Math 454 Fall 2014

David Ross, Department of Mathematics

1 Introduction

1.1 Goals

- Foundational
  - What is a set/integer/function/etc?
  - What rules/axioms do we need to do mathematics?
  - Are these rules/axioms we use consistent?
- Philosophical
  - Resolve paradoxes (Russell, Burali-Forti, Skolem-Lowenheim)
- Mathematical
  - Tools:
    - induction on well-founded sets
    - axiom of choice and equivalents
    - infinite numbers (ordinals, cardinals, infinitary arithmetic)
    - infinitary combinatorics (König lemma, Ramsey Theory, trees)
  - Questions: (G)CH, others (algebra, analysis, topology,...)

1.2 Needs

- Familiarity/comfort with mathematical manipulations/proof (need to use them, need to recognize what we're using)
- Very naive set theory (notation: $\in, \{x\ :\ \cdots\}, \bigcup, \bigcap, \subset, \subseteq, \supset, \subsetneq, \setminus, \times, \emptyset, \triangle, {}^{\complement}, \mathcal{P}\left(\right) \ldots$)
- Induction on $\mathbb{N}$
- 1st-order logic (notation: $\wedge, \vee, \neg, \Rightarrow, \exists, \forall, \nexists, \ldots$)
- Open mind

2 The Language of Set Theory

2.1 Alphabet

- **Variables:** $w, x, y, z, \ldots$ (or $A, B, C, \ldots$; or $x_1, x_2, \ldots$)
- **Connectives:** $\wedge$, $\neg$ (can *define* the others $\vee$, $\implies$, etc. in terms of these)
- **Quantifiers:** $\forall$ (can define $\exists$ by $\exists x \phi(x) := \neg \forall x \neg \phi(x)$)
- **Non-logical symbols:** (,)
- **Equality Symbol:** $=$
- **Membership Symbol:** $\in$

2.2 Formulas

1. **Atomic Formulas:** $x = y$, $x \in y$ (where x, y are variable symbols)
2. **Compound Formulas:**
    - If ϕ is a formula then so is $\neg\phi$
    - If ϕ, ψ are formulas then so is $(\phi \wedge \psi)$
    - If ϕ is a formula and x is a variable symbol then $\forall x \phi$ is a formula
3. The only formulas are the atomic and compound formulas as defined above.

2.3 Remarks

1. We will be fairly profligate with abbreviations (e.g., $\phi \vee \psi$ for $\neg(\neg\phi \wedge \neg\psi)$, $\phi \implies \psi$ for $\neg\phi \vee \psi$), and inconsistent with parentheses.
2. Every formula is a *word* – that is, a string of symbols – on the alphabet $\{x, y, z, \ldots, \wedge, \neg, \forall, (,), =, \in\}$. For anyone familiar with trees (which we will discuss later). we could instead have treated a formula as a tree of subformulas and thereby eliminated the need for parentheses.
3. All the symbols except the $\in$ are common to every *first-order* (or *predicate*) logic. One usually specifies a particular logical *language* solely in terms of the symbols which are unique to the application. So, for example, we would say that the *language of set theory* is $\mathcal{L}_{set} = \{\in\}$.
4. If we had urelements which we wanted to distinguish somehow, we could add *constant symbols* to $\mathcal{L}\ _{set}$ to represent them. For example, if we decided to make the natural numbers $\{0, 1, 2, \ldots\}$ urelements, then we could represent each number n by a numeral $\mathbf{n}$, and our language would be $\{\in$, $\mathbf{0}$, $\mathbf{1}$, $\mathbf{2}$, $\ldots\}$

2.4 Exercises:

Expand the following abbreviations out in full, gory detail as correct formulas:

1. $\emptyset \in x$

2. There exists a *unique* x such that $\phi(x)$ holds. (We will usually abbreviate this as $\exists! x \phi(x)$

3 Axioms

3.1 Extensionality

- Axiom (Informal): Two sets with the same elements are the same set.
- Axiom (Formal): $\forall z(z \in x \iff z \in y) \Rightarrow x = y$
- Idea: A set is determined by its elements, not (for example) its presentation. BTW, this looks like an assertion about $=$, but really is an assertion about $\in$, the reason being that $=$ is a primitive symbol of the *logic* – that is the formal system – and is not specific to the formal study of *set theory*.

3.2 Empty Set

- Axiom (Informal): There exists a set with no elements.
- Axiom (Formal): $\exists x \forall z(z \notin x)$
- Idea: We need at least one set for a reasonable universe. Once we have *any* set a, we can get the empty set from it using the Subset Axiom (see below) and the formula "$x \neq x$", that is, we can define $\emptyset$ to be $\{x \in a \ : \ x \neq x\}$. So, the assertion that the set $\emptyset$ exists is the weakest such assertion (and becomes superfluous if we ever end up asserting the existence of any other set, which we will in fact do later with the Axiom of Infinity).
- Note: By Extensionality, the empty set – once it exists – is unique. From now on, we will often use $\emptyset$ as shorthand for a more complicated statement involving the above formula. For example, the statement $\emptyset \in A$ is really shorthand for $\exists x(\forall z(z \notin x) \wedge x \in A)$. (Equivalently, $\forall x(\forall z(z \notin x) \Rightarrow x \in A)$ — convince yourself that this really is equivalent!)

3.3 Pairing

- Axiom (Informal): Given two sets a and b, there is a third set $\{a, b\}$
- Axiom (Formal): $\forall a \forall b \exists x \forall z (z \in x \iff (z = a \vee z = b))$
- Idea: This is the simplest rule for creating new sets from old. In particular, it lets you create the set $\{a\} = \{a, a\}$ from a set a; and, given a finite set $a_1, \ldots, a_n$ of sets, it lets you create the set $\{a_1, \ldots, a_n\}$. If you peek ahead to see how we define the natural numbers, you will see that we now have enough to construct every natural number (though not the *set* ω).

3.4 Union

- Axiom (Informal): Given any collection (set) of sets, the union of these sets exists.
- Axiom (Formal): $\forall x \exists y \forall z (z \in y \iff \exists w (w \in x \wedge z \in w))$
- Idea: We will frequently write $\bigcup X$ for the union of all the sets in X. Of course, one more often sees either the union of a *finite* number of sets, e.g. $A \cup B$, or of an indexed family of sets, e.g. $\bigcup_{i \in I} A_i$. The former is an obvious special case (especially in light of the Pairing axiom) and the latter will also be expressible in this way, once we formalize the notion of an indexed family.

  By the way, one can get the *intersection* for free using the Subset axiom (below): $z \in \bigcap X \iff \forall w (w \in X \implies z \in w)$ (I leave it to you to write this as an axiom the way we are writing the others). It is interesting that union requires a special axiom and intersection does not.

3.5 Power Set

- Axiom (Informal): Given a set x, there is another set containing all subsets of x
- Axiom (Formal): $\forall x \exists y \forall z (z \in y \iff \forall w (w \in z \Rightarrow w \in x))$
- Idea: This is another set construction axiom, and should be viewed with some suspicion since it is in fact very *non*constructive. You should observe how we formalize it using only the $\in$ symbol, and not the $\subset$ symbol. In fact, we should always view $\subset$ as an abbreviation, e.g. $A \subset B \iff \forall w (w \in A \Rightarrow w \in B)$

  Similarly, we will henceforth freely use the notation $\mathcal{P}\left(X\right)$ for the power set of X, and remember that its use in formulas is as an abbreviation, i.e. that we can always replace it by an expression in our formal language.

3.6 Subset (or Separation, or Selection) *Schema*

- Axiom (Informal): If A is a set and $\phi(x)$ a property of sets then $\{x \in A \mid \phi(x)\}$ is a set.
- Axiom (Formal):For each formula $\phi(x)$, $\forall A \exists B \forall z (z \in B \iff z \in A \wedge \phi(z))$
- Idea: This lets us select a subset of a given set based on some precise property. We haven't yet defined the word "formula" - we will do so rigorously later - but the meaning should be clear by now. The word *schema* refers to the fact that this is not just one axiom, but actually a list of axioms, one for each formula $\phi(x)$.

  Incidentally, for this axiom we can allow the formula $\phi(x)$ to mention other sets. Thus, if C is some other set we know about, ϕ might actually take the form $\phi(x, C)$, and the axiom becomes $\forall A \exists B \forall z (z \in B \iff z \in A \wedge \phi(z, C))$ This might suggest that our list of axioms is actually as big as our universe of sets (which might be very big indeed); fortunately, we can eliminate the need to include C explicitly in the formula by writing instead $\forall C \forall A \exists B \forall z (z \in B \iff z \in A \wedge \phi(z, C))$

3.7 Replacement (or Collection) *Schema*

- Axiom (Informal): If F a function and A is a subset of the domain of F then $F[A]$ is a set.

