formula $\varphi(x, y, z, w_1, \ldots, w_n)$,

$$(\forall w_1) \cdots (\forall w_n)(\forall z)\,((\forall x)(x \in z \to (\exists! y)\,\varphi) \\ \to (\exists u)(\forall x)(x \in z \to (\exists y)(y \in u \wedge \varphi)))\,.$$

(For any set z and any relation R, if each member x of z bears R to exactly one set y_x, then there is a set to which all these y_x belong.)

Remark. By Comprehension, Replacement can be strengthened to give

$$(\forall w_1) \cdots (\forall w_n)(\forall z)\,((\forall x)(x \in z \to (\exists! y)\,\varphi) \\ \to (\exists u)(\forall y)(y \in u \leftrightarrow (\exists x)(x \in z \wedge \varphi)))\,.$$

Define $\mathcal{S}(x) = x \cup \{x\}$. Note that $\emptyset$ exists by Set Existence and Comprehension.

Axiom of Infinity:

$$(\exists x)\,(\emptyset \in x \wedge (\forall y)(y \in x \to \mathcal{S}(y) \in x))\,.$$

(There is a set that has the empty set as a member and is closed under the operation $\mathcal{S}$.)

Let "$z \subseteq x$" abbreviate "$(\forall w)(w \in z \to w \in x)$."

Axiom of Power Set.

$$(\forall x)(\exists y)(\forall z)(z \subseteq x \to z \in y).$$

(For any set x, there is a set to which all subsets of x belong.)

Let $\mathcal{P}(x) = \{z \mid z \subseteq x\}$. (It exists by Power Set and Comprehension.) Let $x \cap y = \{z \mid z \in x \wedge z \in y\}$. (It exists by Comprehension.)

Axiom of Choice:

$$(\forall x)\,((\forall y_1)(\forall y_2)\,((y_1 \in x \,\wedge\, y_2 \in x) \to (y_1 \neq \emptyset \,\wedge\, (y_1 = y_2 \,\vee\, y_1 \cap y_2 = \emptyset))) \\ \to\; (\exists z)(\forall y)(y \in x \to (\exists! w)\, w \in y \cap z))\,.$$

(If x is a set of non-empty sets no two of which have any members in common, then there is a set that has exactly one member in common with each member of x.)

Remark. For all the axioms except Comprehension and Replacement, the formal and informal versions are equivalent. But the formal Comprehension and Replacement Schemata are *prima facie* weaker than the informal versions. The formal schemata apply, not to arbitrary properties and relations, but only to properties and relations characterizable by formulas of the formal language. (Warning: We shall later use the word "relation" in a precise technical sense quite different from the intuitive way we used the word in stating the informal version of Replacement.)

Justifications of the axioms. The ZFC axioms are supposed to be true of the iterative concept of set. Following is an axiom-by-axiom attempt to explain why.

Set Existence. $\emptyset$ belongs to V_1.

Extensionality. It follows from the notion of identity for collections.

Foundation. Assume $x \neq \emptyset$. Let w be the collection of all sets formed before any member of x is formed. Some member of x is formed at some stage s. Since w is a subcollection of V_s, clause (f) of the iterative concept implies that w is formed as a set at some stage s_1 no later than s. No $y \in x$ can be formed at a stage s_2 before s_1, for then w would be a subcollection of V_{s_2} and so would be formed at or before s_2. If no $y \in x$ were formed at s_1, then V_{s_1+1} would be included in w, and so w would belong to itself, an impossibility. Any $y \in x$ formed at s_1 has the right properties.

Comprehension. The desired y is a subcollection of z and so of V_s, where z is formed at s.

Pairing. If x and y are formed at or before s, then they belong to V_{s+1}, which therefore works as z.

Union. If x is formed at s, then all members of x, and so all members of members of x, belong to V_s. Hence V_s works as y.

Replacement. For each $x \in z$, let s_x be the stage at which the unique y such that $\varphi(x, y, z, w_1, \ldots, w_n)$ is formed. The collection of all these s_x is no larger than the set z, so "absolute infinity" demands that there be a stage s later than all the s_x. Then V_s works as u.

Infinity. By absolute infinity, there is an infinite stage s. Let x be the collection of all y in V_s that are formed at finite stages. Then x has the required properties and is formed at or before s.

Power Set. If x is formed at s and if $z \subseteq x$, then $z \subseteq V_s$ and so $z \in V_{s+1}$. Thus V_{s+1} works for y.

Choice. If x is formed at s, then we are looking for a z that might as well be a subcollection of $\mathcal{U}(x) \subseteq V_s$. What we have to convince ourselves is that such a subcollection exists.

The *ordered pair* $\langle x, y\rangle$ of sets x and y is $\{\{x\}, \{x, y\}\}$. Note that

$$\langle x, y\rangle = \langle z, w\rangle \leftrightarrow (x = z \wedge y = w).$$

Exercise 1.1. Write a formula of the formal language expressing the statement that $v_0 = \langle v_1, v_2\rangle$.

The *Cartesian product* $u \times v$ of sets u and v is $\{\langle x, y\rangle \mid x \in u \wedge y \in v\}$.

Theorem 1.1. *$u \times v$ always exists.*

Proof 1. Let $x \in u$. Then $(\forall y \in v)(\exists! w)\, w = \langle x, y\rangle$. Here, and later, we use obvious abbreviations, such as "$(\forall y \in v) \ldots$," without explicit mention. By Replacement and Comprehension, let $z_x = \{w \mid (\exists y \in v)\, w = \langle x, y\rangle\}$. Then $(\forall x \in u)(\exists! z)\, z = z_x$. (Note that there is a formula $\psi(x, z, u, v)$ expressing the statement that $z = z_x$.) By Replacement and Comprehension, let $q = \{z_x \mid x \in u\}$. The Cartesian product of u and v is $\mathcal{U}(q)$. □

Proof 2. $\mathcal{P}(\mathcal{P}(u \cup v))$ exists by Power Set and Comprehension. If $x \in u$ and $y \in v$, then $\langle x, y\rangle \in \mathcal{P}(\mathcal{P}(u \cup v))$. Thus $u \times v$ exists by Comprehension. □

Remark. Proof 1 used Replacement but not Power Set. Proof 2 used Power Set but not Replacement.

A *relation* is a set of ordered pairs. A *function* is a relation f such that

$$(\forall x)(\forall y_1)(\forall y_2)((\langle x, y_1\rangle \in f \wedge \langle x, y_2\rangle \in f) \rightarrow y_1 = y_2).$$

The definitions of a *one-one* function, the *domain* of a function, and the *range* of a function are the obvious ones. The notation $f : x \to y$ means, as usual, that f is a function whose domain is x and whose range is $\subseteq y$.

A set r is a *linear ordering* of a set x if r is a *relation on* x (i.e., $r \subseteq x \times x$) and r linearly orders x in the usual strict sense (i.e., we require that $\langle y, y \rangle \notin r$).

A relation r is *wellfounded* if

$$(\forall x)(x \neq \emptyset \;\to\; (\exists y \in x)(\forall z \in x)\, \langle z, y \rangle \notin r).$$

Example. Let u be a set. Let

$$\in \restriction u = \{\langle z, y \rangle \in u \times u \mid z \in y\}.$$

The Axiom of Foundation says that $\in \restriction u$ is wellfounded for every u.

We say that r *is a wellordering* of x if r is a linear ordering of x and r is wellfounded. We say that r *wellorders* x if r is a relation and $r \cap (x \times x)$ is a wellordering of x.

A set x is *transitive* if $\mathcal{U}(x) \subseteq x$.

An *ordinal number* is a set x such that

(1) x is transitive;

(2) $\in \restriction x$ wellorders x.

Remark. Foundation implies that (2) is equivalent with the assertion that $\in \restriction x$ linearly orders x.

Exercise 1.2. Let x and y be ordinal numbers. Show, without using Foundation, that

$$x \in y \;\vee\; y \in x \;\vee\; x = y.$$

Hint. Let $z = x \cap y$. Assume that $x \not\subseteq y$ and prove that $z \in x \setminus y$ ($=\{a \mid a \in x \wedge a \notin y\}$). To do this, note that the nonempty set $x \setminus y$ must have an $\in$-least member u. Prove that z and u have the same members. Similarly prove that if $y \not\subseteq x$ then $z \in y \setminus x$. Finally, consider the four possible answers as to whether $y \subseteq x$ and $x \subseteq y$.

The set ω is defined as follows:

$$x \in \omega \;\leftrightarrow\; (\forall y)((\emptyset \in y \;\wedge\; (\forall z)(z \in y \;\to\; \mathcal{S}(z) \in y)) \;\to\; x \in y).$$

ω exists by Infinity and Comprehension. Note that

$$\emptyset \in \omega \wedge (\forall z)(z \in \omega \to \mathcal{S}(z) \in \omega).$$

The members of ω are called *natural numbers*.

Remark. In preparation for metamathematical results in 220C, we shall make note of all uses of Foundation or Choice in proving theorems, and we shall avoid using these axioms unnecessarily. In particular, we avoid using Foundation in the following proofs, although using it would simplify matters.

Theorem 1.2. *ω is a set of ordinal numbers; i.e., every natural number is an ordinal number.*

Proof. Let $y = \{n \in \omega \mid n \text{ is an ordinal number}\}$; y exists by Comprehension. It is easy to see that $\emptyset \in y$. Let $n \in \omega$. We assume that $n \in y$ and show that $\mathcal{S}(n) \in y$. This will prove that $\omega \subseteq y$, and so that $y = \omega$.

By the definition of $\mathcal{S}(n)$,

$$(\forall u)(u \in \mathcal{S}(n) \leftrightarrow (u \in n \vee u = n)).$$

Hence, for any v, $v \in \mathcal{U}(\mathcal{S}(n)) \Leftrightarrow (v \in \mathcal{U}(n) \vee v \in n) \Rightarrow$ (since n is transitive) $v \in n \Rightarrow v \in \mathcal{S}(n)$. Hence $\mathcal{S}(n)$ is transitive.

$n \notin n$, since otherwise $\in \restriction n$ is not wellfounded, indeed is not even a linear ordering of n. Moreover n does not belong to any $u \in n$, since otherwise transitivity gives $n \in n$. Thus the relation $\in \restriction \mathcal{S}(n)$ is just the wellordering $\in \restriction n$ with n stuck on at the end. It is easy to prove that $\in \restriction \mathcal{S}(n)$ is a wellordering, using the fact that $\in \restriction n$ is wellordering. □

Remark. The method used to prove the last theorem is mathematical induction. To prove that every natural number has a property (such as being an ordinal number), we prove that $\emptyset$ has the property and that if $n \in \omega$ has the property then so does $\mathcal{S}(n)$. By the definition of ω, this implies that the set of all natural numbers with the property is all of ω, i.e., that every natural number has the property.

Theorem 1.3. *ω is an ordinal number.*

Proof. Let $y = \{n \in \omega \mid n \subseteq \omega\}$. To prove that ω is transitive, we must show that $y = \omega$. We use mathematical induction. Trivially $\emptyset \in y$. Suppose $n \in y$. Then $u \in \mathcal{S}(n) \Leftrightarrow (u \in n \vee u = n) \Rightarrow u \in \omega$. Hence $\mathcal{S}(n) \subseteq \omega$. But also $\mathcal{S}(n) \in \omega$, so $\mathcal{S}(n) \in y$.

Theorem 1.2 and its proof show that $\in\restriction\omega$ is irreflexive ($n \notin n$ for $n \in \omega$) and asymmetric ($m \in n \to n \notin m$ for m and n elements of ω). The fact that every member of ω is transitive implies directly that $\in\restriction\omega$ is a transitive relation ($k \in m \in n \to k \in n$ for k, m, and n elements of ω). Exercise 1.2 and Theorem 1.2 imply that $\in\restriction\omega$ is connected ($m \in n \vee n \in m \vee m = n$ for m and n elements of ω). Thus $\in\restriction\omega$ is a linear ordering of ω.

To show that $\in\restriction\omega$ is wellfounded, we prove that each non-empty subset of ω has an $(\in\restriction\omega)$-least element. Let $v \subseteq \omega$ with $v \neq \emptyset$. Let $n \in v$. If $n \cap v = \emptyset$, then n is the $(\in\restriction\omega)$-least element of v. Suppose then that $n \cap v \neq \emptyset$. By Theorem 1.2, the set $n \cap v$ has an $(\in\restriction n)$-least element m. The transitivity of n implies that m is also the $(\in\restriction\omega)$-least element of v. □

Sometimes we shall want to assert *theorem schemata* rather than simple theorems: we shall want to assert that, for every formula φ, some sentence derived from φ is a theorem. A convenient way to do this is to speak of *classes.* We shall speak of $\{x \mid \varphi(x, \ldots)\}$ as a *class* whether or not there is a set $\{x \mid \varphi(x, \ldots)\}$. When the set exists, we identify the set and the class. When the set does not exist, we call $\{x \mid \varphi(x, \ldots)\}$ a *proper class.* Lower case letters will be used only for sets. Upper case letters will be used mostly for classes.

Terms like *relation, function, domain, wellfounded*, etc. are defined for classes just as they are for sets. In class language, the Comprehension Schema says that the intersection of a class and a set is a set.

Let $V = \{x \mid x = x\}$. V is a proper class, since otherwise Comprehension would yield the self-contradictory *Russell set* $\{x \mid x \notin x\}$.

An example of a proper class relation is $\in = \{\langle x, y\rangle \mid x \in y\}$. In the hint to Exercise 1.2, we wrote "$\in$" instead of $\in\restriction x$ and $\in\restriction y$. Retroactively this notation is now explained.

Exercise 1.3. Prove that $\in$ is a proper class.

If F is a class function and A is a class, then $F\restriction A = \{\langle x, y\rangle \in F \mid x \in A\}$.