- Axiom (Formal): $\forall A[\forall x \forall y \forall z((x \in A \wedge \phi(x, y) \wedge \phi(x, z)) \Rightarrow (y = z)) \Rightarrow \exists B \forall z(z \in B \iff \exists x(x \in A \wedge \phi(x, z)))]$

- Idea: This is another way of building new sets from old. We use 'function-like' formulas $\phi(x, y)$ instead of actual functions F because (a) we haven't defined formally the word *function*, and (b) this is stronger, and we need the full strength.

  Of course, this is a schema like Subset, and as in Subset we allow other sets to appear in the definition of ϕ.

- Note: This implies the Union axiom; just take for $\phi(x, y)$ the formula $\forall z(z \in y \iff \exists w(w \in x \wedge z \in w))$. The 'function-like' nature of ϕ follows from Extension.

3.8 Infinity

- Axiom (Informal):There is an infinite set.

- Axiom (Formal): $\exists x[(\emptyset \in x) \wedge \forall y(y \in x \Rightarrow \mathcal{P}\ (y) \in x)]$

- Idea: This throws the set $\{\emptyset, \mathcal{P}\ (\emptyset), \mathcal{P}\ (\mathcal{P}\ (\emptyset)), \dots\}$ into our universe, which turns out to be enough to guarantee the existence of most infinite sets we will be interested in. Note that with this axiom we no longer need Empty Set as a separate axiom.

  In some developments, this axiom is not stated until ordinals are defined, at which point the axiom is just written as "ω exists".

3.9 Foundation (or Regularity)

- Axiom (Informal): The universe of sets is well-founded under $\in$
- Axiom (Formal): $\forall x(x \neq \emptyset) \Rightarrow (\exists a(a \in x \wedge a \cap x = \emptyset))$
- Idea: This is meant to capture the image of the set-theoretic universe as the cumulative hierarchy, in the "downward" direction. "Well-founded under $\in$" means that there is no infinite decreasing sequence of sets $\cdots x_3 \in x_2 \in x_1 \in x_0$. It would be hard to write this as a finite axiom, at least until we formally define ω and 'function'; the one given turns out to be adequate: put $X = \{x_0, x_1, x_2, \dots\}$ It also captures the idea that any set is of higher *rank* than any of its elements - this is the origin of the term 'regularity'.

  Note this axiom implies *a fortiori* that no set can be an element of itself, since if $x \in x$ then we get the sequence $\cdots x \in x \in x \in x$.

3.10 Choice

- Axiom (Informal): Given a set of disjoint nonempty sets, there is a new set containg exactly one element from each of the others.
- Axiom (Formal): $\forall x[((x \neq \emptyset) \wedge \forall z(z \in x \Rightarrow z \neq \emptyset) \wedge \forall z \forall w(w \neq z \wedge z \in x \wedge w \in x \implies \forall t \neg(t \in z \wedge t \in w)) \Rightarrow \exists y \forall z(z \in x \Rightarrow \exists w(w \in z \wedge w \in y \wedge \forall t((t \in z \wedge t \in y) \Rightarrow t = w)))]$
- Idea: Yet another way to build new sets, this axiom was controversial early in the century. The reason is that while it 'selects' an element from each of a set of nonempty set, it doesn't specify how the element is to be selected. In cases where such a recipe for selection exists, for example where all the sets in x are well-ordered, this axiom is unnecessary. Lest one wonder if it is really necessary, note that it is equivalent to the assertion that a Cartesian product of nonempty sets is nonempty, and is used to prove many important results in mathematics. In particular, it is equivalent to the assertion that vector spaces have bases, or that sets have cardinalities. For this reason, while some philosophers still occasionally question the axiom, few mathematicians eschew its use.

4 Coding the Universe

4.1 Ordered Pairs

Definition 4.1 *(*Kuratowski*) The* ordered pair *of x and y is the set* $\langle x, y\rangle := \{\{x\}, \{x, y\}\}$

Proposition 4.1 $\langle x, y\rangle = \langle z, w\rangle$ *if and only if* $x = z$ *and* $y = w$

(Proof: exercise)

- Question: Why the funny definition? (Examples of other plausible definitions...)
- Note: If $x, y \in A$ then $\langle x, y\rangle \in \mathcal{P}\left(\mathcal{P}\left(A\right)\right)$
- Note: If $\langle x, y\rangle \in A$ the $x, y \in \cup \cup A$
- The statement "v is an ordered pair" is a first-order formula in the variable v (Exercise: show!)
- We say that x is the *first coordinate* of $\langle x, y\rangle$, and that y is the *second coordinate* of $\langle x, y\rangle$

4.2 Cartesian Products

Definition 4.2 *The* Cartesian product *of A and B is defined by* $A \times B := \{\langle a, b\rangle \mid a \in A, b \in B\}$

- By an earlier remark, $A \times B$ is a subset of $\mathcal{P}\left(\mathcal{P}\left(A \cup B\right)\right)$; it follows that we can prove that $A \times B$ actually exists using subset selection applied to this larger set.
- Question: Is $(A \times B) \times C = A \times (B \times C)$?

4.3 Relations

Definition 4.3 *A (binary)* relation *is a set of ordered pairs.*

- We will define "n-ary relation" later, after defining "n-tuple"; in general, the word "relation" used by itself means a binary relation.
- Exercise: write "R is a relation" in our formal language.
- If R is a relation, we often write xRy instead of $\langle x, y\rangle \in R$
- If R is a relation, the *domain* of R is $dom(R) := \{x \mid \exists y xRy\}$ and the *range* of R is $ran(R) := \{y \mid \exists x xRy\}$. Note that $R \subset dom(R) \times ran(R)$, but that in general the inclusion will be proper.

The *field* of a relation R is $fld(R) := dom(R) \cup ran(R) = \cup \cup R$. (This is not standard terminology). We will often say "R is a relation on A" to mean R is a relation with $fld(R) \subseteq A$. (This *is* standard terminology!)

4.4 Properties of Relations

Definition 4.4 *A relation R is:*

1. symmetric *if* $\forall a, b(aRb \implies bRa)$

2. antisymmetric *if* $\forall a, b(aRb \implies \neg bRa)$

3. reflexive *if* $\forall a \in dom(R)\ aRa$

4. irreflexive *if* $\forall a \neg aRa$

5. transitive *if* $\forall a, b, c(aRb \wedge bRc \implies aRc)$

6. connected *if* $\forall a, b \in fld(R)(a \neq b \implies aRb \vee bRa)$

Some examples (all on N): $<$; $\leq$; "divides"; 'is relatively prime to"; "equal mod 9"; "has same prime factors"

Notation: Let R be a relation.

- $R^{-1} := \{\langle b, a\rangle \ : \ \langle a, b\rangle \in R\}$. (Inverse of R.)
- If A a set, then $R\restriction_A := \{\langle a, b\rangle \ : \ aRb$ and $a \in A\}$ (Restriction of R to A.)
- $R[A] := ran(R\restriction_A)$ (Image of A under R.)
- If S is another relation, then $R \circ S := \{\langle a, b\rangle \ : \ \exists c(aSc \ \& \ cRb)\}$ (Composition of relations.)
- R is *one-to-one* or *injective* or *single-rooted* provided $\forall y \in ran(R)\ \exists! x \in dom(R)(xRy)$

Theorem 4.1 *Let F, G be relations, $\mathcal{A}$ a set. Then: (a) $(F \circ G)^{-1} = G^{-1} \circ F^{-1}$; (b) $F[\cup\mathcal{A}] = \cup\{F[A] : A \in \mathcal{A}\}$; (c) $F[\cap\mathcal{A}] \subseteq \cap\{F[A] : A \in \mathcal{A}\}$; (d) $F[A] - F[B] \subseteq F[A - B]$; (e) $(F^{-1})^{-1} = F$; (f) $F[F^{-1}(y)] \supseteq \{y\}$ for all $y \in ran(F)$*

If F is injective then (c) and (d) are equalities. (f) is an equality if F is a function.

Proof: class and text

4.5 Functions

Definition 4.5 *A* function *is a relation F such that $\forall a \in dom(F)\ \exists! b\ (aFb)$*

- For functions F we usually write $F(a) = b$ instead of aFb. If $A = dom(F)$ and $ran(F) \subseteq B$ then we often write $F\ A \to B$
- Note a relation F is *one-to-one* if and only if F^{-1} is a function.
- In particular, if F is a function then F^{-1} respects all Boolean operations:
  - $F^{-1}[\cup\mathcal{A}] = \cup\{F^{-1}[A] : A \in \mathcal{A}\}$
  - $F^{-1}[\cap\mathcal{A}] = \cap\{F^{-1}[A] : A \in \mathcal{A}\}$
  - $F^{-1}[A] - F^{-1}[B] = F^{-1}[A - B]$
- If F, G are functions, then $F \circ G$ is a function, and $(F \circ G)(x) = F(G(x))$ for every $x \in dom(G) = dom(F \circ G)$

Some remarks on inverting functions:

- If $F\ A \to B$ is a one-to-one function then $F^{-1} \circ F = Id_A$ and $F \circ F^{-1} = Id_{ran(F)}$. Conversely, if there is a function $G\ ran(F) \to A$ satisfying $G \circ F = Id_A$ then F is one-to-one.
- If F is not one-to-one then F^{-1} is a relation that need not be a function; for one or more $a \in ran(F)$, $F^{-1}(a)$ might contain more than one element. Suppose there is a $G\ : ran(F) \to A$ such that $G(a) \in F^{-1}(a)$ for every a. Then G is certainly a "right inverse" in the sense that $F \circ G = Id_{ran(F)}$. Does such a G exist?

Theorem 4.2 *The following are equivalent: (a) AC; (b) If F is a relation, there is a function f such that $dom(f) = dom(F)$ and $f \subset F$*

5 n-ary Relations and Functions, Arbitrary Cartesian Products

Definition 5.1 *The* n-tuple $\langle a_1, \cdots, a_n\rangle$ *is defined by induction:*

$$\langle a_1\rangle := a_1; \qquad \langle a_1, a_2, \ldots, a_{n+1}\rangle := \langle\langle a_1, a_2, \ldots, a_n\rangle, a_{n+1}\rangle$$

- Note that this uses our previous definition for ordered pair, and agrees with it when $n = 1$
- We can now define an *n-ary relation* to be a set of ordered n-tuples.
- If $A_1, \ldots, A_n$ are sets then $A_1 \times A_2 \times \cdots \times A_n$ is defined to be the set of n-tuples $\langle a_1, \ldots, a_n\rangle$ such that $a_i \in A_i$ for all i. Equivalently, $A_1 \times A_2 \times \cdots \times A_n := (\cdots((A_1 \times A_2) \times A_3) \times \cdots) \times A_n$
- We usually denote the n-fold Cartesian product of one fixed set A by $A^n := A \times A \times \cdots \times A$ (n times)
- An ordered n-tuple is also an ordered pair; it follows that an n-ary relation is also a binary relation. Call an n-ary relation f an *n-ary function* if f is a function when viewed as a binary relation; $dom(f)$ and $ran(f)$ are defined accordingly. We often write $f(a_1, \cdots, a_n)$ for $f(\langle a_1, \cdots, a_n\rangle)$

Definition 5.2 *If A, B are sets then ${}^A B$ is the set of functions from A to B.*

Note that any $f\ A \to B$ is a subset of $A \times B$, so ${}^A B$ is a subset of $\mathcal{P}\ (A \times B)$ (and an actual set, by subset selection).

We can view ${}^A B$ as a kind of infinite Cartesian product. For each $a \in A$ let $B_a = B$; then ${}^A B$ corresponds to our intuition for the infinite product $\prod\limits_{a\in A} B_a$.

More generally, let A be a set, and B a function on A. (That is, $B(a)$ can vary with A, as opposed to the constant B we just discussed.) Then $\prod\limits_{a\in A} B(a)$ is the set of all functions f with domain A satisfying $f(a) \in B(a)$ for every $a \in A$. (Note that $\prod\limits_{a\in A} B(a)$ will be a subset of ${}^A[\cup ran(B)]$). Other notations: $\prod\limits_{a\in A} B_a$, $\bigotimes\limits_{a\in A} B(a)$, etc. A is called the *index set* for the product.

6 Equivalence relations

Definition 6.1 *An* equivalence relation *is a relation R which is reflexive, symmetric, and transitive.*

Examples include:

- $\{\langle m, n\rangle \in \mathbb{Z}^2\ :\ m = n \mod p\}$ where p is a prime
- $\{\langle m, n\rangle \in \mathbb{N}^2\ :\ m, n \text{ have the same prime factors}\}$

- $\{\langle x, y\rangle \in \mathbb{R}^2 \ : \ x - y \in \mathbb{Q}\}$

- $\{\langle \overline{u}, \overline{v}\rangle \in \mathbb{R}^{2n} \ : \ (\overline{u} - \overline{v}) \perp \overline{w}\}$ where $\overline{w} \in \mathbb{R}^n$ is a fixed vector.

Fix an equivalence relation R.

- Note $dom(R) = ran(R)$.

- For any x, write $[x]_R$ for the *equivalence class* of x (with respect to R), $[x]_R := \{y \ : \ xRy\}$. Note that $[x]_R \neq \emptyset$ if and only if $x \in dom(R)$

- For any x, y either $[x]_R = [y]_R$ or $[x]_R \cap [y]_R = \emptyset$ (Why?) It follows that $dom(R)$ can be written as the union of a set of pairwise disjoint sets, $dom(R) = \{[x]_R \ : \ x \in dom(R)\}$.

- If X is any set, a *partition* of X is a set P of sets such that (a) $X = \cup P$ and (b) $\forall x, y \in P(x = y \vee x \cap y = \emptyset)$. We say that P *partitions* X. If P and R are as in the last paragraph, we say that P is the partition *induced* or *generated* by the equivalence relation R, and write X/R (or $X \mod R$) for the set of equivalence classes (which is of course the same as the set of elements of the partition).

- Conversely, suppose that P is a partition of a set X. We can define a relation R on X by xRy if and only if x and y are in the same element of P; that is, $R := \{\langle x, y\rangle \in X^2 \ : \ \exists w \in P(x, y \in w)\}$ **Exercise:** Show that this is an equivalence relation. (We say that R is the equivalence relation *induced* or *generated* by P.)

- If R is an equivalence relation on X, and f is a function with domain X, then f is *compatible* with R provided $\forall x, y \in X\ (xRy \implies f(x) = f(y))$ Note that every function f compatible with R induces a function on X/R, and vice versa.

- Example: If $xRy \iff x = y \mod 6$ on $X = \mathbb{Z}$ then $f(x) = x^2 \mod 6$ is compatible with R (why?), but $f(x) = [x/2] \mod 6$ is not (where $[n]$ is the greatest integer less than or equal to n).

- One can similarly say that the n-ary function f is compatible with R provided whenever $x_1Ry_1, \ldots, x_nRy_n$ we have $f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n)$.

- Most often we will be looking at functions from X (or X^n) to X, with the goal of inducing a function from X/R to X/R. In such cases say that f is compatible with R provided $\forall x, y \in X\ (xRy \implies f(x)Rf(y))$. (This is the definition of 'compatible' in some texts.)

- Similarly, the *relation* $F \subseteq X^2$ is compatible with an equivalence relation R on X provided whenever xRy and zRw, $xFz \iff yFw$. Example: $<$ is not compatible with $\equiv_6$:="congruence mod 6" on the integers. Of course, in this example we would normally define a relation $<$ on $\mathbb{Z}/\equiv_6$ by writing $[x]_6 < [y]_6$ provided there exist integers m, n, r, s such that $x = 6r + m,\ y = 6s + n,\ 0 \leq m < n < 6$

- Call x a *representative* of the equivalence class $[x]_R$. In the last example, 5 is a representative of the equivalence class $[11]_6$

- The function $f\ X \to X/R$ defined by $f(x) = [x]_R$ is called the *quotient map* for the equivalence relation R. It is obviously compatible with R in the first sense above.

7 Partial orders

Definition 7.1 *A binary relation R with field X is a* partial order *on X provided R is reflexive, antisymmetric, and transitive. R is a* strict *partial order on X if it is irreflexive and transitive. If R is a partial order on X then (X, R) is called a* poset. *R is* linear *(or* total*) on X if it is a connected relation, i.e. any two elements are* comparable *w/r to R.*

- Note that some texts use "partial order" for our "strict partial order". Use the text's definition when working text problems. My convention will be to use symbols like $\leq, \subseteq, \preceq, \sqsubseteq$, etc. for (non-strict) partial orders, and $<, \subset, \prec, \sqsubset$, etc. for strict partial orders.
- *Exercise:* Show that if R is a partial order on X then $R - \{\langle x, x\rangle \ : \ x \in X\}$ is a strict partial order on X. State and prove a reasonable converse.
- $\leq$, $<$ on $\mathbb{R}$
- $\subseteq$ on $\mathcal{P}\ (X)$ for any set X
- $a|b$ on the positive natural numbers.
- A relation R satisfies *trichotomy* on X provided $\forall x, y \in X$ exactly one of xRy, yRx, or $x = y$ hold. For a relation R on X, TFAE: (a) R is a strict linear order; (b) R is transitive and satisfies trichotomy.