Theorem 1.4 (Schema of Definition by Recursion). *Let $F : V \to V$. There is a unique (set) $g : \omega \to V$ such that*

$$(\forall n \in \omega)\, g(n) = F(g\restriction n).$$

Proof. We first show that

$$(\forall n \in \omega)(\exists! g)(g : n \to V \ \wedge\ (\forall m \in n)\, g(m) = F(g\restriction m)).$$

For $n = \emptyset$, the empty g (i.e., $\emptyset$) works. Suppose $g : n \to V$ is the unique function with the property $(\forall m \in n)\, g(m) = F(g \restriction m)$. Let $g' = g \cup \{\langle n, F(g)\rangle\}$. Clearly $g' : \mathcal{S}(n) \to V$ and $(\forall m \in \mathcal{S}(n))\, g'(m) = F(g' \restriction m)$. If $h : \mathcal{S}(n) \to V$ satisfies $(\forall m \in \mathcal{S}(n))\, h(m) = F(h \restriction m)$, then $h \restriction n = g$ by the uniqueness property of g. But then $h(n) = F(h \restriction n) = F(g) = g'(n)$, and so $h = g'$. Our conclusion follows by induction.

By Replacement and Comprehension, let

$$z = \{y \mid (\exists n \in \omega)(y : n \to V \ \wedge \ (\forall m \in n)\, y(m) = F(y \restriction m))\}.$$

Suppose y_1 and y_2 belong to z. Let $y_1 : n_1 \to V$ and $y_2 : n_2 \to V$. If $n_1 = n_2$ then the uniqueness part of the assertion proved in the last paragraph gives $y_1 = y_2$. If $n_1 \in n_2$ then uniqueness gives $y_1 = y_2 \restriction n_1$; if $n_2 \in n_1$ then uniqueness gives $y_2 = y_1 \restriction n_1$. Thus $y_1 \subseteq y_2$ or $y_2 \subseteq y_1$. Let $g = \mathcal{U}(z)$. It is easy to see that g is a function and that $\text{domain}\,(g) \subseteq \omega$. To see that $\omega \subseteq \text{domain}\,(g)$, use the existence part of the assertion of the last paragraph to get, for each $n \in \omega$, a $y \in z$ with $y : \mathcal{S}(n) \to V$. It is easy to see that $(\forall n \in \omega)\, g(n) = F(g \restriction n)$. For uniqueness, assume that $(\forall n \in \omega)\, h(n) = F(h \restriction n)$. For each $n \in \omega$, $g \restriction \mathcal{S}(n) = h \restriction \mathcal{S}(n)$, and so $g(n) = h(n)$. □

Remark. We needed Replacement only to get that g is a set (rather than a proper class).

Theorem 1.5. $(\forall x)(\exists y)(y \text{ is transitive } \wedge x \subseteq y)$.

Proof. Define $F : V \to V$ by

$$F(z) = u \ \leftrightarrow \ \begin{cases} z \text{ is not a function and } u = \emptyset \\ \text{or } z \text{ is a function and } u = x \cup \mathcal{U}(\mathcal{U}(\text{range}\,(z))). \end{cases}$$

Let g be given by Theorem 1.4. Let $y = \mathcal{U}(\text{range}\,(g))$. Suppose $v \in y$. Then $v \in g(n)$ for some $n \in \omega$. Hence $v \in \mathcal{U}(\text{range}\,(g \restriction \mathcal{S}(n)))$. Therefore

$$v \subseteq \mathcal{U}(\mathcal{U}(\text{range}\,(g \restriction \mathcal{S}(n)))) \subseteq F(g \restriction \mathcal{S}(n)) = g(\mathcal{S}(n)) \subseteq y.$$

Since $x = g(0)$, it follows that $x \subseteq y$. □

For any class A, let

$$\bigcap A = \{z \mid (\forall y \in A)\, z \in y\}.$$

Comprehension gives that $\bigcap A$ is a set if A is non-empty. Note that $\omega = \bigcap\{y \mid \emptyset \in y \wedge (\forall z \in y)\, \mathcal{S}(z) \in y\}$. The operation dual, in a natural sense, to $\bigcap$ is the operation $\mathcal{U}$. We shall hence sometimes write $\bigcup x$ for $\mathcal{U}(x)$.

For any set x let

$$\operatorname{trcl}(x) = \bigcap\{y \mid y \text{ is transitive } \wedge\ x \subseteq y\}.$$

Theorem 1.5 implies that $\operatorname{trcl}(x)$, the *transitive closure* of x, is always a set.

Theorem 1.6. *Let*

$$\mathrm{ON} = \{x \mid x \text{ is an ordinal number}\}.$$

The (class) relation $\in \restriction \mathrm{ON}$ *is a wellordering of* ON. *Indeed* $\in \restriction \mathrm{ON}$ *is well-founded in the strong sense that every non-empty subclass of* ON *has an* $\in$*-minimal element. Furthermore* ON *is transitive.*

Proof. The proofs that $\in \restriction \mathrm{ON}$ is irreflexive, asymmetric, transitive, and connected are just like the corresponding parts of the proof of of Theorem 1.3.

Suppose that $A \subseteq \mathrm{ON}$ is a non-empty class. Let $x \in A$. If $x \cap A = \emptyset$, then we are done. Otherwise apply the fact that $x \in \mathrm{ON}$ to $x \cap A$. This gives a $y \in x \cap A$ with $y \cap x \cap A = \emptyset$. If $z \in y \cap A$ then $z \in y \in x \in \mathrm{ON}$, and so $z \in x$.

To prove that ON is transitive, suppose $x \in y \in \mathrm{ON}$. By the transitivity of y, we have that $x \subseteq y$. The fact that $\in \restriction x$ is a wellordering thus follows easily from the fact that $\in \restriction y$ is a wellordering. To show that x is transitive, and so that x is an ordinal number, let $z \in w \in x$. We have that w, and hence z, belongs to y. Since $\in \restriction y$ is a transitive relation, we get that $z \in x$. □

When we talk of $\emptyset$ in its role as an ordinal number, we shall call it 0. We denote $\in \restriction \mathrm{ON}$ by $<$. For ordinals α and β, we write the natural $\alpha < \beta$ to mean that $\langle \alpha, \beta \rangle \in <$, i.e., that $\alpha \in \beta$.

Exercise 1.4. Show, for any ordinal number α, that $\mathcal{S}(\alpha)$ is the immediate successor of α with respect to $<$.

Exercise 1.5. Let x be any set of ordinal numbers. Prove that $\mathcal{U}(x)$ is an ordinal number.

Theorem 1.6 makes possible proof by *transfinite induction*. If we want to show that all ordinal numbers have some property expressed by a formula φ, it is enough to show that, for every ordinal number α,

$$(\forall \beta < \alpha)\, \varphi(\beta, \ldots) \;\rightarrow\; \varphi(\alpha, \ldots).$$

For then Theorem 1.6 implies that the class of $\alpha \in \mathrm{ON}$ such that $\neg\varphi(\alpha, \ldots)$ cannot be non-empty. The following theorem gives us a useful division into cases when we are using transfinite induction.

Theorem 1.7. *If α is an ordinal number, then one of the following holds:*

(1) $(\exists \beta < \alpha)\, \alpha = \mathcal{S}(\beta)$;

(2) $\alpha = \mathcal{U}(\alpha)$.

Proof. Let α be an ordinal number, and assume that (1) fails. Since $\mathcal{U}(\alpha) \subseteq \alpha$ for any ordinal α, we need only show that $\alpha \subseteq \mathcal{U}(\alpha)$. Let $\beta \in \alpha$. By Exercise 1.4, $\mathcal{S}(\beta)$ is an ordinal number $\leq \alpha$. Since (1) fails, we must have $\mathcal{S}(\beta) < \alpha$. But then $\beta \in \mathcal{S}(\beta) \in \alpha$, so $\beta \in \mathcal{U}(\alpha)$. □

Ordinals satisfying (1) are called *successor* ordinals. Non-zero ordinals satisfying (2) are called *limit* ordinals.

Theorem 1.8 (Schema of Definition by Transfinite Recursion). *Let $F : V \to V$. There is a (unique) $G : \mathrm{ON} \to V$ such that*

$$(\forall \alpha \in \mathrm{ON})\, G(\alpha) = F(G \restriction \alpha).$$

Proof. We first show that

$$(\forall \alpha \in \mathrm{ON})(\exists! g)(g : \alpha \to V \;\wedge\; (\forall \beta < \alpha)\, g(\beta) = F(g \restriction \beta)).$$

We argue by transfinite induction. Let α be an ordinal and assume that the statement holds for all smaller ordinals. The case $\alpha = 0$ is trivial. If $\alpha = \mathcal{S}(\beta)$ for some ordinal β, then we argue as in the proof of Theorem 1.4. If α is a limit ordinal, then we use Replacement as for the special case $\alpha = \omega$ in the last part of the proof of Theorem 1.4 to get a z that is the set of all g' that work for ordinals $\beta < \alpha$. We let $g = \mathcal{U}(z)$.

Let

$$G = \mathcal{U}(\{g \mid (\exists \alpha \in \mathrm{ON})(g : \alpha \to V \;\wedge\; (\forall \beta < \alpha)\, g(\beta) = F(g \restriction \beta))\}).$$

It is easy to check that G, and only G, has the required property. □

Remark. Note that the proof gives an explicit definition of G from a definition of F. Thus the theorem really is a theorem schema, and the quantification over proper classes in its statement could be avoided.

Theorem 1.9. *There is a unique* $\mathbf{V} : \mathrm{ON} \to V$ *such that (where we write* V_α *for* $\mathbf{V}(\alpha)$*)*

(a) $V_0 = \emptyset$*;*

(b) $V_{\mathcal{S}(\alpha)} = \mathcal{P}(V_\alpha)$*;*

(c) $V_\lambda = \mathcal{U}(\{V_\alpha \mid \alpha < \lambda\})$ *if* λ *is a limit ordinal.*

Proof. Let $F(x) = \emptyset$ if $x = \emptyset$ or x is not a function whose domain is an ordinal number. If α an ordinal and $x : \mathcal{S}(\alpha) \to V$, then let $F(x) = \mathcal{P}(x(\alpha))$. If λ is a limit ordinal and $x : \lambda \to V$, let $F(x) = \mathcal{U}(\mathrm{range}\,(x))$. The desired function is given by Theorem 1.8. □

Exercise 1.6. Show that $\alpha < \beta \to V_\alpha \subseteq V_\beta$.

Theorem 1.10. (Uses Foundation) $(\forall x)(\exists \alpha)\, x \in V_\alpha$.

Proof. Suppose x belongs to no V_α. Let

$$z = \{u \in \mathrm{trcl}\,(x) \cup \{x\} \mid (\forall \alpha \in \mathrm{ON})\, u \notin V_\alpha\}.$$

Since $z \neq \emptyset$, Foundation gives a $u \in z$ such that $u \cap z = \emptyset$. Every member of u belongs to $\mathrm{trcl}\,(x)$, and so every member of u belongs to some V_α. For $y \in u$, let α_y be the least α such that $y \in V_\alpha$. By Replacement and Comprehension, let $\alpha = \mathcal{U}(\{\alpha_y \mid y \in u\})$. By Exercise 1.5, $\alpha \in \mathrm{ON}$. By Exercise 1.6, $u \subseteq V_\alpha$. This gives the contradiction that $u \in V_{\mathcal{S}(\alpha)}$. □

By transfinite recursion, one can define addition, multiplication, and exponentiation of ordinal numbers as follows:

$$\begin{aligned} \alpha + 0 &= \alpha\,; \\ \alpha + \mathcal{S}(\beta) &= \mathcal{S}(\alpha + \beta)\,; \\ \alpha + \lambda &= \mathcal{U}(\{\alpha + \beta \mid \beta < \lambda\}) \text{ if } \lambda \text{ is a limit ordinal.} \end{aligned}$$

$$\begin{aligned} \alpha \cdot 0 &= 0\,; \\ \alpha \cdot \mathcal{S}(\beta) &= \alpha \cdot \beta + \alpha\,; \\ \alpha \cdot \lambda &= \mathcal{U}(\{\alpha \cdot \beta \mid \beta < \lambda\}) \text{ if } \lambda \text{ is a limit ordinal.} \end{aligned}$$

$$\begin{aligned}
\alpha^0 &= 1 \ \ (=\mathcal{S}(0))\,; \\
\alpha^{\mathcal{S}(\beta)} &= \alpha^\beta \cdot \alpha\,; \\
\alpha^\lambda &= \mathcal{U}(\{\alpha^\beta \mid \beta < \lambda\}) \text{ if } \lambda \text{ is a limit ordinal.}
\end{aligned}$$

The way this is done is as follows: Consider the definition of $+$. We can define a function $F : \mathrm{ON} \times V \to V$, so that, e.g., if α and β are ordinals and $x : \mathcal{S}(\beta) \to V$, then $F(\langle \alpha, x\rangle) = \mathcal{S}(x(\beta))$. If we define $F_\alpha : V \to V$ by $F_\alpha(x) = F(\langle \alpha, x\rangle)$, then Theorem 1.8 applied to F_α gives a function $+_\alpha : \mathrm{ON} \to \mathrm{ON}$. Since the proof of Theorem 1.8 gives us a definition of the $+_\alpha$ from the parameter α, we get an explicit definition of $+$.

Note that $\alpha + 1 = \mathcal{S}(\alpha)$ for every ordinal α. We shall often write $\alpha + 1$ instead of $\mathcal{S}(\alpha)$. For the rest of this section, however, we shall continue to write $\mathcal{S}(\alpha)$ in order to avoid confusion with the different kind of addition that we shall shortly define.