Definition 7.2 *Let $(X, \leq)$ be a poset, and $x \in X$.*

- *x is* minimal *if $\nexists y \in X\ (y < x)$*
- *x is* maximal *if $\nexists y \in X\ (y > x)$*
- *x is* minimum *if $\forall y \in X\ (x \leq y)$*
- *x is* maximum *if $\forall y \in X\ (x \geq y)$*
- Note that 'least', 'smallest', etc. are often used in place of 'minimum'; similarly, 'largest', 'greatest', etc. are often used for 'maximum'.
- If it exists, a maximum element is maximal, and is the *unique* maximal element. If it exists, a minimum element is minimal, and is the *unique* minimal element. In the absence of a maximum/minimum, there can be more than one maximal/minimal element. Question: If there is only one maximal element, is it a maximum?
- The ideas of 'minimal' and 'maximal' can be generalized to arbitrary relations. If R is any binary relation (or even any set) and X a set, then say $x \in X$ is *R-minimal on X* provided $\nexists y \in X - \{x\}\ (yRx)$. (Define R-maximal similarly.)

Definition 7.3 *A relation R is* well founded *(on X) provided every nonempty set $D \subseteq X$ contains an R-minimal element.*

- Note mention of X is not necessary in this definition!
- if $\leq$ is a partial order on X, then $\leq$ is well-founded iff for every $Y \subseteq X$, the poset $(Y, \leq)$ has a minimal element.
- Examples: $<$ on $\mathbb{R}$ and on $\mathbb{N}$; $aRb \iff a|b$ on $\mathbb{N} - \{0\}$; $pRq \iff q' = p$ on the set of real polynomials in x.

Definition 7.4 *A relation R is a* wellorder *if it is a well-founded linear order.*

Examples:

- $<$ on $\mathbb{N}$ is a wellorder
- $X = \mathbb{N} \cup \{a_0\},\ x \prec y \iff (y = a_0$ or $x, y \in \mathbb{N}$ and $x < y)$ is a wellorder
- $X = \mathbb{N}^2,\ \langle a, b\rangle \prec \langle c, d\rangle \iff (b < d \ \&\ (b = d \vee a < c))$ is a wellorder
- $\subseteq$ on $\mathcal{P}_{fin}(X)$:=finite subsets of X is well-founded (why?)

Some interesting results and examples:

Theorem 7.1 *Let R be a relation on X. TFAE: (a) R is irreflexive and well founded on X; (b) There is no sequence $\{x_n\}_{n\in\omega} \subseteq X$ such that $\forall n \in \omega(x_{n+1} R x_n)$*

Cor 7.1 *If $R \subseteq S$ are two relations on X and S is well-founded on X then so is R.*

Definition 7.5 *Let (X, R) and (Y, S) be posets, and $\phi\, X \to Y$.*

- *ϕ is a* homomorphism *(from X to Y) provided $\forall x_1, x_2\ (x_1 R x_2 \iff \phi(x_1) S \phi(x_2))$*
- *ϕ is an* isomorphism *(from X to Y) provided it is an injective homomorphism from X to Y.*
- *X and Y are* isomorphic *provided there is an isomorphism from X onto Y.*

Remarks:

- The definitions above can be applied to *any* binary relations R and S. In the case of posets we often emphasize the order by calling ϕ an *order*-homomorphism/isomorphism
- We will use the weaker term *order-preserving function* for a function $\phi X \to Y$ satisfying $\forall x_1, x_2\ (x_1 R x_2 \implies \phi(x_1) S \phi(x_2))$
- Note mention of the order is supressed in the expression "X and Y are isomorphic"; why is this screwy?
- We often write $X \cong Y$ in place of "X and Y are isomorphic"
- Examples:
  - arctan from $(\mathbb{R}, \leq)$ into itself, or onto a subinterval.
  - *id* $(X, \prec_X) \to (Y, \prec_Y)$, where $\prec_X = \prec_Y \restriction_X$ (embedding)
  - $(\mathbb{Q}, \leq)$ and $(((0,1) \cup (3,8)) \cap \mathbb{Q}, \leq)$
- Question: Is order-isomorphism an equivalence relation?
- If X and Y are linearly-ordered sets, and $\phi\, X \to Y$ is order-preserving, then it is an isomorphism.

Proposition 7.1 *If $(X, \leq)$ is a poset then there is a $Y \subseteq \mathcal{P}(X)$ such that $(X, \leq) \cong (Y, \subseteq)$*

Theorem 7.2 Induction on well-founded sets *Suppose R is a well-founded relation on X, and that $E \subseteq X$ is a set such that (i) $\forall x \in X,\ x\ R{-}minimal \implies x \in E$; (ii) $\forall x \in X$, if $\forall y \in X - \{x\}(yRx \implies y \in E)$ then $x \in E$. Then $E = X$*

8 Natural Numbers

Definition 8.1 $a^+ := a \cup \{a\}$ *(= the* successor *of a)*

A set A is inductive *provided (1)* $\emptyset \in A$*, and (2)* $\forall a \in A\ (a^+ \in A)$

$\omega := \bigcap\{A\ :\ A \textit{ is an inductive set}\}$ *(= the* natural numbers*)*

$0 := \emptyset;\ 1 := 0^+ = \{0\} = \{\emptyset\};\ 2 := 1^+ = 1 \cup \{1\} = \{0, 1\} = \{\emptyset, \{\emptyset\}\};\ \ldots$

- Note that the statement "A is inductive" is a formula in the language of set theory.

- Recall that the Axiom of Infinity gives us a set E_∞ such that for every $x \in E_\infty$, $x^+ \in E_\infty$. Thus, E_∞ is an inductive set. One could have waited until now to define state Axiom of Infinity, and just taken it to be the existence of an inductive set.

- **Exercise:** If X is a set of inductive sets, then $\cap X$ is inductive.

- For the definition of ω above to work, one must make sure that the intersection is over a set, not a class. One can replace the above with $\omega := \bigcap\{A \in \mathcal{P}(E_\infty)\ :\ A \text{ is an inductive set}\}$ This is certainly a set (subset selection). Using the previous exercise, convince yourself that this is the same thing as $\bigcap\{A\ :\ A \text{ is an inductive set}\}$.

- ω is the *smallest* inductive set, in the sense that if A is any other inductive set then $\omega \subset A$. Why?

- As defined, $0, 1, 2, \cdots \in \omega$. Note that we could in general define $n = \{0, 1, \ldots, n-1\}$. It therefore makes sense to use $<$ and $\in$ interchangeably on ω. It follows that $\forall n \in \omega\ \forall x\ (x < n \iff x \in n)$ *trivially.*

- We would like to verify that our (intuitive) natural numbers map bijectively onto ω, and the order is the same. This follows from the following **Exercise:** Define f inductively on $\mathbb{N}$ by $f(0) := \emptyset$, $f(n+1) := \{f(0), \ldots, f(n)\}$. Then $\omega = range(f)$

8.1 Ordering on ω

Goal:

Theorem 8.1 *$(\omega, <)$ is a well-ordered set.*

(Proof later)

Definition 8.2 *A set A is* transitive *provided $\forall a \in A \; a \subseteq A$*

- Note that A is transitive provided $x \in a \in A \implies x \in A$. (Other equivalents: $\bigcup A \subseteq A, \; A \subseteq \mathcal{P}(A), \dots$)
- This is not quite the same thing as transitivity for a relation - get the specific meaning from context.
- **Lemma:** If a is transitive then $\bigcup(a^+) = a$
- **Corollary:** $\forall n \in \omega, \; n$ is transitive.
- **Lemma:** ω is transitive

Now we proceed to show that $(\omega, <)$ is a woset. We need to show that $<$ is (1) irreflexive (since it is meant to be a *strict* partial order on ω), (2) transitive, (3) linear, and (4) well-founded (hence a well-order) on ω. Note that both (1) and (4) follow from the Axiom of Foundation, but it is interesting to see that this axiom isn't really necessary for $\in$ on ω.

1. **Irreflexive:** To show that $x \in x$ never happens, let $E = \{x \in \omega \ : \ x \notin x\}$. $\emptyset \in E$ since *nothing* is in the empty set, let alone itself. Suppose $n \in E$; is $n^+ \in E$? If not, then $n^+ \in n^+ = n \cup \{n\}$, so $n^+ \in n$ or $n^+ = n$, so $n \in n^+ \in n$ or $n \in n^+ = n$. Either way, $n \in n$ (we're using the result that n is transitive), contradicting our hypothesis, so in fact $n^+ \in E$. Then E is an inductive subset of ω, so is all of ω.

2. **Transitive:** $x < y \& y < z \implies x \in y \in z$ which implies $x \in z$ by transitivity of $z \in \omega$.

3. **Linear:** We need to show that any two elements are comparable. We start with a weaker assertion. Some notation: for $x \in \omega$ write $[0, x] := \{y \in \omega \ : \ y \leq x\}$

   (Do NOT infer from this notation that $[0, x]$ is a segment, or is itself linearly ordered, or even includes 0! In fact, this is just a suggestive way of writing x^+.)

   First, we'll show that $<$ satisfies trichotomy on $[0, x]$ for any x.

   Put $E = \{x \in \omega \ : \ [0, x] \text{ satisfies trichotomy}\}$.

   Since $[0, 0] = \{\emptyset\}$, $\emptyset \in E$. Suppose $n \in E$. If $x, y \in [0, n^+]$ then there are 3 cases (why?):

   (a) $x, y \leq n$: Then $x, y \in [0, n]$ so x and y are comparable.

   (b) $x = n^+ = y$: Then $x = y$, so they're comparable.

   (c) $x \leq n$ and $y = n^+$, or $y \leq n$ and $x = n^+$: For definiteness, assume the first case. Then $x \in n \in n^+$, so $x < y$.

   Either way, they're comparable; it follows that $n^+ \in E$

   This proves that $<$ satisfies trichotomy on $[0, x]$ for any x; we'll return to the full statement of linearity later.

4. **Well-founded:** This one is harder than it looks! First, we prove a

   Proposition 8.1 *0 is the least element of* ω

   Proof: Put $E = \{x \in \omega \ : \ 0 \leq x\}$. $0 \in E$, and if $n \in E$ then $0 = n \in n^+$ or $0 \in n \in n^+$, either way $0 \leq n^+$. So, E is inductive.

   Now, let $D \subseteq \omega$ be nonempty, and suppose (for a contradiction) that D has no least element. Put $E = \{x \in \omega \ : \ [0, x] \subseteq \omega - D$. $0 \in E$ since 0 is least in all of ω. Let $n \in E$, and suppose $n^+ \notin E$. Then n^+ must be least in D, as any $x < n^+$ is in $[0, n]$, a contradiction.

5. **Linearity (concluded):** We need another proposition.

   Proposition 8.2 *$\forall n \in \omega, n = 0$ or $\exists m \in \omega (n = m^+)$*

   Proof: Put $E = \{n \in \omega : n = 0$ or $\exists m \in \omega(n = m^+)$. It is very easy to see that E is inductive, which proves the proposition.

   Now, let $x, y \in \omega$, and assume they are not comparable. Let n least in $[0, x] \setminus [0, y]$; this set is nonempty since x and y are not comparable, and a least element exists since $[0, x]$ is linearly ordered by the above, and $<$ is well-founded on this set. Similarly, let m least in $[0, y] \setminus [0, x]$. Since $m, n \neq 0$ there are $a, b \in \omega$ with $m = a^+, n = b^+$. By minimality of m and n, $a, b \in [0, x] \cap [0, y]$ In particular, they are comparable since the ordering here is linear. Suppose for definiteness $a \leq b$. b must also be less than m (why? - draw picture!), so $a \leq b < m = a^+$, so $b = a$, but then $m = n$, contradiction. This finishes the proof of linearity, and thus finishes the proof that $<$ (ie, $\in$) well-orders ω.

8.2 Induction on ω

The principal of induction for well-founded sets, applied to $(\omega, <)$ now becomes:

Theorem 8.2 *Let $E \subseteq \omega$ satisfy (i) $0 \in E$, and (ii) $\forall x \in \omega,\ [0, x) \subseteq E \implies x \in E$. Then $E = \omega$*

(Note the the notation $[0, x)$ for $\{y \in \omega \ : \ y < x\}$, which of course just equals x!)

Usually, E will be $\{x \in \omega \ : \ \phi(x)\}$ for some formula of $\mathcal{L}_{\ \in}$. The result becomes:

Cor 8.1 *Let $\phi(x)$ be a formula, and suppose (i) $\phi(0)$, and (ii) $\forall x \in \omega$, if $\phi(y)$ holds for all $y < x$ then $\phi(x)$ holds. Then $\forall x \in \omega\ \phi(x)$*

This is sometimes called *strong induction.* "Standard" induction holds as well:

Cor 8.2 *Let $E \subseteq \omega$ satisfy (i) $0 \in E$, and (ii) $\forall x \in X,\ x \in E \implies x^+ \in E$. Then $E = \omega$*

Cor 8.3 *Let $\phi(x)$ be a formula, and suppose (i) $\phi(0)$, and (ii) $\forall x \in \omega$, $\phi(x) \implies \phi(x^+)$. Then $\forall x \in \omega\ \phi(x)$*

Proof: Exercise.

Remark: It is occasionally useful to allow *parameters* in ϕ; that is, $\phi(x; a_1, \ldots, a_m)$ refers to some fixed sets $a_1, \ldots, a_m$.

8.3 Recursion

We *prove* results in ω using induction; we *construct* objects from ω using *recursion.*

- EG: Fibonacci numbers
- EG: X^n
- EG: $X^n \leftrightarrow {}^nX$

In each case above, the function's value at n was a function of one *or more* earlier values. In other words, we start with a function G whose domain includes any function g whose domain is an initial segment from ω (i.e., $dom(g) = n$ for some $n \in \omega$.) To define f on ω, we specify $f(0)$, then put $f(n) := G(f\restriction_n)$. We then assert the existence of a unique function f satisfying these conditions for all $n \in \omega$. Before proceeding to make this precise, we note a couple of potential problems:

- There is *no* function G whose domain includes any function g whose domain is an initial segment from ω! There are two ways around this:

  1. Probably we already know the intended range Z of f. In this case, we can just require that the recursion start with a function $G\,\{{}^{<\omega}Z\} \to Z$, where ${}^{<\omega}Z := \bigcup_{n\in\omega} {}^nZ$.
  2. Occasionally we might only have an idea of how, given $f(0), f(1), \ldots, f(n-1)$, to construct $f(n)$ (or even – weaker! – to assert that such an $f(n)$ exists). We can formalize this by requiring (in place of the G above) a formula $\gamma(x, y)$ such that $\forall x\ \exists! y\ \gamma(x, y)$.

- Our G (or γ) might only make sense for some subset W of ${}^{<\omega}Z$; in that case we can define G arbitrarily off W, then prove inductively that the f we get satisfies $\forall n\ f\restriction_n \in W$.

So, here are versions of *Definition by Recursion*, in increasingly general form:

Theorem 8.3 *(Recursion on ω)Let Z be a set and G be a function from ${}^{<\omega}Z$ to Z. Then there is a unique function $f\ :\omega \to Z$ such that $f(n) = G(f\restriction_n)$ for every $n \in \omega$*

Here ${}^{<\omega}Z$ is the set of finite sequences from Z. Note that $f\restriction_n$ is such a sequence, and that G can be a function of n as well (since the definition of G can certainly include reference to the domain of $f\restriction_n$).

For the more general version, let $\leq$ be a well-founded relation on a set X. (Note that my use of the $\leq$ symbol here does *not* mean that it is a partial order, though it usually will be the case in applications.) For $x \in X$ write $\mathrm{seg}(x) := \{y \in X\ :\ y < x\}$ (note $x \notin \mathrm{seg}(x)$).