We now turn to the subject of cardinal numbers. If x and y are sets, let us say that $x \preceq y$ if there is a one-one $f : x \to y$. By $x \approx y$ we mean that there is a one-one onto $f : x \to y$.

Theorem 1.11 (Schröder–Bernstein Theorem). *If $x \preceq y$ and $y \preceq x$ then $x \approx y$.*

Proof. Let $f : x \to y$ and $g : y \to x$ be one-one. Using Theorem 1.4, define $h : x \times \omega \to x$ by

$$\begin{aligned}
h(z, 0) &= z\,; \\
h(z, \mathcal{S}(n)) &= g(f(h(z, n)))\,.
\end{aligned}$$

Let

$$u = \{z \in x \mid (\exists v \in x)(\exists n \in \omega)(h(v, n) = z \ \wedge \ v \notin \mathrm{range}\,(g))\}.$$

Note that if $z \notin u$ then $z \in \mathrm{range}\,(g)$. Let $k : x \to y$ be given by

$$k(z) = \begin{cases} f(z) & \text{if } z \in u; \\ g^{-1}(z) & \text{if } z \notin u. \end{cases}$$

(If r is any relation, $r^{-1} = \{\langle w, w'\rangle \mid \langle w', w\rangle \in r\}$. Since g is a one-one function, we have that $g^{-1} : \mathrm{range}\,(g) \to y$.)

To see that k is one-one, assume that $k(z_1) = k(z_2)$. Exchanging z_1 and z_2 if necessary, we may assume that either $z_1 = z_2$ or else $z_1 \in u$ and $z_2 \notin u$. Assume for a contradiction that the latter is the case. Then $f(z_1) = g^{-1}(z_2)$, and so $g(f(z_1)) = z_2$. Let v and n witness that $z_1 \in u$. Since $h(v, n) = z_1$, we get that $g(f(h(v, n))) = g(f(z_1)) = z_2$. This means that $h(v, \mathcal{S}(n)) = z_2$, contradicting the fact that $z_2 \notin u$.

Assume that $z \in y \setminus \text{range}\,(k)$. Then $g(z) \in u$, since otherwise $k(g(z)) = g^{-1}(g(z)) = z$. Let v and n witness that $g(z) \in u$. Obviously $n \neq 0$. Thus $n = \mathcal{S}(m)$ for some m. We have then that $g(z) = h(v, \mathcal{S}(m)) = g(f(h(v, m)))$. Hence $z = f(h(v, m))$. But $h(v, m) \in u$, and so we get the contradiction that

$$k(h(v, m)) = f(h(v, m)) = z.$$

□

A *cardinal number* is an ordinal number α such that $(\forall \beta < \alpha)\, \beta \not\approx \alpha$.

Theorem 1.12. *Every natural number is a cardinal number. ω is a cardinal number.*

Proof. For the first assertion, we show that

$$(*) \qquad (\forall n \in \omega)(\forall f)((f : n \to n \,\wedge\, f \text{ one-one}) \;\to f \text{ onto}).$$

The case $n = 0$ is trivial. Let $f : \mathcal{S}(n) \to \mathcal{S}(n)$ be one-one. We must have that $n \in \text{range}\,(f)$, since otherwise $f \restriction n : n \to n$ is not onto. Let $a = f(n)$ and let $f(b) = n$. Define $g : n \to n$ by

$$g(m) = \begin{cases} f(m) & \text{if } m \neq b; \\ a & \text{if } m = b. \end{cases}$$

By the induction hypothesis, range $(g) = n$. Thus

$$\text{range}\,(f) = \{n\} \cup \text{range}\,(g) = \mathcal{S}(n).$$

For the second assertion, note that if $n \in \omega$ and $f : \omega \to n$ is one-one, then $f \restriction \mathcal{S}(n) : \mathcal{S}(n) \to n$ contradicts $(*)$. □

Theorem 1.13. *Let $\alpha \in \text{ON} \setminus \omega$. Then $\mathcal{S}(\alpha)$ is not a cardinal number.*

Proof. Define $f : \mathcal{S}(\alpha) \to \alpha$ by

$$f(\beta) = \begin{cases} \mathcal{S}(n) & \text{if } n < \omega; \\ \beta & \text{if } \omega \le \beta < \alpha; \\ 0 & \text{if } \beta = \alpha. \end{cases}$$

□

Let card (x) $(= |x|)$ be the least cardinal number κ such that $x \approx \kappa$, if it exists. Note that card (α) exists for all ordinals α. The following theorem implies that card (x) exists if x can be wellordered, i.e., if there is a wellordering of x.

Theorem 1.14. *Let r be a wellordering of x. Then there is an ordinal number α such that $\langle x, r\rangle$ is isomorphic to $\langle \alpha, \in\restriction\alpha\rangle$, i.e., there is a one-one onto $f : \alpha \to x$ such that*

$$\beta < \gamma < \alpha \to \langle f(\beta), f(\gamma)\rangle \in r.$$

Furthermore, both α and the isomorphism f are unique.

Proof. Note that α and f must satisfy

$$(\forall\beta < \alpha)\, f(\beta) \text{ is the } r\text{-least element of } x \setminus \text{range}\,(f \restriction \beta).$$

Define $F : V \to V$ as follows. Let $F(z)$ be the r-least element of $x \setminus \text{range}\,(z)$ if $(\exists\beta \in \text{ON})(\, z : \beta \to x \wedge \text{range}\,(z) \neq x)$, and let $F(z) = \emptyset$ otherwise. Let G be given by Theorem 1.8.

For each ordinal β, if range $(G \restriction \beta) \subsetneq x$ then $G(\beta) \in x \setminus \text{range}\,(G \restriction \beta)$.

Suppose that range $(G \restriction \beta) \subsetneq x$ for every ordinal β. Then $G : \text{ON} \to x$ and G is one-one. By Replacement (and Comprehension), we get that ON is a set. By Theorem 1.6, this implies that ON $\in$ ON, which contradicts Theorem 1.6.

Thus there is a $\beta \in$ ON such that range $(G\restriction\beta)$ is not a proper subset of x. Let α be the least such ordinal. If α is a limit ordinal, then range $(G\restriction\alpha) \subseteq x$ and so range $(G \restriction \alpha) = x$. This follows also if $\alpha = \mathcal{S}(\beta)$, since $G(\beta) \in x$. In both cases is it easy to see that $G \restriction \alpha$ is the desired isomorphism. □

For cardinal numbers κ and δ, we define the *cardinal sum* $\kappa + \delta$ of κ and δ by

$$\kappa + \delta = \text{card}\,(\{0\} \times \kappa) \cup (\{1\} \times \delta)),$$

if it exists. Our notation is ambiguous; we use the same symbol "+" both for the cardinal sum and for the *ordinal sum*, i.e., for the + operation on ordinal numbers defined on page 13. For the rest of this section, we shall avoid confusion by writing $\alpha +_{\mathrm{ON}} \beta$ for the ordinal sum of α and β.

Theorem 1.15. *(a) For all cardinal numbers κ and δ, $\kappa + \delta$ exists.*

(b) For m and $n \in \omega$, $m + n = m +_{\mathrm{ON}} n \in \omega$.

(c) If either of κ and δ does not belong to ω, then $\kappa + \delta = \max\{\kappa, \delta\}$ $(= \mathcal{U}(\{\kappa, \delta\}))$.

Proof. (a) Define an ordering $r_{\kappa,\delta}$ of $(\{0\} \times \kappa) \cup (\{1\} \times \delta)$ by placing $\langle i, \alpha\rangle$ before $\langle j, \beta\rangle$ if and only if

$$\alpha < \beta \vee (\alpha = \beta \wedge i < j).$$

It is easy to show that $r_{\kappa,\delta}$ is a wellordering. Let $f_{\kappa,\delta} : \alpha_{\kappa,\delta} \to (\{0\} \times \kappa) \cup (\{1\} \times \delta)$ be given by Theorem 1.14. Then $\kappa + \delta = \mathrm{card}\,(\alpha_{\kappa,\delta})$.

(b) For fixed $m \in \omega$, we prove by induction on n that $m +_{\mathrm{ON}} n \in \omega$ and $m +_{\mathrm{ON}} n \approx (\{0\} \times m) \cup (\{1\} \times n)$. By definition, $m +_{\mathrm{ON}} 0 = m \in \omega$, and we can define a one-one onto $f : m \to \{0\} \times m$ by setting $f(k) = \langle 0, k\rangle$ for each $k < m$. Assume that $m +_{\mathrm{ON}} n \in \omega$ and that $f : m +_{\mathrm{ON}} n \to (\{0\} \times m) \cup (\{1\} \times n)$ is one-one and onto. Then $m +_{\mathrm{ON}} \mathcal{S}(n) = \mathcal{S}(m +_{\mathrm{ON}} n) \in \omega$. Let

$$f' = f \cup \{\langle m +_{\mathrm{ON}} n, \langle 1, n\rangle\rangle\}.$$

It is easy to see that $f' : m +_{\mathrm{ON}} \mathcal{S}(n) \to (\{0\} \times m) \cup (\{1\} \times \mathcal{S}(n))$ is one-one and onto.

(c) It is enough to prove that $\kappa + \kappa = \kappa$ for every cardinal number $\kappa \notin \omega$. Assume that this is false, and let κ be the $<$-least counterexample. Note that $r_{\kappa,\kappa}$ is a wellordering of $2 \times \kappa$, where $2 = \{0, 1\}$. We have that

$$\kappa < \kappa + \kappa \leq \alpha_{\kappa,\kappa}.$$

Let $f_{\kappa,\kappa}(\kappa) = \langle i, \beta\rangle$. Thus

$$\kappa \approx \{\langle j, \gamma\rangle \mid \langle j, \gamma\rangle\, r_{\kappa,\kappa}\, \langle i, \beta\rangle\} \subseteq (2 \times \beta) \cup \{\langle 0, \beta\rangle\} \approx \mathcal{S}(\mathrm{card}\,(\beta) + \mathrm{card}\,(\beta)).$$

If $\beta \in \omega$, then we would also have $\kappa \in \omega$. Hence the minimality of κ gives that $\kappa \preceq \mathcal{S}(\mathrm{card}\,(\beta))$, and Theorems 1.11 and 1.13 then give the contradiction that $\kappa \approx \mathrm{card}\,(\beta)$. □

For cardinal numbers κ and δ, we define the *cardinal product* $\kappa \cdot \delta$ of κ and δ by

$$\kappa \cdot \delta = \mathrm{card}\,(\kappa \times \delta),$$

if it exists. Our notation is once more ambiguous, so for the rest of this section we shall write $\cdot_{\mathrm{ON}}$ for the *ordinal product* defined on page 13.

Theorem 1.16. *(a) For all cardinal numbers κ and δ, $\kappa \cdot \delta$ exists.*

(b) For m and $n \in \omega$, $m \cdot n = m \cdot_{\mathrm{ON}} n \in \omega$.

(c) If either of κ and δ does not belong to ω and neither of κ and δ is 0, then $\kappa \cdot \delta = \max\{\kappa, \delta\}$.

Exercise 1.7. Prove Theorem 1.16.

Hint: (a) Define an ordering $s_{\kappa,\delta}$ of $\kappa \times \delta$ as follows:

$$\langle \alpha, \beta \rangle \, s_{\kappa,\delta} \, \langle \alpha', \beta' \rangle \leftrightarrow \begin{cases} \max\{\alpha, \beta\} < \max\{\alpha', \beta'\} \ \vee \\ \max\{\alpha, \beta\} = \max\{\alpha', \beta'\} \wedge \alpha < \alpha' \ \vee \\ \max\{\alpha, \beta\} = \max\{\alpha', \beta'\} \wedge \alpha = \alpha' \wedge \beta < \beta'. \end{cases}$$

Show that $s_{\kappa,\delta}$ is a wellordering. Let $f^*_{\kappa,\delta} : \alpha^*_{\kappa,\delta} \to \kappa \times \delta$ be given by Theorem 1.14. Then $\kappa \cdot \delta = \mathrm{card}\,(\alpha^*_{\kappa,\delta})$.

(b) For fixed $m \in \omega$, prove by induction that, for all $n \in \omega$, $m \cdot_{\mathrm{ON}} n = m \cdot n \in \omega$. The case $n = 0$ is trivial. Assume that $m \cdot_{\mathrm{ON}} n = m \cdot n \in \omega$. Then

$$m \cdot_{\mathrm{ON}} \mathcal{S}(n) = m \cdot_{\mathrm{ON}} n +_{\mathrm{ON}} m = m \cdot n + m \in \omega,$$

where the last equality is by the induction hypothesis and Theorem 1.15. Show that $m \cdot n + m = m \approx m \times \mathcal{S}(n)$.

(c) It is enough to prove that $\kappa \cdot \kappa = \kappa$ for every cardinal number $\kappa \notin \omega$. Assume that this is false, and let κ be the $<$-least counterexample. Let $f^*_{\kappa,\kappa} : \alpha^*_{\kappa,\kappa} \to \kappa \times \kappa$ be defined as in the hint for part (a). Then

$$\kappa < \kappa \cdot \kappa \leq \alpha^*_{\kappa,\kappa}.$$

Let $\langle \alpha, \beta \rangle = f^*_{\kappa,\kappa}(\kappa)$. Let $\rho = \max\{\alpha, \beta\}$. Use the definition of $s_{\kappa,\kappa}$, the minimality of κ, and Theorem 1.15 to deduce the contradiction that $\kappa \approx \mathrm{card}\,(\rho) \leq \rho < \kappa$.

For sets x and y, let ${}^x y = \{f \mid f : x \to y\}$. (Note that ${}^x y$ is contained in the set $\mathcal{P}(x \times y)$.) Since we do not have a convenient special notation for the *ordinal exponentiation* defined on page 14, we defer defining cardinal exponentiation until after the next theorem, which concerns ordinal exponentiation.