Theorem 8.4 *(Recursion on a well-founded relation) Suppose $\leq$ is a well-founded relation on a set X, Z another set, and G is a function from $(\bigcup_{x\in X} {}^{\mathrm{seg}(x)}Z)\times X$ to Z. Then there is a unique function $f\ : X \to Z$ such that $f(x) = G(f\restriction_{\mathrm{seg}(x)}, x)$ for every $x \in X$.*

Finally, the most general version:

Theorem 8.5 *(Recursive construction on a well-founded relation) Suppose $\leq$ is a well-founded relation on a set X, and $\gamma(w, x, y)$ is a formula (possibly with parameters) such that for every $x \in X$ and function f with domain seg(x) there is exactly one z such that $\gamma(f, x, z)$ holds. Then there is a unique function f with domain X such that $\gamma(f\restriction_{\mathrm{seg}(x)}, x, f(x))$ holds for every $x \in X$.*

8.4 Arithmetic

We now give recursive definitions of the usual operations on the natural numbers:

Definition 8.3 *Addition:* $x + 0 := x; \quad x + (n^+) := (x + n)^+$

- Note that this actually defines, for any $x \in \omega$, a unary function $A_x \ : \omega \to \omega$; we then define the binary $+ \ : \omega^2 \to \omega$ by $+(m, n) := A_m(n)$
- What is G here?
- For any x, A_x is a proper set-theoretic object (by the recursion result). What about $+$?

Definition 8.4 *Multiplication:* $m \cdot 0 := 0; \quad m \cdot (n^+) := (m \cdot n) + m$

Definition 8.5 *Exponentiation:* $m^0 := 1; \quad m^(n^+) := (m^n) \cdot m$

Theorem 8.6 $1 + 1 = 2$

Proof: $1 + 1 = 1 + 0^+ = (1 + 0)^+ = 1^+ = 2$

Theorem 8.7 *Properties of $+$ and $\cdot$:* $\forall m, n, p \in \omega$,

1. $m + n = n + m$
2. $m + (n + p) = (m + n) + p$
3. $m \cdot n = n \cdot m$
4. $m \cdot (n \cdot p) = (m \cdot n) \cdot p$
5. $m \cdot (n + p) = m \cdot n + m \cdot p$

Proof of (1): Class. The rest are similar

8.5 Peano Arithmetic

A common way to characterize arithmetic on the natural numbers is by specifying a set of basic operations, and listing a set of axioms the operations should obey. For the purposes of this section, take as basic the zero element and *successor* operation, $\mathbf{0}, \quad \mathbf{S}(x)$. The *Peano Postulates* are then:

1. $\forall x\ \mathbf{S}(x) \neq \mathbf{0}$
2. $\forall x, y\ \mathbf{S}(x) = \mathbf{S}(y) \implies x = y$
3. $\forall y\ y \neq \mathbf{0} \implies \exists x\ y = \mathbf{S}(x)$
4. (Induction schema) $\forall z\ (\phi(\mathbf{0}, z)\ \wedge\ \forall x\ (\phi(x, z) \implies \phi(\mathbf{S}(x), z)) \implies \forall x\ \phi(x, z)$ for every formula ϕ in the language of $\mathbf{S}$ and $\mathbf{0}$.

(Note that the Peano Postulates are often written in a language including including $+$ and possibly $\cdot$, and include the postulates which define these operations. We will revisit them in Math 455.)

It is easy to see that $\omega,\ 0$, and the operation $x \mapsto x^+$ satisfy these axioms:

1. Since $0 \in x^+$ for any x, and $0 \notin 0$
2. $n < m \implies n^+ < m^+$: Else $m \in m^+ \leq n^+ = n \cup \{n\}$, so $m \in n$ or $m = n$, contradicting trichotomy.
3. This is an easy induction: let $E = \{n \in \omega\ :\ n \neq 0 \implies \exists m\ n = m^+\}$, etc.
4. This is a form of induction, as already discussed. (Note the parameters z.)

Finally, attention is called to some of the other order and arithmetic properties of $(\omega, <)$, as discussed in this chapter of the text. They are all relatively straightforward applications of properties we've already proved here.

One we will need is the following:

Lemma 8.1 *If $m < n$ then for some p, $m + p = n$.*

8.6 Integers

Once we have the natural numbers, the integers, rationals, reals, etc. are all quite easy.

Definition 8.6 *For $a, b, c, d \in \omega$ write $\langle a, b\rangle \sim \langle c, d\rangle$ provided $a + d = b + c$*

Lemma 8.2 *$\sim$ is an equivalence relation on ω^2.*

Definition 8.7 *$\mathbb{Z} := \omega^2/\sim$ (the integers)*

The class $[\langle a, b\rangle]_\sim$ is meant to represent our intuitive idea of the integer $a-b$.

There are plenty of other ways we could have formalized this idea, e.g. we could have looked at ordered pairs $\langle a, b\rangle$ with $b \in \omega$ and $a = 0$ or 1 (with 0 meaning "nonnegative" and 1 meaning "negative"). The advantages of the given approach are (a) operations can be defined by simgle formulas, and (b) it captures the philosophical idea that we extend ω primarily so as to be able to subtract.

Lemma 8.3 *The following functions and relation are well-defined:*

- $[\langle a, b\rangle]_\sim + [\langle c, d\rangle]_\sim := [\langle a + c, b + d\rangle]_\sim$
- $[\langle a, b\rangle]_\sim \cdot [\langle c, d\rangle]_\sim := [\langle ac + bd, bc + ad\rangle]_\sim$
- $-[\langle a, b\rangle]_\sim := [\langle b, a\rangle]_\sim$
- $[\langle a, b\rangle]_\sim < [\langle c, d\rangle]_\sim \iff (a + d <_\in b + c)$

The proof is a tedious exercise; see also the text.

Write $0 := [\langle 0, 0\rangle]_\sim,\ 1 = [\langle 1, 0\rangle]_\sim$

Theorem 8.8 *Properties of $+$, $\cdot$, and $<$: $\forall x, y, z \in \mathbb{Z}$,*

1. $x + 0 = x$
2. $x \cdot 1 = x$, $x \cdot 0 = 0$
3. $x + y = y + x$
4. $x + (y + z) = (x + y) + z$
5. $x \cdot y = y \cdot x$
6. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
7. $x \cdot (y + z) = x \cdot y + x \cdot z$
8. $-x = 0 - x = (-1) \cdot x$
9. $x + (-x) = 0$
10. *$<$ is a linear order*

11. $x < y \implies x + z < y + z$

12. $x < y \wedge z > 0 \implies xz < yz$

13. $x < y \wedge z < 0 \implies xz > yz$

The proof follows from the corresponding properties of the operations on ω, and some tedious work.

Lemma 8.4 *The function $\psi : x \mapsto [\langle x, 0\rangle]_\sim$ satisfies the following:*

- *it is order-preserving from $(\omega, <)$ to $(\mathbb{Z}, <)$, with range $\{x \in \mathbb{Z} : x \geq 0\}$.*
- *It preserves addition and multiplication in the sense that for every $m, n \in \omega$, $\psi(m) + \psi(n) = \psi(m + n)$ and $\psi(m) \cdot \psi(n) = \psi(m \cdot n)$*

In light of the last lemma, we can view $\mathbb{Z}$ as an extension of ω, and *identify* each $n \in \omega$ with $[\langle n, 0\rangle]_\sim$, writing the latter n as well.

8.7 Rationals

On to the rationals:

Definition 8.8 *For $a, b, c, d \in \mathbb{Z}$ write $\langle a, b\rangle \sim \langle c, d\rangle$ provided $a \cdot d = b \cdot c$*

Lemma 8.5 *$\sim$ is an equivalence relation on $\mathbb{Z}^2$.*

Definition 8.9 $\mathbb{Q} := \{\langle a, b\rangle \in \mathbb{Z}^2 \ : \ b \neq 0\}/\sim$ *(the rational numbers)*

The class $[\langle a, b\rangle]_\sim$ is meant to represent our intuitive idea of the fraction $\frac{a}{b}$.

Lemma 8.6 *The following functions and relation are well-defined:*

- $[\langle a, b\rangle]_\sim + [\langle c, d\rangle]_\sim := [\langle ad + bc, bd\rangle]_\sim$
- $[\langle a, b\rangle]_\sim \cdot [\langle c, d\rangle]_\sim := [\langle ac, bd\rangle]_\sim$
- $-[\langle a, b\rangle]_\sim := [\langle -a, b\rangle]_\sim$
- $[\langle a, b\rangle]_\sim/[\langle c, d\rangle]_\sim := [\langle ad, bc\rangle]_\sim$ *provided $c \neq 0$*
- $[\langle a, b\rangle]_\sim$ *is* positive *provided $ab > 0$*
- $[\langle a, b\rangle]_\sim < [\langle c, d\rangle]_\sim \iff [\langle c, d\rangle]_\sim - [\langle a, b\rangle]_\sim$ *is positive*

(Note the use of $x - y$ for $x + (-y)$.)

The proof is another tedious exercise.

Write $0 := [\langle 0, c\rangle]_\sim$, $1 = [\langle c, c\rangle]_\sim$, and note that neither depends on the choice of $c \neq 0$.

Theorem 8.9 *Properties of $+$, $\cdot$, and $<$: $\forall x, y, z \in \mathbb{Q}$,*

1. $x + 0 = x$
2. $x \cdot 1 = x,\ x \cdot 0 = 0$
3. $x + y = y + x$
4. $x + (y + z) = (x + y) + z$
5. $x \cdot y = y \cdot x$
6. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
7. $x \cdot (y + z) = x \cdot y + x \cdot z$
8. $-x = 0 - x = (-1) \cdot x$
9. $x + (-x) = 0$
10. *$<$ is a linear order*

11. $x < y \implies x + z < y + z$

12. $x < y \wedge z > 0 \implies xz < yz$

13. $x < y \wedge z < 0 \implies xz > yz$

The proof follows from the corresponding properties of the operations on $\mathbb{Z}$, and some tedious work.

Lemma 8.7 *The function* $\psi : \ x \mapsto [\langle x, 1\rangle]_\sim$ *satisfies the following:*

- *it is order-preserving from* $(\mathbb{Z}, <)$ *to* $(\mathbb{Q}, <)$
- *It preserves addition and multiplication in the sense that for every* $m, n \in \omega$, $\psi(m) + \psi(n) = \psi(m + n)$ *and* $\psi(m) \cdot \psi(n) = \psi(m \cdot n)$

9 Real Numbers

The choice of construction for the Real numbers depends on what we decide is the major reason one wants to invent $\mathbb{R}$ rather than sticking with $\mathbb{Q}$. Here are some commonly-cited reasons:

1. $\mathbb{R}$ is metrically *complete* - that is, every Cauchy-convergent sequence actually converges to something.
2. $\mathbb{R}$ satisfies the *LUB Property*: every subset of $\mathbb{R}$ which is bounded above has a least upper bound.
3. $\mathbb{R}$ is *Real-closed*: polynomials of odd degree have roots, and nonnegative elements have square roots.

It turns out that (1) and (2) are equivalent, and that (3) is strictly weaker, so it makes sense to use (1) or (2) as our guiding intuition.

One common way to construct $\mathbb{R}$ is to let X be the set of Cauchy-convergent sequences from $\mathbb{Q}$, to define an equivalence relation $\sim$ on X by $(s_n) \sim (t_n) \iff \lim\limits_{n\to\infty}(s_n - t_n) = 0$, and to take $\mathbb{R} = X/\sim$. This works well, and it is easy to prove from this construction that $\mathbb{R}$ is complete.

Probably I'll give you some extra exercises on this.

We'll instead use (2) to guide us.

Definition 9.1 *A* cut *(or* Dedekind cut*) in $\mathbb{Q}$ is a set $A \subset \mathbb{Q}$ such that*

1. *$A \neq \emptyset$*
2. *A is bounded above, that is, $\exists M \in \mathbb{Q}\ \forall x \in A\ x < M$*
3. *A has no greatest element, that is, $\forall x \in A\ \exists y \in A\ x < y$*
4. *A is an initial segment, $\forall x \in A\ \forall y < x\ y \in A$*

Remarks:

- Other definitions of Dedekind Cut
- Intuitively, the real number x is the cut $(-\infty, x) \cap \mathbb{Q}$

Definition 9.2 *Write $\mathbb{R}$ for the set of cuts of $\mathbb{Q}$, together with operations:*

1. $A + B := \{a + b \ : \ a \in A, b \in B\}$

2. $0_{\mathbb{R}} := \{a \in \mathbb{Q} \ : \ a < 0\}$; $1_{\mathbb{R}} := \{a \in \mathbb{Q} \ : \ a < 1\}$

3. $A <_{\mathbb{R}} B \iff A \subsetneq B$

4. $A \cdot B := 0_{\mathbb{R}} \cup \{a \cdot b \ : \ a \in A, b \in B, a, b \geq 0\}$ *if* $0 \leq A, B$

Remarks:

- As we have no equivalence relations, there is no need to show that anything is well-defined.
- Warning: Some operations must have tricky definitions. For example, had we tried to define $A \cdot B$ to be $a \cdot b \ : \ a \in A, b \in B$, what would $0 \cdot 0$ equal? Similarly, we cannot define $-A := \{-a \ : \ a \in A\}$, as that is not a cut! Similarly, see our definition above for multiplication.

Definition 9.3 *Some more operations on* $\mathbb{R}$*:*

1. $-A := \{a \in \mathbb{Q} \ : \ \exists x > a \ \ -x \notin A\} = \{-x \in \mathbb{Q} \ : \ \exists y < x \ y \in \mathbb{Q} \setminus A\}$

2. $sgn(A) := \begin{cases} 1 & if \ \exists a \in A \ a > 0 \\ 0 & if \ A = 0_{\mathbb{R}} \\ -1 & otherwise \end{cases}$

3. $|A| := A \cup -A$

4. *If* $A = 0$ *or* $B = 0$ *then* $A \cdot B := 0_{\mathbb{R}}$

5. $A \cdot B := \begin{cases} |A| \cdot |B| & if \ sgn(A) = sgn(B) \\ -(|A| \cdot |B|) & otherwise \end{cases}$

6. $-1_{\mathbb{R}} := \{a \in \mathbb{Q} \ : \ a < -1\}$

7. *If* $0_{\mathbb{R}} <_{\mathbb{R}} A$, $A^{-1} := 0_{\mathbb{R}} \cup \{1/a : \ a \in A, 0 <_{\mathbb{Q}} a\}$

8. *If* $0_{\mathbb{R}} >_{\mathbb{R}} A$, $A^{-1} := -((-A)^{-1})$

9. *If* $B \neq 0_{\mathbb{R}}$, $A/B := A \cdot B^{-1}$

Remarks:

- Why not just put $-A := \mathbb{R} \setminus \{-a \ : \ a \in A\}$?
- We could instead have defined, for $0_{\mathbb{R}} <_{\mathbb{R}} A$, $A^{-1} := \{1/a : \ a \in A, a \neq 0_{\mathbb{Q}}\} \cup \{0\}$

Theorem 9.1 $\mathbb{R}$*, together with the constants and operations defined above, forms an ordered field.*

The proof is even more tedious than the ones for $\mathbb{Q}$ and $\mathbb{Z}$.

Theorem 9.2 $\mathbb{R}$ *has the LUB property*

Proof:

Cor 9.1 $\mathbb{R}$ *is complete*

Proof:

Lemma 9.1 *The function* $\psi : \mathbb{Q} \to \mathbb{R}$ *defined by* $\psi(x) := \{a \in \mathbb{Q} \ : \ a <_{\mathbb{Q}} x\}$ *satisfies:*

- *it is order-preserving from* $(\mathbb{Q}, <_{\mathbb{Q}})$ *to* $(\mathbb{R}, <_{\mathbb{R}})$
- *It preserves* 0, 1, *addition and multiplication in the sense that* $\psi(0) = 0$, $\psi(1) = 1$, *and for every* $m, n \in \omega$, $\psi(m) + \psi(n) = \psi(m + n)$ *and* $\psi(m) \cdot \psi(n) = \psi(m \cdot n)$

Proof:

Remark: $\mathbb{R}$ is the unique *complete* ordered field.

10 Cardinals

Notation: Write ω for the natural numbers (usually $\mathbb{N}$, but for technical reasons it is useful to have the alternate notation). $<_\omega$ is the usual order on ω.

Recall the following notions: *function, relation, one-to-one, onto, on-to-one correspondence, injection, surjection, bijection, equivalence relation,* A^n, ${}^AB$

- If A B are two ordered sets, we can ask:

  1. Do A, B have the same order structure?
  2. Do A, B have the same number of elements?

- Note that the second question is more general (doesn't really refer to the order structure).

- An affirmative answer to the first is an affirmatve answer to the second, since an order isomorphism is *a fortiori* a bijection.

- The converse is not true, for example $(\omega, <_\omega)$ and $(\omega, {<_\omega}^{-1})$ are not order-isomorphic (one has a least element, one a greatest) but the base sets are the same.

In this section I will formalize the notion of "have the same number of elements;" later we will show that this is really a notion involving order structures and order isomorphism!