Theorem 1.17. *For m and $n \in \omega$, ${}^m n \approx n^m \in \omega$, where n^m is as defined on page 14.*

Proof. Fix $n \in \omega$. For the case $m = 0$, note that ${}^0 n = \{\emptyset\} = 1 = n^0$. Assume that $n^m \in \omega$ and that $n^m \approx {}^m n$. Then $n^{\mathcal{S}(m)} = n^m \cdot_{\mathrm{ON}} n \in \omega$. Moreover

$$n^m \cdot_{\mathrm{ON}} n = n^m \cdot n \approx n^m \times n \approx {}^m n \times n \approx {}^{\mathcal{S}(m)} n.$$

(For the last $\approx$, define a one-one onto f by setting $f(\langle g, k\rangle) = g \cup \{\langle m, k\rangle\}$ for $g : m \to n$ and $k < n$.) □

We now define *cardinal exponentiation* by setting $\kappa^\lambda = \mathrm{card}\,({}^\lambda \kappa)$, if it exists, for cardinal numbers κ and λ. We shall make no more use of ordinal exponentiation in this section.

Theorem 1.18. *If $0 \neq n \in \omega$ and $\kappa \notin \omega$ is a cardinal number, then $\kappa^n = \kappa$.*

Proof. Fix a cardinal number $\kappa \notin \omega$. For $n \in \omega$, define $f_n : {}^{\mathcal{S}(n)}\kappa \to {}^n\kappa \times \kappa$ by setting $f_n(g) = \langle g \restriction n, g(n)\rangle$. The functions f_n are one-one and onto.

Clearly ${}^1\kappa \approx \kappa$. Assume that $n > 0$ and that ${}^n\kappa \approx \kappa$. Then

$${}^{\mathcal{S}(n)}\kappa \approx {}^n\kappa \times \kappa \approx \kappa \times \kappa \approx \kappa.$$ □

For ordinal numbers α and sets y, let ${}^{<\alpha}y = \{f \mid (\exists \beta < \alpha)\, f : \beta \to y\}$. For cardinal numbers κ and λ, let $\kappa^{<\lambda} = \mathrm{card}\,({}^{<\lambda}\kappa)$, if it exists.

Theorem 1.19. *If $\kappa \notin \omega$ is a cardinal number, then $\kappa^{<\omega} = \kappa$.*

Proof. The theorem is an easy consequence of Theorem 1.18 and the Axiom of Choice, but we wish to avoid the latter. Let f_n be as in the proof of Theorem 1.18. Let $h : \kappa \times \kappa \to \kappa$ be one-one and onto.

Define $g_n : {}^{\mathcal{S}(n)}\kappa \to \kappa$ and $g_n^* : {}^{\mathcal{S}(n)}\kappa \times \kappa \to \kappa \times \kappa$ simultaneously by recursion as follows. Let g_0 be given by h. Given g_n, let

$$g_n^*(\langle q, \alpha\rangle) = \langle g_n(q), \alpha\rangle.$$

Now let

$$g_{\mathcal{S}(n)} = h \circ g_n^* \circ f_{\mathcal{S}(n)},$$

where $\circ$ means composition. (It is easy to justify this method of definition via Theorem 1.4.) By induction we see that each g_n is one-one and onto.

Next define a one-one $p : \omega \times \kappa \to {}^{<\omega}\kappa$ by setting $p(n, \alpha) = {g_n}^{-1}(\alpha)$. (Here we write $p(n, \alpha)$ for $p(\langle n, \alpha\rangle)$.) Since ${}^{<\omega}\kappa = \mathrm{range}\,(p) \cup \{1\}$, we get that ${}^{<\omega}\kappa \approx (\omega \times \kappa) \cup \{1\} \approx \kappa$. □

Theorem 1.20. *For every set x, $x \prec {}^x2$, i.e., $x \preceq {}^x2$ and $x \not\approx {}^x2$.*

Proof. Fix x. It is easy to see that ${}^x2 \approx \mathcal{P}(x)$. We show that $x \prec \mathcal{P}(x)$.

To show that $x \preceq \mathcal{P}(x)$ define a one-one $f : x \to \mathcal{P}(x)$ by setting $f(y) = \{y\}$ for all $y \in x$.

Suppose that $f : x \to \mathcal{P}(x)$ is onto. Let $z = \{y \in x \mid y \notin f(y)\}$. Let $z = f(y)$. Then $y \in f(y) \Leftrightarrow y \notin z \Leftrightarrow y \notin f(y)$. □

Theorem 1.21. *There is no greatest cardinal number.*

Proof. Let κ be a cardinal number. Let

$$a = \{\langle x, r\rangle \mid x \subseteq \kappa \wedge r \text{ is a wellordering of } x\}.$$

For $\langle x, r\rangle \in a$, let $g(\langle x, r\rangle)$ be the unique α such that $\langle \alpha, \in\restriction\alpha\rangle$ is isomorphic to $\langle x, r\rangle$. If α is an ordinal number and $\alpha \preceq \kappa$, then there is an $\langle x, r\rangle \in a$ with $\alpha = g(\langle x, r\rangle)$. (Let $f : \alpha \to \kappa$ be one-one; let $x = \text{range}(f)$; let $\langle f(\beta), f(\gamma)\rangle \in r \Leftrightarrow \beta < \gamma$.) Let $\delta = \mathcal{U}(\text{range}(g))$. Then $\delta \in \text{ON}$ and $\kappa \prec \delta$. Indeed, δ is the least cardinal number $> \kappa$. □

For any set x such that $\text{card}\,(x)$ exists, let x^+ be the least cardinal number greater than $\text{card}\,(x)$.

By transfinite recursion define

$$\begin{aligned} \aleph_0 &= \omega\,; \\ \aleph_{\mathcal{S}(\alpha)} &= {\aleph_\alpha}^+\,; \\ \aleph_\lambda &= \bigcup\{\aleph_\beta \mid \beta < \lambda\} \text{ for limit ordinals } \lambda. \end{aligned}$$

It is easy to see that the $\aleph_\alpha$, $\alpha \in \text{ON}$, are all the cardinal numbers $\geq \omega$.

Theorem 1.22. (Uses Choice) *Every set can be wellordered.*

Proof. Fix a set x. For $y \subsetneq x$, let $a_y = \{y\} \times (x \setminus y)$. Let $u = \{a_y \mid y \subsetneq x\}$. Let v be given by Choice. Define $F : V \to V$ as follows. Let $F(z)$ be the unique w such that $\langle \text{range}\,(z), w\rangle \in v$ if $(\exists\beta \in \text{ON})(\, z : \beta \to x \wedge \text{range}\,(z) \neq x)$, and let $F(z) = \emptyset$ otherwise. Let G be given by transfinite recursion. Just as in the proof of Theorem 1.14, one can show that there is an ordinal α such that $G \restriction \alpha$ is a one-one onto function from α to x. □

Corollary 1.23. (Uses Choice) *For every set x, $\text{card}\,(x)$ exists. For all cardinals κ and λ, both κ^λ and $\kappa^{<\lambda}$ are defined.*

By Theorem 1.20, we have that $2^{\aleph_\alpha} > \aleph_\alpha$ for every ordinal α. The *Continuum Hypotheses* (CH) asserts that $2^{\aleph_0} = \aleph_1$, and the *Generalized Continuum Hypothesis* (GCH) asserts that $2^{\aleph_\alpha} = \aleph_{\mathcal{S}(\alpha)}$ for all ordinals α.

2 Models, compactness, and completeness

Informally we shall consider a *language* to be a set of symbols, the union of the following:

(1) a set of *constant symbols*;

(2) for each n, $0 < n \in \omega$, a set of *n-place function symbols*;

(3) for each n, $0 < n \in \omega$, a set of *n-place relation symbols*.

Since we want to use theorems of set theory in doing model theory (and for other reasons concerning 220C), we adopt the following purely set theoretic definition as our official one.

A *language* is a pair $\langle f, p\rangle$ where

(a) $f : \omega \to V$;

(b) $p : \omega \setminus \{0\} \to V$;

(c) $(\forall m \in \omega)(\forall n \in \omega)\ (m \neq n \to f(m) \cap f(n) = \emptyset)$;

(d) $(\forall m \in \omega \setminus \{0\})(\forall n \in \omega \setminus \{0\})\ (m \neq n \to p(m) \cap p(n) = \emptyset)$;

(e) $(\forall m \in \omega)(\forall n \in \omega \setminus \{0\})\ f(m) \cap p(n) = \emptyset$;

(f) each $f(n)$ and each $p(n)$ is disjoint from $\{2{\cdot}n \mid n \in \omega\} \cup \{1, 3, 5, 7, 9, 11\}$;

(g) no function whose domain is in $\omega \setminus \{\emptyset\}$
belongs to any $f(n)$ or $p(n)$.

If $\mathcal{L} = \langle f, p\rangle$, then $f(0)$ is the set of *constant symbols* of $\mathcal{L}$; for $n > 0$, $f(n)$ is the set of *n-place function symbols* of $\mathcal{L}$; for $n > 0$, $p(n)$ is the set of *n-place relation symbols* of $\mathcal{L}$. Clauses (c)–(e) say that nothing has two uses as a symbol of $\mathcal{L}$. Clause (f) says that no symbol of $\mathcal{L}$ is also one of the *logical symbols* specified below. Clause (g), as we shall explain later, avoids still another kind of double use of a symbol of $\mathcal{L}$. This will be explained later.

Logical symbols. The following symbols will be used with every language:

Informal	Official
$v_0, v_1, v_2, \ldots$	$0, 2, 4, \ldots$
(	1
)	3
$=$	5
$\neg$	7
$\wedge$	9
$\exists$	11

The symbols $v_0, v_1, v_2, \ldots$ (officially $0, 2, 4, \ldots$) are *variables*.

Terms. Informally we can describe the *terms* of a language $\mathcal{L}$ as constituting the smallest set such that

(i) all variables and constant symbols are terms;

(ii) if F is an n-place function symbol and $t_1, \ldots, t_n$ are terms, then the expression $F(t_1 \ldots t_n)$ is a term.

More informally, we shall often add commas for clarity: $F(t_1, \ldots, t_n)$.

Officially terms of $\mathcal{L}$ are finite sequences of symbols, where a *finite sequence* is a function whose domain is a natural number. To give the official set-theoretic definition we first define some operations on finite sequences.

If $g : m \to V$ and $h : n \to V$ are finite sequences, let $g^\frown h : m + n \to V$ be given by

$$(g^\frown h)(k) = \begin{cases} g(k) & \text{if } k < m\,; \\ h(j) & \text{if } k = m + j \text{ with } j < n\,. \end{cases}$$

For finite sequences h of finite sequences, we define concat (h), the *concatenation* of h, by recursion on domain (h) as follows:[1]

$$\text{concat}\,(h) = \begin{cases} \emptyset & \text{if domain}\,(h) = 0\,; \\ (\text{concat}\,(h \restriction n))^\frown h(n) & \text{if domain}\,(h) = n + 1\,. \end{cases}$$

For finite sequences f, let $\ell\text{h}(f) = \text{domain}\,(f)$. For any a, let $\langle a \rangle$ be the unique element of ${}^1\{a\}$, i.e., let it be $\{\langle 0, a \rangle\}$.

Now let

$$\text{Term}_0^{\mathcal{L}} = \{\langle a \rangle \mid a \text{ is a variable or a constant symbol}\}\,.$$

For $n \in \omega$, let $\text{Term}_{n+1}^{\mathcal{L}}$ be the set of all concat (h) such that, for some $k \in \omega \setminus \{0\}$,

(a) $h : k + 3 \to V$;

(b) $h(0) \in {}^1(f(k))$, where $\mathcal{L} = \langle f, p \rangle$;

(c) $h(1) = \langle (\rangle$ (i.e., $h(1) = \langle 1 \rangle$);

[1] The definition of the class function concat is more complicated than the phrase "by recursion" suggests. We cannot define by recursion a function $q : \omega \to V$ such that each $q(n)$ is the restriction of concat to those h whose domain is n. For $n > 1$, the desired $q(n)$ is a proper class. Instead, we define by recursion q_α for each fixed $\alpha \in \text{ON}$, where $q_\alpha(n)$ is the restriction of concat to those h with domain n and range $\subseteq V_\alpha$. Then we define concat to be $\bigcup(\{q_\alpha(n) \mid n \in \omega \wedge \alpha \in \text{ON}\})$.

(d) $h(k+2) = \langle) \rangle$;

(e) $(\forall j < k)\, h(2+j) \in \bigcup\{\text{Term}_m^{\mathcal{L}} \mid m \leq n\}$.

A *term* of $\mathcal{L}$ is any member of $\bigcup\{\text{Term}_n^{\mathcal{L}} \mid n \in \omega\}$.

Exercise 2.1. (a) Prove unique readability for terms. That is, show that if t is a term of a language $\mathcal{L}$ not belonging to $\text{Term}_0^{\mathcal{L}}$, then there are unique $k \in \omega$ and $h : k+3 \to V$ such that $t = \text{concat}\,(h)$ and (a)–(e) above hold of k and h, with (e) modified by replacing "$m \leq n$" by "$m \in \omega$." You may (informally) prove the informal version of this fact.

(b) Would unique readability for terms still hold if we dropped the parentheses? Prove your answer.

Formulas. Informally we can describe the *formulas* of $\mathcal{L}$ as forming the smallest set satisfying the conditions

(i) if t_1 and t_2 are terms, then $t_1 = t_2$ is a formula;

(ii) if P is a k-place relation symbol and $t_1, \ldots, t_k$ are terms, then $P(t_1 \ldots t_k)$ is a formula;

(iii) if φ is a formula, then so is $\neg\varphi$;

(iv) if φ and ψ are formulas, then so is $(\varphi \wedge \psi)$;

(v) if φ is a formula and x is a variable, then $(\exists x)\, \varphi$ is a formula.

Officially we take formulas, like terms, to be finite sequences of symbols. We let $\text{Formula}_0^{\mathcal{L}}$ be the set of all *atomic formulas*, i.e., the set of all finite sequences corresponding to clauses (i) and (ii) above. For $n \in \omega$, we let $\text{Formula}_{n+1}^{\mathcal{L}}$ be the set of all the sequences gotten from $\bigcup\{\text{Formula}_m^{\mathcal{L}} \mid m \leq n\}$ via clauses (iii), (iv) and (v). We omit the official definition, which is similar to that of the sets Term_n.

Exercise 2.2. (a) Prove unique readability for formulas. That is, show that every formula either is atomic or else has a unique analysis via (iii), (iv), or (v).

(b) Would unique readability for formulas still hold if we dropped the parentheses? Prove your answer.

Officially let us define an *occurrence* of a variable x in a formula φ to be $\langle m, \varphi \rangle$ for any $m < \ell\text{h}(\varphi)$ such that $\varphi(m) = x$. Similarly define the notion of an *occurrence* of a variable in a term.

By the *complexity* of a formula φ, we mean the least n such that $\varphi \in \mathrm{Formula}_n^{\mathcal{L}}$. By recursion on complexity of formulas, we define the *free* occurrences of a variable in a formula. Every occurrence of a variable in an atomic formula is *free*. An occurrence $\langle m+1, \neg\varphi\rangle$ is *free* just in case the corresponding occurrence $\langle m, \varphi\rangle$ is free. An occurrence $\langle m+1, (\varphi \wedge \psi)\rangle$ with $m < \ell\mathrm{h}(\varphi)$ is *free* just in case $\langle m, \varphi\rangle$ is free. An occurrence $\langle \ell\mathrm{h}(\varphi) + m + 2, (\varphi \wedge \psi)\rangle$ is *free* just in case $\langle m, \psi\rangle$ is free. An occurrence $\langle 2, (\exists x)\varphi\rangle$ is not free. An occurrence $\langle m + 4, (\exists y)\varphi\rangle$ of x is free just in case $\langle m, \varphi\rangle$ is free and x and y are different variables.

Models. A *model* $\mathfrak{A}$ for a language $\mathcal{L}$ is an ordered pair consisting of (a) a non-empty set $A = |\mathfrak{A}|$, the *universe* or *domain* of the model, and (b) a function assigning

(1) to each constant symbol c, an element $c_{\mathfrak{A}}$ of A;

(2) to each k-place function symbol F, a function $F_{\mathfrak{A}} : {}^kA \to A$;

(3) to each k-place relation symbol P, a subset $P_{\mathfrak{A}}$ of ${}^kA$.

As a convention, when we denote a model by a Fraktur letter, then we denote the universe of the model by the corresponding italic Roman letter.

In order to define the notions of *satisfaction* and *truth*, let us fix a language $\mathcal{L}$ and a model $\mathfrak{A}$ for $\mathcal{L}$.

The *complexity* of a term t is the least n such that $t \in \mathrm{Term}_n^{\mathcal{L}}$. For a term t and for $s \in {}^{<\omega}A$ such that all variables occurring in t belong to $\{v_i \mid i < \ell\mathrm{h}(s)\}$, we define, by recursion on the complexity of t, an element $t_{\mathfrak{A}}^s$ of A:

$$\begin{aligned} c_{\mathfrak{A}}^s &= c_{\mathfrak{A}} \quad \text{for } c \text{ a constant;} \\ {v_i}_{\mathfrak{A}}^s &= s(i)\,; \\ (F(t_1 \ldots t_n))_{\mathfrak{A}}^s &= F_{\mathfrak{A}}({t_1}_{\mathfrak{A}}^s, \ldots, {t_n}_{\mathfrak{A}}^s)\,, \end{aligned}$$

where "$F_{\mathfrak{A}}({t_1}_{\mathfrak{A}}^s, \ldots, {t_n}_{\mathfrak{A}}^s)$" is an abbreviation for "$F_{\mathfrak{A}}(q)$, where $q : n \to A$ and $q(i) = {t_{i+1}}_{\mathfrak{A}}^s$ for all $i < n$." Note that $t_{\mathfrak{A}}^s$ is independent of s if no variables occur in t.

Satisfaction. We define, by recursion, for each $n \in \omega$ a relation

$$\mathrm{Sat}_n^{\mathfrak{A}} \subseteq \mathrm{Formula}_n^{\mathcal{L}} \times {}^{<\omega}A.$$

If $\langle \varphi, s\rangle \in \mathrm{Sat}_n^{\mathfrak{A}}$, then the variables having free occurrences in φ must be among $\{v_i \mid i < \ell\mathrm{h}(s)\}$. Also φ must of course belong to $\mathrm{Formula}_n^{\mathcal{L}}$. We shall omit mentioning these two requirements below.

(i) $\langle t_1 = t_2, s\rangle \in \mathrm{Sat}_0^{\mathfrak{A}} \leftrightarrow {t_1}^s_{\mathfrak{A}} = {t_2}^s_{\mathfrak{A}}$.

(ii) $\langle P(t_1 \dots t_k), s\rangle \in \mathrm{Sat}_0^{\mathfrak{A}} \leftrightarrow q \in P_{\mathfrak{A}}$, where $q : k \to A$ and $q(i) = {t_{i+1}}^s_{\mathfrak{A}}$ for each $i < k$.

(iii) $\langle \neg\varphi, s\rangle \in \mathrm{Sat}_{n+1}^{\mathfrak{A}} \leftrightarrow \langle \varphi, s\rangle \notin \bigcup\{\mathrm{Sat}_m^{\mathfrak{A}} \mid m \le n\}$.

(iv) $\langle (\varphi \wedge \psi), s\rangle \in \mathrm{Sat}_{n+1}^{\mathfrak{A}} \leftrightarrow (\langle \varphi, s\rangle \in \bigcup\{\mathrm{Sat}_m^{\mathfrak{A}} \mid m \le n\} \wedge \langle \psi, s\rangle \in \bigcup\{\mathrm{Sat}_m^{\mathfrak{A}} \mid m \le n\})$.

(v) $\langle (\exists v_j)\,\varphi, s\rangle \in \mathrm{Sat}_{n+1}^{\mathfrak{A}} \leftrightarrow (\exists s')(s' \supseteq s{\restriction}\mathrm{domain}\,(s)\backslash\{j\} \wedge j \in \mathrm{domain}\,(s') \wedge \langle \varphi, s'\rangle \in \bigcup\{\mathrm{Sat}_m^{\mathfrak{A}} \mid m \le n\})$.

We let $\mathrm{Sat}^{\mathfrak{A}} = \bigcup\{\mathrm{Sat}_n^{\mathfrak{A}} \mid n \in \omega\}$. We say that $\mathfrak{A}$ *satisfies* $\varphi[s]$ (in symbols, $\mathfrak{A} \models \varphi[s]$) if $\langle \varphi, s\rangle \in \mathrm{Sat}^{\mathfrak{A}}$. If only $v_{i_1}, \dots, v_{i_n}$ have free occurrences in φ, then we may indicate this by writing $\varphi(v_{i_1}, \dots, v_{i_n})$ for φ. Moreover we write $\mathfrak{A} \models \varphi[a_1, \dots, a_n]$ to mean that, for some (or equivalently, every) s such that $s(i_j) = a_j$ for each j, $\mathfrak{A} \models \varphi[s]$.

If a term t contains no variables, then we write $t_{\mathfrak{A}}$ for $t^s_{\mathfrak{A}}$. If a formula σ has no free occurrences of variables (σ is a *sentence*), then we write $\mathfrak{A} \models \sigma$ for $\mathfrak{A} \models \sigma[s]$. If σ is a sentence and $\mathfrak{A} \models \sigma$ then we say that $\mathfrak{A}$ is a *model of* σ and that σ is *true* in $\mathfrak{A}$. If Σ is a set of sentences then we define

$$\mathfrak{A} \textit{ satisfies } \Sigma \;\leftrightarrow\; \mathfrak{A} \models \Sigma \;\leftrightarrow\; \mathfrak{A} \text{ is a } \textit{model} \text{ of } \Sigma \;\leftrightarrow\; (\forall \sigma \in \Sigma)\, \mathfrak{A} \models \sigma.$$

Exercise 2.3. Theorem 1.4 shows that the definition above of $\mathrm{Sat}^{\mathfrak{A}}$ yields an explicit definition of $\mathrm{Sat}^{\mathfrak{A}}$ from the parameter $\mathfrak{A}$ and so gives us a proper class function $\mathfrak{A} \mapsto \mathrm{Sat}^{\mathfrak{A}}$. Consider the language $\mathcal{L}$ of set theory, which (informally) is the set {"∈"}. Think of V as giving a "model" $\mathfrak{V}$ with $|\mathfrak{V}| = V$ and with "∈"$_{\mathfrak{V}} = \in$. Can Theorem 1.4 be used define, via clauses like (i)–(v) above, a proper class $\mathrm{Sat}^{\mathfrak{V}} \subseteq \mathrm{Formula}^{\mathcal{L}} \times {}^{<\omega}V$? Explain.

A sentence or a set of sentences of a language $\mathcal{L}$ is *valid* in $\mathcal{L}$ if every model $\mathfrak{A}$ for $\mathcal{L}$ satisfies it. A sentence or a set of sentences of $\mathcal{L}$ is *consistent* (*satisfiable*) in $\mathcal{L}$ if some model $\mathfrak{A}$ for $\mathcal{L}$ satisfies it. It is easy to see by induction that validity and consistency in $\mathcal{L}$ of a sentence σ or set Σ of sentences is independent of $\mathcal{L}$ (for $\mathcal{L}$ containing all symbols in σ or Σ respectively), so we shall usually omit "in $\mathcal{L}$." A sentence σ *logically implies* a sentence τ in $\mathcal{L}$ (in symbols, $\sigma \models_{\mathcal{L}} \tau$) if every model for $\mathcal{L}$ that is a model of σ is a model of τ. Similarly define Σ *logically implies* τ in $\mathcal{L}$ ($\Sigma \models_{\mathcal{L}} \tau$) for sets Σ of sentences and sentences τ. It is easy to see that $\sigma \models_{\mathcal{L}} \tau$ and $\Sigma \models_{\mathcal{L}} \tau$ are independent of $\mathcal{L}$, so we shall usually omit the subscript "$\mathcal{L}$" and the phrase "in $\mathcal{L}$."

A set Σ of sentences has *Henkin witnesses* if whenever $(\exists x)\,\varphi(x) \in \Sigma$ then there is a constant symbol c such that $\varphi(c) \in \Sigma$, where $\varphi(c)$ is the result of substituting c for the free occurrences of x in $\varphi(x)$.

Theorem 2.1 (Henkin Models). (Uses Choice) *Let Σ be a set of sentences of a language $\mathcal{L}$. Suppose that*

(1) every finite subset of Σ is consistent in $\mathcal{L}$;

(2) Σ has Henkin witnesses;

(3) for each sentence σ of $\mathcal{L}$, either $\sigma \in \Sigma$ or $\neg\sigma \in \Sigma$.

Then Σ has a model $\mathfrak{A}$ such that $\mathrm{card}\,(\mathfrak{A}) \leq$ *the cardinal number of the set of constant symbols of $\mathcal{L}$, where we mean by "*$\mathrm{card}\,(\mathfrak{A})$*" not the literal* $\mathrm{card}\,(\mathfrak{A})$ *(namely 2) but rather* $\mathrm{card}\,(A)$*.*

(The model $\mathfrak{A}$ will be constructed without using Choice. We need Choice to guarantee that the set of all constant symbols of $\mathcal{L}$ has a cardinal number.)

We call a set x *finite* (e.g., in hypothesis (1)), if $\mathrm{card}\,(x) \in \omega$.

Proof. In preparation for the proof of the Completeness Theorem, we shall explicitly record all facts about logical implication needed for the proofs of Theorem 2.1 and Theorem 2.8. (We shall later see that all these facts correspond to facts about a proof-theoretic notion of implication.)

Note that

$$\Delta \text{ consistent } \leftrightarrow \neg(\exists\tau)(\Delta \models \tau \,\wedge\, \Delta \models \neg\tau).$$

For the purpose of listing facts about $\models$, let us take this as the *definition* of consistency.

$$\text{(I)} \qquad \{\tau\} \models \tau$$

$$\text{(II)} \qquad (\Delta_1 \models \tau \,\wedge\, \Delta_1 \subseteq \Delta_2) \,\rightarrow\, \Delta_2 \models \tau$$

Lemma 2.2. *Assume that $\Delta \subseteq \Sigma$ is finite and such that $\Delta \models \tau$. Then $\tau \in \Sigma$.*

Proof. Otherwise hypothesis (3) gives that $\neg\tau \in \Sigma$. By (I) and (II),

$$\Delta \cup \{\neg\tau\} \models \neg\tau \;\wedge\; \Delta \cup \{\neg\tau\} \models \tau.$$

This contradicts hypothesis (1). □

Let us call a formula φ *prime* if φ is either atomic or of the form $(\exists x)\,\psi$. The formulas of $\mathcal{L}$ constitute the smallest set containing the prime formulas of $\mathcal{L}$ and closed under the operations $\varphi \mapsto \neg\varphi$ and $\langle\varphi,\psi\rangle \mapsto (\varphi \wedge \psi)$. This gives rise to a variant notion of complexity of formulas, with respect to which we may use induction and definition by recursion.

A *valuation* for $\mathcal{L}$ is a function v from the set of prime formulas of $\mathcal{L}$ to $\{0,1\}$. Given any valuation v for $\mathcal{L}$ we can define by recursion a canonical $v^* : \text{Formula}^{\mathcal{L}} \to \{0,1\}$ such that v^* extends v:

$$\begin{aligned} v^*(\varphi) &= v(\varphi) \text{ for } \varphi \text{ prime;} \\ v^*(\neg\varphi) &= 1 - v^*(\varphi)\,; \\ v^*((\varphi \wedge \psi)) &= \min\{v^*(\varphi), v^*(\psi)\}\,. \end{aligned}$$

(For $n \leq m \in \omega$, $m - n$ is the k such that $n + k = m$. It is easy to show the existence and uniqueness of such a k.)

A formula φ of $\mathcal{L}$ is *true under* a valuation v if $v^*(\varphi) = 1$. We say that a set Φ of formulas of $\mathcal{L}$ *truth-functionally implies* in $\mathcal{L}$ a formula φ of $\mathcal{L}$ if, for every valuation v for $\mathcal{L}$, if each member of Φ is true under v then φ is true under v. A *tautology* of $\mathcal{L}$ is a formula true under every valuation for $\mathcal{L}$. It is easy to show by induction that truth-functional implication and being a tautology are, in the natural sense, independent of $\mathcal{L}$, so we shall usually omit "in $\mathcal{L}$" and "of $\mathcal{L}$." We write $\Phi \models_{\text{tf}} \varphi$ to mean that Φ truth-functionally implies φ.

Lemma 2.3. *Suppose that Δ is a set of sentences of $\mathcal{L}$ and that τ is a sentence of $\mathcal{L}$. If $\Delta \models_{\text{tf}} \tau$ then $\Delta \models \tau$.*

Proof. Suppose that $\mathfrak{A}$ is a model for $\mathcal{L}$ such that $\mathfrak{A} \models \Delta$ but $\mathfrak{A} \not\models \tau$. Define a valuation v for $\mathcal{L}$ as follows:

$$v(\varphi) = \begin{cases} 0 & \text{if } \varphi \text{ is not a sentence;} \\ 0 & \text{if } \varphi \text{ is a sentence and } \mathfrak{A} \not\models \varphi\,; \\ 1 & \text{if } \varphi \text{ is a sentence and } \mathfrak{A} \models \varphi\,. \end{cases}$$

It is easy to prove by induction on complexity that, for any sentence σ of $\mathcal{L}$, σ is true under v if and only if $\mathfrak{A} \models \sigma$. Hence v witnesses that $\Delta \not\models_{\text{tf}} \tau$. □

The next fact in our list is a weakening of Lemma 2.3.

(III) $$(\Delta \text{ finite } \wedge\ \Delta \models_{\text{tf}} \tau) \to \Delta \models \tau$$

The reason for not taking the full lemma as (III) will be explained later.

Let us write $\models \sigma$ to mean that $\emptyset \models \sigma$, i.e., that σ is valid.

For constants (constant symbols) c_1 and c_2 of $\mathcal{L}$, set

$$c_1 \sim c_2 \;\leftrightarrow\; c_1 = c_2 \in \Sigma\,.$$

Lemma 2.4. $\sim$ *is an equivalence relation.*

Proof. Note that

(IV) $$\models c = c \quad \text{for } c \text{ a constant.}$$

By Lemma 2.2, this gives $c \sim c$.

Assume that $c_1 \sim c_2$.

(V) $$\models (t_1 = t_2 \to (\varphi(t_1) \to \varphi(t_2)))$$ $$\text{for } \varphi(x) \text{ atomic, } t_1 \text{ and } t_2 \text{ terms without variables}$$

Here $\varphi(t_i)$ is the result of replacing the free occurrences of x in $\varphi(x)$ by occurrences of t_i. Here also we make use of the abbreviation "$\rightarrow$." (See page 2.)

With $x = c_1$ for $\varphi(x)$, we get from (V) that

$$\models \;(c_1 = c_2 \;\to\; (c_1 = c_1 \to\; c_2 = c_1))\,.$$

Lemma 2.2 then implies that this sentence belongs to Σ. Now one readily checks that $\{\sigma, (\sigma \to \tau)\} \models_{\mathrm{tf}} \tau$ for any σ and τ. By (III) and two applications of Lemma 2.2, we get that $c_2 = c_1 \in \Sigma$ and so that $c_2 \sim c_1$.

Assume that $c_1 \sim c_2$ and $c_2 \sim c_3$. Applying (V) with $x = c_3$ for $\varphi(x)$, we get that

$$\models \;(c_2 = c_1 \;\to\; (c_2 = c_3 \to c_1 = c_3))\,.$$

Since $c_2 = c_1 \in \Sigma$ and $c_2 = c_3 \in \Sigma$, it follows by (III) and Lemma 2.2 that $c_1 = c_3 \in \Sigma$ and so that $c_1 \sim c_3$. □

For constants c of $\mathcal{L}$, let $[c] = \{c' \mid c' \sim c\}$. Let

$$A = \{[c] \mid c \text{ is a constant of } \mathcal{L}\}\,.$$

(VI) $$\models (\exists v_1)\, v_1 = v_1$$

Lemma 2.5. *The set A is non-empty.*

Proof. By (VI) and Lemma 2.2, the sentence $(\exists v_1)\, v_1 = v_1$ belongs to Σ. Hypothesis (2) yields a constant c of $\mathcal{L}$ such that $c = c \in \Sigma$. Hence there is a constant of $\mathcal{L}$. $\square$

Define $c_{\mathfrak{A}} = [c]$ for each constant c of $\mathcal{L}$.

(VII) $\models (\exists x)\, F(c_1 \ldots c_k) = x$
for F a k-place function symbol
and $c_1, \ldots, c_k$ constants

For F and $c_1, \ldots, c_k$ as in (VII), we get by (VII), Lemma 2.2, and hypothesis (2) that there is a constant c with $F(c_1 \ldots c_k) = c \in \Sigma$. Define

$$F_{\mathfrak{A}}([c_1], \ldots, [c_k]) = [c]\,.$$

Here and hereafter we use the following notational convention: $a_1, \ldots, a_k$ denotes the sequence q of length k such that $q(i) = a_{i+1}$ for each $i < k$. We must show that this does not depend on the representatives $c_1, \ldots, c_k$ and on the choice of c.

(VIII) $\models (t_1 = t_2 \rightarrow u(t_1) = u(t_2))$
for $u(x)$ a term, t_1 and t_2 terms without variables

Suppose that $F(c_1 \ldots c_k) = c$ and $F(c'_1 \ldots c'_k) = c'$ both belong to Σ and that $c_i \sim c'_i$ for $1 \leq i \leq k$. For $1 \leq j \leq k+1$, let t_j be the term

$$F(c'_1 \ldots c'_{j-1} c_j \ldots c_k)\,.$$

(VIII) and (III) give us that $t_j = t_{j+1}$ belongs to Σ for $1 \leq j \leq k$. Let $0 \leq i < k$ and assume that $t_{k+1-i} = t_{k+1} \in \Sigma$. By (V),

$$\models (t_{k+1-i} = t_{k+1} \rightarrow (t_{k+1-(i+1)} = t_{k+1-i} \rightarrow t_{k+1-(i+1)} = t_{k+1}))\,.$$

(III) and Lemma 2.2 then give that $t_{k+1-(i+1)} = t_{k+1} \in \Sigma$. By induction we get that $t_1 = t_{k+1} \in \Sigma$, that is, $F(c_1 \ldots c_k) = F(c'_1 \ldots c'_k)$ belongs to Σ. (V) and (III) give that $F(c'_1 \ldots c'_k) = c$ belongs to Σ; (V) and (III) again give that $c = c' \in \Sigma$.

Exercise 2.4. Prove that, for all terms t without variables, $t_{\mathfrak{A}} = [c]$ if and only if $t = c$ belongs to Σ.

We complete the definition of $\mathfrak{A}$ by stipulating that

$$P_{\mathfrak{A}}([c_1], \ldots, [c_k]) \leftrightarrow P(c_1 \ldots c_k) \in \Sigma\,.$$

Here we let $P_{\mathfrak{A}}(q) \leftrightarrow q \in P_{\mathfrak{A}}$, and we also use the notational convention introduced above. The proof that the $P_{\mathfrak{A}}$ are well-defined is like the corresponding proof for the $F_{\mathfrak{A}}$.

Lemma 2.6. *Let $\varphi(x)$ be a formula of $\mathcal{L}$, let c be a constant of $\mathcal{L}$, and let $\mathfrak{B}$ be a model for $\mathcal{L}$. Then $\mathfrak{B} \models \varphi[c_{\mathfrak{B}}]$ if and only if $\mathfrak{B} \models \varphi(c)$, where $\varphi(c)$ is the result of replacing the free occurrences of x in $\varphi(x)$ by occurrences of c.*

We omit the proof, an easy induction on the complexity of $\varphi(x)$.

The following lemma completes the proof of the theorem.

Lemma 2.7. *For every sentence σ of $\mathcal{L}$, $\mathfrak{A} \models \sigma$ if and only if $\sigma \in \Sigma$.*

Proof. We proceed by induction of the complexity of σ.

Suppose σ is $t_1 = t_2$. Let $t_{1\mathfrak{A}} = [c_1]$ and $t_{2\mathfrak{A}} = [c_2]$. The $\mathfrak{A} \models \sigma \Leftrightarrow [c_1] = [c_2] \Leftrightarrow c_1 = c_2 \in \Sigma \Leftrightarrow$ (by Exercise 2.4, (V), and (III)) $t_1 = t_2 \in \Sigma$.

The case that σ is $P(t_1 \ldots t_k)$ is similar to the case that σ is $t_1 = t_2$.

If σ is $\neg\tau$, then $\mathfrak{A} \models \sigma \Leftrightarrow \mathfrak{A} \not\models \tau \Leftrightarrow \tau \notin \Sigma \Leftrightarrow$ (by (1) and (3)) $\sigma \in \Sigma$.

We have the following truth-functional implications:

$$\{(\tau_1 \wedge \tau_2)\} \models_{\text{tf}} \tau_1 \quad \{(\tau_1 \wedge \tau_2)\} \models_{\text{tf}} \tau_2 \quad \{\tau_1, \tau_2\} \models_{\text{tf}} (\tau_1 \wedge \tau_2)\,.$$

If σ is $(\tau_1 \wedge \tau_2)$ then $\mathfrak{A} \models \sigma \Leftrightarrow (\mathfrak{A} \models \tau_1$ and $\mathfrak{A} \models \tau_2) \Leftrightarrow (\tau_1 \in \Sigma$ and $\tau_2 \in \Sigma) \Leftrightarrow$ (by (III) and Lemma 2.2) $(\tau_1 \wedge \tau_2) \in \Sigma$.

(IX) $$\models (\varphi(c) \rightarrow (\exists x)\, \varphi(x) \quad \text{for } c \text{ a constant}$$

Suppose that σ is $(\exists x)\, \varphi(x)$. Then $\mathfrak{A} \models \sigma \Leftrightarrow$ there is an $a \in A$ such that $\mathfrak{A} \models \varphi[a] \Leftrightarrow$ there is a constant c of $\mathcal{L}$ such that $\mathfrak{A} \models \varphi[[c]] \Leftrightarrow$ (by Lemma 2.6) there is a constant c of $\mathcal{L}$ such that $\mathfrak{A} \models \varphi(c) \Leftrightarrow$ there is a constant c of $\mathcal{L}$ such that $\varphi(c) \in \Sigma \Leftrightarrow$ ($\Rightarrow$ by (IX), (III), and Lemma 2.2; $\Leftarrow$ by hypothesis (2)) $(\exists x)\, \varphi(x) \in \Sigma$. □

Theorem 2.8. (Uses Choice) *Let $\mathcal{L}$ be a language and let $\mathcal{L}^*$ be obtained from $\mathcal{L}$ by adding* $\max\{\text{card}\,(\mathcal{L}), \aleph_0\}$ *new constant symbols, where* $\text{card}\,(\mathcal{L})$ *is the cardinal number of the set of all non-logical symbols of $\mathcal{L}$. Let Σ be a set of sentences of $\mathcal{L}$ such that every finite subset of Σ is consistent (in $\mathcal{L}$).*

Then there is a set $\Sigma^* \supseteq \Sigma$ of sentences of $\mathcal{L}^*$ such that (1) every finite subset of Σ^* is consistent (in $\mathcal{L}^*$), (2) Σ^* has Henkin witnesses, and (3) for each sentence σ of $\mathcal{L}^*$, either $\sigma \in \Sigma^*$ or $\neg\sigma \in \Sigma^*$.

Proof. Let

$$\kappa = \max\{\aleph_0, \mathrm{card}(\mathcal{L})\}.$$

By Theorem 1.19, $\kappa^{<\omega} = \kappa$. Since κ is the cardinal of the set of all symbols of $\mathcal{L}^*$, the cardinal of the set of all sentences of $\mathcal{L}^*$ is $\leq \kappa^{<\omega}$. There are at least κ sentences of $\mathcal{L}^*$. (Consider sentences $c = c$ for constants c.) Thus κ is the cardinal of the set of all sentences of $\mathcal{L}^*$. Let

$$\alpha \mapsto \sigma_\alpha$$

be a one-one onto function from κ to the set of all sentences of $\mathcal{L}^*$.

Let r be a wellordering of the set of all constant symbols of $\mathcal{L}^*$.

By transfinite recursion, we define sets Σ_α of sentences of $\mathcal{L}^*$ for $\alpha \leq \kappa$. We shall arrange that

(a) $\Sigma_0 = \Sigma$;

(b) $\Sigma_\lambda = \bigcup\{\Sigma_\beta \mid \beta < \lambda\}$ for limit ordinals $\lambda \leq \kappa$;

(c) for $\beta \leq \alpha \leq \kappa$, $\Sigma_\beta \subseteq \Sigma_\alpha$;

(d) for $\alpha \leq \kappa$, every finite subset of Σ_α is consistent (in $\mathcal{L}^*$);

(e) $\mathrm{card}\,(\Sigma_{\alpha+1} \setminus \Sigma_\alpha) \leq 2$ for $\alpha < \kappa$;

(f) for $\alpha < \kappa$, either $\sigma_\alpha \in \Sigma_{\alpha+1}$ or $\neg\sigma_\alpha \in \Sigma_{\alpha+1}$;

(g) if $\alpha < \kappa$, if σ_α is $(\exists x)\,\varphi(x)$, and if $\sigma_\alpha \in \Sigma_{\alpha+1}$, then $\varphi(c) \in \Sigma_{\alpha+1}$ for some constant c of $\mathcal{L}^*$.

Once we carry out this construction, we can finish the proof by setting $\Sigma^* = \Sigma_\kappa$.

For $\alpha = 0$ and for limit α, we define Σ_α as required by conditions (a) and (b) respectively. Since consistency in $\mathcal{L}$ implies consistency in $\mathcal{L}^*$, (d) holds for $\alpha = 0$. Furthermore (d) holds for limit Σ_λ unless (c) fails for some β and $\alpha < \lambda$ or (d) fails for some $\alpha < \lambda$. for λ in place of κ. This is because, as is not difficult to prove, if Δ is a finite subset of Σ_λ then there is a $\beta < \lambda$ such that $\Delta \subseteq \Sigma_\beta$.

It follows that, however we define Σ_α for successor ordinals α, the smallest ordinal $\gamma \leq \kappa$ such that (a)–(g) fail for the Σ_β, $\beta \leq \gamma$, would have to be a successor ordinal.

Assume then that $\alpha < \kappa$ and that we are given Σ_β, $\beta \leq \alpha$, violating none of (a)–(g).

Suppose first that $\Delta \cup \{\neg\sigma_\alpha\}$ is consistent for every finite $\Delta \subseteq \Sigma_\alpha$. Set

$$\Sigma_{\alpha+1} = \Sigma_\alpha \cup \{\neg\sigma_\alpha\}.$$

Clearly none of (a)–(g) are violated by the Σ_β, $\beta \leq \alpha + 1$.

Before considering the other case, we prove the following lemma.

Lemma 2.9. *Let Δ be a set of sentences and let σ be a sentence. If $\Delta \cup \{\neg\sigma\}$ is inconsistent, then $\Delta \models \sigma$.*

Proof. We use two more facts about $\models$:

(X) $$\Delta \cup \{\sigma\} \models \tau \;\rightarrow\; \Delta \models (\sigma \rightarrow \tau)$$

(XI) $$(\Gamma \models \tau \;\wedge\; (\forall\sigma \in \Gamma)\, \Delta \models \sigma) \;\rightarrow\; \Delta \models \tau$$

We also need that

$$\{(\neg\sigma \rightarrow \tau), (\neg\sigma \rightarrow \neg\tau)\} \models_{\mathrm{tf}} \sigma.$$

Suppose that $\Delta \cup \{\neg\sigma\}$ is inconsistent. For some sentence τ, we have that

$$\Delta \cup \{\neg\sigma\} \models \tau\,;$$
$$\Delta \cup \{\neg\sigma\} \models \neg\tau\,.$$

By (X) we get that $\Delta \models$ both $(\neg\sigma \rightarrow \tau)$ and $(\neg\sigma \rightarrow \neg\tau)$. By (III) and (XI) we get that $\Delta \models \sigma$. □

Now suppose that there is a finite $\Delta \subseteq \Sigma_\alpha$ such that $\Delta \cup \{\neg\sigma_\alpha\}$ is inconsistent. Fix such a Δ. By Lemma 2.9, we have that $\Delta \models \sigma_\alpha$.

The cardinal number of $\Sigma_\alpha \setminus \Sigma$ is $\leq 2 \cdot \mathrm{card}(\alpha) < \kappa$. Therefore the cardinal number of the set of all new constants of $\mathcal{L}^*$ (i.e., those that are not constants of $\mathcal{L}$) occurring in $\Sigma_\alpha \cup \{\sigma_\alpha\}$ is $< \kappa$. Since κ is the cardinal number of the set of all new constants of $\mathcal{L}^*$, let c_α be the r-least constant of $\mathcal{L}^*$ not occurring in $\Sigma_\alpha \cup \{\sigma_\alpha\}$.

Let

$$\Sigma_{\alpha+1} = \Sigma_\alpha \cup \{\sigma_\alpha\},$$

unless σ_α is $(\exists x)\, \varphi_\alpha(x)$ for some formula φ_α, in which case let

$$\Sigma_{\alpha+1} = \Sigma_\alpha \cup \{\sigma_\alpha, \varphi_\alpha(c_\alpha)\}.$$

If we can prove that every finite subset of $\Sigma_{\alpha+1}$ is consistent, then we will have shown that (a)–(g) do not fail for the Σ_β, $\beta \leq \alpha + 1$, and so we will have completed the proof of the theorem.

Assume that $\Delta' \cup \{\sigma_\alpha\}$ is inconsistent for some finite subset Δ' of Σ_α. By (XI), (III), and the fact that $\{\neg\neg\sigma_\alpha\} \models_{\mathrm{tf}} \sigma_\alpha$, we get that $\Delta' \cup \{\neg\neg\sigma_\alpha\}$

is inconsistent. By Lemma 2.9, we get that $\Delta' \models \neg\sigma_\alpha$. But then $\Delta \cup \Delta'$ is an inconsistent finite subset of Σ_α.

$$\text{(XII)} \qquad \begin{array}{l} \Delta \cup \{\psi(c)\} \models \tau \;\rightarrow\; \Delta \cup \{(\exists x)\, \psi(x)\} \models \tau \\ \quad \text{for } c \text{ is a constant not occurring in } \Delta,\ \psi(x), \text{ or } \tau \end{array}$$

(If $\mathfrak{B}$ is a model satisfying $\Delta \cup \{(\exists x)\psi(x)\}$ but not τ, then let $b \in B$ be such that $\mathfrak{B} \models \psi[b]$. Let $\mathfrak{B}'$ be like $\mathfrak{B}$, except that $c_{\mathfrak{B}'} = b$. Then $\mathfrak{B}'$ satisfies $\Delta \cup \{\psi(c)\}$ but not τ.)

Assume that some finite subset of $\Sigma_{\alpha+1}$ is inconsistent. Then $\Sigma_{\alpha+1} = \Sigma_\alpha \cup \{\sigma_\alpha, \varphi_\alpha(c_\alpha)\}$, and there is a finite $\bar{\Delta} \subseteq \Sigma_\alpha$ and there is a sentence τ such that

$$\begin{array}{l} \bar{\Delta} \cup \{\sigma_\alpha, \varphi_\alpha(c_\alpha)\} \models \tau\,; \\ \bar{\Delta} \cup \{\sigma_\alpha, \varphi_\alpha(c_\alpha)\} \models \neg\tau\,. \end{array}$$

Using the the truth-functional implication $\{\tau, \neg\tau\} \models_{\mathrm{tf}} \tau'$, we may assume that c_α does not occur in τ. By (XII) we have

$$\begin{array}{l} \bar{\Delta} \cup \{\sigma_\alpha, (\exists x)\, \varphi_\alpha(x)\} \models \tau\,; \\ \bar{\Delta} \cup \{\sigma_\alpha, (\exists x)\, \varphi_\alpha(x)\} \models \neg\tau\,. \end{array}$$

But σ_α is $(\exists x)\, \varphi_\alpha(x)$, so we have the contradiction that $\bar{\Delta} \cup \{\sigma_\alpha\}$ is inconsistent. □

Theorem 2.10. (Compactness I and Weak Löwenheim–Skolem Theorem) (Uses Choice) *Let Σ be a set of sentences of a language $\mathcal{L}$ such that every finite subset of Σ is consistent. Then there is a model $\mathfrak{A}$ of Σ such that* $\mathrm{card}(\mathfrak{A}) \le \max\{\aleph_0, \mathrm{card}(\mathcal{L})\}$.

Proof. Let $\mathcal{L}^*$ be as in the statement of Theorem 2.8. Let Σ^* be given by that theorem. Let $\mathfrak{A}^*$ be the model of Σ^* given by Theorem 2.1. Let $\mathfrak{A}$ be the *reduct* of $\mathfrak{A}^*$ to $\mathcal{L}$. Clearly $\mathfrak{A} \models \Sigma$. □

Theorem 2.11 (Compactness II). (Uses Choice) *Let Σ be a set of sentences and let σ be a sentence. If $\Sigma \models \sigma$ then there is a finite $\Delta \subseteq \Sigma$ such that $\Delta \models \sigma$.*

Proof. Suppose that $\Sigma \models \sigma$. Then $\Sigma \cup \{\neg\sigma\}$ is inconsistent. By Theorem 2.10, there is a finite $\Delta \subseteq \Sigma$ such that $\Delta \cup \{\neg\sigma\}$ is inconsistent. But then $\Delta \models \sigma$. □

Exercise 2.5. Let $\mathcal{L}$ be any language. A class K of models for $\mathcal{L}$ is EC (is an *elementary class*) if there is a sentence σ of $\mathcal{L}$ such that

$$K = \{\mathfrak{A} \mid \mathfrak{A} \models \sigma\}\,.$$

A class K is EC_Δ if there is a set Σ of sentences of $\mathcal{L}$ such that

$$K = \{\mathfrak{A} \mid \mathfrak{A} \models \Sigma\}\,.$$

Which of the following are EC_Δ ?

(i) $\{\mathfrak{A} \mid A \text{ is infinite}\}$;

(ii) $\{\mathfrak{A} \mid A \text{ is finite}\}$.

Show that neither is EC.

Theorem 2.12. *Assume that ZFC (i.e., the set of axioms of ZFC) is consistent. For variables x, let* Number (x) *be the formula "x is a natural number."*

There is a model $\mathfrak{A}$ of ZFC and an $a \in A$ such that $\mathfrak{A} \models \mathrm{Number}\,[a]$ and such that $\in_\mathfrak{A} \restriction \{b \mid b \in_\mathfrak{A} a\}$ is not wellfounded.

Proof. For $n \in \omega$, let $\chi_n(x)$ be the formula "$x = n$." ($\chi_n(x)$ is defined by recursion on n.) Let $\mathcal{L}^*$ be the result of adding to the language of set theory a constant c. Let

$$\Sigma = \mathrm{ZFC} \cup \{\mathrm{Number}\,(c)\} \cup \{(\forall v_0)(\chi_n(v_0) \to v_0 \in c) \mid n \in \omega\}\,.$$

Let Δ be a finite subset of Σ. Then there is some $m \in \omega$ such that

$$\Delta \subseteq \mathrm{ZFC} \cup \{\mathrm{Number}\,(c)\} \cup \{(\forall v_0)(\chi_n(v_0) \to v_0 \in c) \mid n < m\}\,.$$

Let $\mathfrak{B}$ be a model of ZFC. For each $n \in \omega$ there is a unique $b \in B$ such that $\mathfrak{B} \models \chi_n[b]$; let $n^\mathfrak{B}$ be this unique b. Expand $\mathfrak{B}$ to a model $\mathfrak{B}^*$ for $\mathcal{L}^*$ by letting $c_{\mathfrak{B}^*} = m^\mathfrak{B}$. Clearly $\mathfrak{B}^* \models \Delta$.

Since every finite subset of Σ is consistent, there is by Theorem 2.10 a model $\mathfrak{A}^*$ of Σ. Let $\mathfrak{A}$ be the reduct of $\mathfrak{A}^*$ to $\mathcal{L}$, and let $a = c_{\mathfrak{A}^*}$.

To see that $\in_\mathfrak{A} \restriction \{b \mid b \in_\mathfrak{A} a\}$ is not wellfounded, let

$$y = \{b \mid b \in_\mathfrak{A} a \wedge (\forall n \in \omega)\mathfrak{A} \not\models \chi_n[b]\}.$$

Since the $\in_\mathfrak{A}$-immediate predecessor of a belongs to y, y is nonempty. For any $b \in y$, the $\in_\mathfrak{A}$-immediate predecessor of b belongs to y, so y has no $\in_\mathfrak{A}$-least element. □

Remark. If $\mathfrak{A}$ and a are as in the statement of Theorem 2.12, then a is a *non-standard* natural number of $\mathfrak{A}$. In §3, we shall construct models with non-standard real numbers.

If $\mathfrak{A}$ and $\mathfrak{B}$ are models for a language $\mathcal{L}$, then $\mathfrak{A}$ and $\mathfrak{B}$ are *elementarily equivalent* ($\mathfrak{A} \equiv \mathfrak{B}$) if they satisfy the same sentences of $\mathcal{L}$.

Theorem 2.13. *Let $\mathcal{L}$ be a language and let $\kappa = \max\{\aleph_0, \operatorname{card}(\mathcal{L})\}$. Every model for $\mathcal{L}$ is elementarily equivalent to a model of cardinal $\leq \kappa$.*

Proof. Let $\mathfrak{B}$ be a model for $\mathcal{L}$. The *theory of* $\mathfrak{B}$ ($\mathrm{Th}(\mathfrak{B})$), the set of all sentences σ such that $\mathfrak{B} \models \sigma$, is consistent. Apply Theorem 2.10. □

Formal Deduction

Fix a language $\mathcal{L}$.

Logical Axioms:

(1) All tautologies.

(2) Identity Axioms:

(a) $t = t$
for t a term;

(b) $(t_1 = t_2 \to (\varphi(t_1, y_1, \ldots, y_n) \to \varphi(t_2, y_1, \ldots, y_n)))$
for t_1 and t_2 terms and $\varphi(x, y_1, \ldots, y_n)$ an atomic formula.

(3) Quantifier Axioms:

$$(\psi(t, y_1, \ldots, y_n) \to (\exists x)\, \psi(x, y_1, \ldots, y_n)),$$

for $\psi(x, y_1, \ldots, y_n)$ a formula and t a term such that no occurrence of a variable in t gives a bound occurrence of the variable in $\psi(t, y_1, \ldots, y_n)$.

Rules:

(1) Modus Ponens: $$\frac{\varphi \qquad (\varphi \to \psi)}{\psi}$$
for φ and ψ formulas;

(2) Quantifier Rule: $$\frac{(\varphi \to \psi)}{((\exists x)\varphi \to \psi)}$$

for φ and ψ formulas with x not free in ψ.

Remark. In stating the axioms and rules, we have used abbreviations involving the symbol "$\to$" (introduced on page 2).

A *deduction* in $\mathcal{L}$ from a set Σ of sentences is a finite sequence of formulas (the *lines* of the deduction) such that every formula in the sequence either (i) belongs to Σ, (ii) is a logical axiom, or (iii) follows from earlier formulas by one of the two rules. A *deduction* in $\mathcal{L}$ *of* a sentence τ from Σ is a deduction in $\mathcal{L}$ from Σ with last line τ.

A set Σ of sentences *deductively implies* in $\mathcal{L}$ a sentence τ ($\Sigma \vdash_{\mathcal{L}} \tau$) if there is a deduction in $\mathcal{L}$ of τ from Σ.

Remark. It will turn out that deductive implication is independent of $\mathcal{L}$, but this is not as easy to prove as the corresponding fact for the semantical notion of logical implication.

Theorem 2.14 (Soundness). *For any language $\mathcal{L}$, if $\Sigma \vdash_{\mathcal{L}} \tau$ then $\Sigma \models \tau$.*

Proof. Let $\mathcal{D}$ be a deduction from Σ in $\mathcal{L}$ and let $\mathfrak{A}$ be any model of Σ. By induction one can show that, for all lines φ of $\mathcal{D}$ and for every s (with large enough domain), $\mathfrak{A} \models \varphi[s]$. This is trivial for $\varphi \in \Sigma$ and is easily checked for logical axioms. Moreover it is easy to see that applications of the rules preserve this property. □

Theorem 2.15. *For any language $\mathcal{L}$,* (I)–(XII) *hold with "$\vdash_{\mathcal{L}}$" in place of "$\models$."*

Remark. The modified (III), like the original (III), remains true if the restriction that Δ be finite, is removed. This is because—as is not difficult to show—compactness holds for truth-functional implication. Our reason for the restriction to finite Δ is to save ourselves the effort of proving the unrestricted version.

Proof. (I), (II), and (XI) follow directly from the notion of a deduction, and do not depend on our particular axioms and rules.

(IV) and (V) are Identity Axioms, and (VIII) follows from Identity Axioms (a) and (b) using Modus Ponens.

For (III), suppose that $\Delta \models_{\mathrm{tf}} \tau$ with Δ finite. Let Δ be $\{\sigma_i \mid i < n\}$. Then

$$(\sigma_0 \to (\sigma_1 \to \ldots \to (\sigma_{n-1} \to \tau) \cdots)$$

is a tautology. By n applications of Modus Ponens, we can get a deduction of τ from Δ.

(VI) follows by Modus Ponens from the Identity Axiom $v_1 = v_1$ and the Quantifier Axiom $(v_1 = v_1 \rightarrow (\exists v_1)\, v_1 = v_1)$.

For (VII), note that

$$F(c_1, \ldots, c_k) = F(c_1, \ldots, c_k)$$

is an Identity Axiom and that

$$(F(c_1, \ldots, c_k) = F(c_1, \ldots, c_k) \;\rightarrow\; (\exists x)\, F(c_1, \ldots, c_k) = x)$$

is a Quantifier Axiom. (VII) follows from these axioms by Modus Ponens.

(IX) is a Quantifier Axiom.

(X) is commonly called the *Deduction Theorem*. To prove it, let $\mathcal{D}$ be a deduction in $\mathcal{L}$ of τ from $\Delta \cup \{\sigma\}$. Get a new sequence $\mathcal{D}'$ of formulas by replacing each line φ of $\mathcal{D}$ by $(\sigma \rightarrow \varphi)$. We shall show how to turn $\mathcal{D}'$ into a deduction of $(\sigma \rightarrow \tau)$ from Δ by inserting additional lines.

If a line φ of $\mathcal{D}$ belongs to Δ or is a logical axiom, then insert φ and the tautology $(\varphi \rightarrow (\sigma \rightarrow \varphi))$. The line $(\sigma \rightarrow \varphi)$ then comes by Modus Ponens.

If a line of $\mathcal{D}$ is σ, then the corresponding line of $\mathcal{D}'$ is the tautology $(\sigma \rightarrow \sigma)$.

If a line φ of $\mathcal{D}$ comes from earlier lines ψ and $(\psi \rightarrow \varphi)$ by Modus Ponens, then insert the tautology

$$(\dagger) \qquad ((\sigma \rightarrow \psi) \rightarrow ((\sigma \rightarrow (\psi \rightarrow \varphi)) \rightarrow (\sigma \rightarrow \varphi)))$$

and the formula

$$(\ddagger) \qquad ((\sigma \rightarrow (\psi \rightarrow \varphi)) \rightarrow (\sigma \rightarrow \varphi))\,.$$

($\ddagger$) comes from the ($\dagger$) and $(\sigma \rightarrow \psi)$ by Modus Ponens, and $(\sigma \rightarrow \varphi)$ then comes from the ($\ddagger$) and $(\sigma \rightarrow (\psi \rightarrow \varphi))$ by another application of Modus Ponens.

Suppose finally that a line of $\mathcal{D}$ is $((\exists x)\varphi \rightarrow \psi)$ and that it comes from an earlier line $(\varphi \rightarrow \psi)$ by the Quantifier Rule. That earlier line corresponds to the line $(\sigma \rightarrow (\varphi \rightarrow \psi))$ of $\mathcal{D}'$. Insert the following lines:

$$\begin{array}{l}
((\sigma \rightarrow (\varphi \rightarrow \psi)) \rightarrow (\varphi \rightarrow (\sigma \rightarrow \psi))) \\
(\varphi \rightarrow (\sigma \rightarrow \psi)) \\
((\exists x)\varphi \rightarrow (\sigma \rightarrow \psi)) \\
(((\exists x)\varphi \rightarrow (\sigma \rightarrow \psi)) \rightarrow (\sigma \rightarrow ((\exists x)\varphi \rightarrow \psi))) \\
(\sigma \rightarrow ((\exists x)\varphi \rightarrow \psi))
\end{array}$$

The first and fourth of these lines are tautologies. The second and fifth come by Modus Ponens. The third comes by the Quantifier Rule. Finally, the line $(\sigma \to \varphi)$ comes by Modus Ponens.

It remains only to show that (XII) holds. Assume that $\Delta \cup \{\psi(c)\} \vdash_{\mathcal{L}} \tau$ and that the conditions of (XII) are met. By (X) we have that $\Delta \vdash_{\mathcal{L}} (\psi(c) \to \tau)$. Let $\mathcal{D}$ be a deduction witnessing this fact. Let y be a variable not occurring in $\mathcal{D}$. We get a deduction $\mathcal{D}'$ from Δ with last line $(\psi(y) \to \tau)$ by replacing each occurrence of c in $\mathcal{D}$ by an occurrence of y. Applying the Quantifier Rule to the last line of $\mathcal{D}'$, we get $((\exists y)\psi(y) \to \tau)$. From this, the Quantifier Axiom $(\psi(x) \to (\exists y)\psi(y))$, and tautologies and Modus Ponens, we get $(\psi(x) \to \tau)$. The Quantifier Rule now gives $((\exists x)\psi(x) \to \tau)$. This argument shows that $\Delta \vdash_{\mathcal{L}} ((\exists x)\psi(x) \to \tau)$. Using Modus Ponens, we can deduce that $\Delta \cup \{(\exists x)\psi(x)\} \vdash_{\mathcal{L}} \tau$. □

Let us say that a set Σ of sentences of a language $\mathcal{L}$ is *deductively consistent* in $\mathcal{L}$ if there is no sentence τ of $\mathcal{L}$ such that $\Sigma \vdash_{\mathcal{L}} \tau$ and $\Sigma \vdash_{\mathcal{L}} \neg\tau$. Otherwise Σ is *deductively inconsistent* in $\mathcal{L}$. Since deductions are finite, a set Σ of sentences is deductively consistent in $\mathcal{L}$ if and only if every finite subset of Σ is deductively consistent in $\mathcal{L}$.

Theorem 2.16. (Uses Choice) *Let Σ be a set of sentences of a language $\mathcal{L}$. Suppose that*

(1) Σ is deductively consistent in $\mathcal{L}$;

(2) Σ has Henkin witnesses;

(3) for each sentence σ of $\mathcal{L}$, either $\sigma \in \Sigma$ or $\neg\sigma \in \Sigma$.

Then Σ has a model $\mathfrak{A}$ such that $\mathrm{card}\,(\mathfrak{A}) \leq$ *the cardinal number of the set of constant symbols of $\mathcal{L}$.*

(As with Theorem 2.1, Choice is needed only to guarantee that the set of all constant symbols of $\mathcal{L}$ has a cardinal number.)

Proof. The proof is exactly like that of Theorem 2.1, using Theorem 2.15. □

Theorem 2.17. *Let Σ be a set of sentences of a language $\mathcal{L}$ such that Σ is deductively consistent in $\mathcal{L}$. Let $\mathcal{L}^*$ be obtained from $\mathcal{L}$ by adding new constant symbols. Then Σ is deductively consistent in $\mathcal{L}^*$.*

Proof. Assume that Σ is deductively inconsistent in $\mathcal{L}^*$. Then there is a sentence τ, which we may without loss of generality assume to be a sentence of $\mathcal{L}$, such that $\Sigma \vdash_{\mathcal{L}*} \tau$ and $\Sigma \vdash_{\mathcal{L}*} \neg\tau$. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be deductions witnessing these facts. Let $c_1, \ldots, c_n$ be distinct and be all the constants of $\mathcal{L}^*$ occurring in either of $\mathcal{D}_1$ or $\mathcal{D}_2$ that are not constants of $\mathcal{L}$. Let $y_1, \ldots, y_n$ be distinct variables not occurring in $\mathcal{D}_1$ or $\mathcal{D}_2$. Obtain $\mathcal{D}'_1$ and $\mathcal{D}'_2$ from $\mathcal{D}_1$ and $\mathcal{D}_2$ respectively by replacing, for each i, each occurrence of c_i by an occurrence of y_i. Then $\mathcal{D}'_1$ and $\mathcal{D}'_2$ witness that $\Sigma \vdash_{\mathcal{L}} \tau$ and $\Sigma \vdash_{\mathcal{L}} \neg\tau$ respectively.

Theorem 2.18. (Uses Choice) *Let $\mathcal{L}$ be a language and let $\mathcal{L}^*$ be obtained from $\mathcal{L}$ by adding* $\max\{\mathrm{card}\,(\mathcal{L}), \aleph_0\}$ *new constant symbols. Let Σ be a set of sentences of $\mathcal{L}$ such that Σ is deductively consistent in $\mathcal{L}$.*

Then there is a set $\Sigma^* \supseteq \Sigma$ of sentences of $\mathcal{L}^*$ such that (1) Σ^* is deductively consistent in $\mathcal{L}^*$, (2) Σ^* has Henkin witnesses, and (3) for each sentence σ of $\mathcal{L}^*$, either $\sigma \in \Sigma^*$ or $\neg\sigma \in \Sigma^*$.

Proof. The proof is exactly like that of Theorem 2.8, using Theorem 2.15 and using Theorem 2.17 to get that $\Sigma_0 = \Sigma$ is deductively consistent in $\mathcal{L}^*$. □

Theorem 2.19. (Uses Choice) *Let Σ be a set of sentences of a language $\mathcal{L}$ that is deductively consistent in $\mathcal{L}$. Then there is a model $\mathfrak{A}$ of Σ such that* $\mathrm{card}(\mathfrak{A}) \leq \max\{\aleph_0, \mathrm{card}(\mathcal{L})\}$.

Proof. The proof is like that of Theorem 2.10. □

Theorem 2.20 (Gödel Completeness Theorem). (Uses Choice) *Let Σ be a set of sentences of a language $\mathcal{L}$ and let σ be a sentence of $\mathcal{L}$. If $\Sigma \models \sigma$ then $\Sigma \vdash_{\mathcal{L}} \sigma$.*

Proof. Assume that $\Sigma \nvdash_{\mathcal{L}} \sigma$. Then, by the analogue of Lemma 2.9, $\Sigma \cup \{\neg\sigma\}$ is deductively consistent in $\mathcal{L}$. By Theorem 2.19, there is a model $\mathfrak{A}$ for $\mathcal{L}$ such that $\mathfrak{A} \models \Sigma \cup \{\neg\sigma\}$. But then $\Sigma \not\models \sigma$. □

Because of the Soundness and Completeness Theorems, the symbol "$\vdash_{\mathcal{L}}$," is superfluous, and we shall make no further use of it.

Exercise 2.6. Let $\mathcal{L}$ be a language with a one-place relation symbol F. Give a deduction witnessing the following

$$\{\neg(\exists v_1)\neg F(v_1)\} \;\vdash_{\mathcal{L}}\; \neg(\exists v_2)\neg F(v_2).$$

Exercise 2.7. Suppose we replaced our Quantifier Rule with the following additional Logical Axioms:

$$((\varphi \to \psi) \to ((\exists x)\varphi \to \psi))$$
$$\text{for } x \text{ not occurring free in } \psi.$$

Would Soundness still hold? Would Completeness still hold? Prove your answers.