Definition 10.1 *A is* equinumerous *with B provided there is a bijection from A onto B.*

Other, equivalent notation/terminology for "A is equinumerous with B":

a. A and B *have the same cardinality*

b. $A \approx B$

c. $\mathrm{card}(A) = \mathrm{card}(B)$

We seem to be referring to a concept called "cardinality" without really defining it; this is especially true for (c), where the notation *looks* like we are setting two functions equal to one another. For the time being, we should view all the above as simply complicated ways of expressing a formula in the two sets A and B. Later we ill indeed define a function-like operation card() which will assign a "cardinal number" to ever set; however, this will have to wait until after we introduce the more subtle concept of *ordinal numbers*. Note that when we finally *do* introduce this card(), it won't be a function, since otherwise dom(card) would be the set of all sets!

Some examples:

1. $\omega \approx \mathbb{Z}$ (Proof: Let $f(x) = 2x$ if $x \geq 0$, $f(x) = -2x - 1$ otherwise; this is evidently a bijection from $\mathbb{Z}$ to ω)

2. $\omega \approx \omega^2$ (Proof 1: traverse the square array, invoke recursive definition. Proof 2: Define $J : \omega^2 \to \omega$ by $J(m, n) = [(m + n)^2 + 3m + n]/2$, show a bijection. Proof 3: The map $(m, n) \mapsto 2^m 3^n$ is clearly an injection. This shows that in some sense $omega^2$ has "no greater" cardinality than ω. Does this suffice?)

3. $\mathbb{Q}^{>0} \approx \omega$ (Proof 1: Array $\mathbb{Q}$ in a matrix, traverse skipping duplicates.)

4. $\mathbb{Q} \approx \omega$ (Proof: Let ϕ be a bijection from ω onto $\mathbb{Z}$ which takes 0 to 0, and ψ a bijection from ω to $\mathbb{Q}^{>0}$, then $x \mapsto \text{sign}(\phi(x))\psi(|\phi(x)|)$ will map ω bijectively onto $\mathbb{Q}$.

The following easy theorem proves that $\approx$ is "like" an equivalence relation. (Why isn't "$\approx$" an actual equivalence relation?)

Theorem 10.1 *For any sets A, B, and C:*

1. $A \approx A$

2. $A \approx B \implies B \approx A$

3. $A \approx B \, \& B \approx C \implies A \approx C$

Proof: class and text.

It will be convenient to have a slightly more general notation:

Definition 10.2 *A has no greater cardinality than B (or $\text{card}(A) \leq \text{card}(B)$) provided there is an injection from A into B.*

Theorem 10.2 *For any nonempty sets A and B the following are equivalent:*

1. $\text{card}(A) \leq \text{card}(B)$

2. For some $C \subseteq B$, $\quad \text{card}(A) = \text{card}(C)$

3. There is a function g from B onto A.

Proof: Class

Remark: If $A = \emptyset$ then the first two of these statements are trivially true, while the third is trivially false.

The following very deep theorem will be proved in more general form later.

Theorem 10.3 (Cantor-Schroder-Bernstein)*If* card$(A) \leq$ card(B) *and* card$(B) \leq$ card(A) *then* card$(A) =$ card(B)

Proof: later.

Theorem 10.4 *(Cantor)For any set A,* card$(A) \neq$ card$(\mathcal{P}\ (A))$ *(in fact,* card$(A) <$ card$(\mathcal{P}\ (A))$

Proof: Let g be any $1-1$ finction from A into $\mathcal{P}\ (A)$. Put $B = \{x \in A\ : x \notin g(x)\}$ Claim: $B \notin range(g)$. To see this, let $x_0 \in A$, and consider 2 cases: (i) $x_0 \in g(x_0)$; then $x_0 \notin B$, so $B \neq g(x_0)$. (ii) $x_0 \notin g(x_0)$; then $x_0 \in B$, so $B \neq g(x_0)$. Either way, $B \neq g(x_0)$. Since x_0 was arbitrary in A, $B \notin range(g)$. Since g was arbitrary, $A \not\approx \mathcal{P}\ (A)$. (Note that $a \mapsto \{a\}$ is an injection from A into $\mathcal{P}\ (A)$, so card$(A) \leq$ card$(\mathcal{P}\ (A))$.)

Proposition 10.1 *For every set A, $\mathcal{P}\ (A) \approx {}^A2$*

Proof: Define $f : \mathcal{P}\ \to {}^A2$ by $f(\alpha) = \chi_\alpha$, where $\chi_\alpha(x) = 1$ if $x \in \alpha$, $= 0$ otherwise (i.e., χ_α is the characteristic function of the set α). Claim: f is a bijection. $1-1$: If $\alpha \neq \beta$ are in $\mathcal{P}\ (A)$ then WOLG there is some $n \in \beta - \alpha$. Then $\chi_\alpha(n) = 0, \chi_\beta(n) = 1$, so $f(\alpha) \neq f(\beta)$. Onto: If $\phi \in {}^A2$ then put $\alpha = \phi^{-1}(1)$. Then $\chi_\alpha(n) = 1 \iff n \in \alpha \iff \phi(n) = 1$, so $f(\alpha) = \phi$.

Cor 10.1 $\omega \not\approx \mathbb{R}$

Proof 1: Text Thm. 6B(a). Note how unsatisfying this is (one needs to develop all the machinery of convergence of series for real numbers)
Proof 2: It suffices to show that ${}^\omega2$ can be embedded into $\mathbb{R}$ in a $1-1$ way. One way is to take the sequence $\tau \in {}^\omega2$ to the real number $\sum\limits_{n=0}^{\infty} \frac{2\tau(n)}{3^{n+1}}$. Again, one needs to show that this series converges and that the map is 1-1. (Remark: the set of real numbers obtained this way is the usual "Cantor middle-thirds set".)
Proof 3:A direct "Dedekind cut" argument is possible. We'll see this after we define "Dedekind cuts".
Remark: Proof 2 uses that ${}^\omega2 \not\approx \omega$, which of course follows from results above. However, we can prove this directly, emulating the proof of Cantor's Theorem. Suppose $g\ :\ \omega \to {}^\omega2$; for any n, $g(n)$ will be an infinite sequence $g(n) = (x_0^n, x_1^n, x_2^n, \ldots, x_n^n, \ldots)$ with each $x_i^n = 0$ or 1. We can identify this sequence with the subset of $A_n \subseteq \omega$ defined by $i \in A_n \iff x_i^n = 1$. Then it is easy to see that the diagonal sequence $(x_0^0, x_1^1, x_2^2, \ldots, x_n^n, \ldots)$ is exactly the set A we produced in the earlier proof.

10.1 Finite sets

Definition 10.3

- *A is* finite *if* $\exists m \in \omega\ A \approx m$, not finite *otherwise*
- *A is* infinite *if* card$(\omega) \leq$ card(A)*,* not infinite *otherwise*

- *A is* Dedekind-infinite *if $\exists B \subsetneq A$ $A \approx B$,* Dedekind-finite *otherwise*

Theorem 10.5 *Suppose $A \neq \emptyset$. The following are equivalent:*

1. *A is finite*
2. *A is not infinite*
3. *A is Dedekind-finite*
4. *There exists an injection $f : A \to m$ for some $m \in \omega$*
5. *There exists a surjection $g : m \to A$ for some $m \in \omega$*

Proof: (1) $\implies$ (4) (and (5)) is immediate.

(4)$\Leftrightarrow$(5) is just like the proof of Theorem 10.2.

(3) $\implies$ (2) If A is infinite then there is an injection $\phi : \omega \to A$; then the function f which is the identity on $A \setminus \text{range}(\phi)$ and takes $\phi(n)$ to $\phi(n+1)$ on range(ϕ) is a bijection between A and $A \setminus \{\phi(0)\}$, so A is Dedekind-infinite.

(2) $\implies$ (1) If A is not finite we can recursively define a function $f : \omega \to A$ by $f(0) \in A, f(n+1) \in A \setminus \{f(0), f(1), \ldots, f(n)\}$. The construction can be carried out since by the assumption $A \setminus \{f(0), f(1), \ldots, f(n)\}$ is always nonempty.

(1) $\implies$ (3) Suppose there exists a finite set A which is not Dedekind-finite. Since finite sets are in 1-1 correspondence with natural numbers, we may assume that A is natural number. Let $m \in \omega$ be *least* such that there is a proper subset $B \subsetneq m$ and bijection $\phi : m \to B$. Since $\emptyset$ has no proper subsets, $m > 0$. By composing (if necessary) ϕ with a permutation that switches $m - 1$ and some element $b \in B$, we may assume that $\phi(m-1) = m-1$. But then the restriction of ϕ to $m - 1$ is a bijection between $m - 1$ and a proper subset of $m - 1$, comtradicting minimality.

For (4) $\implies$ (1), induct on $m \geq 1$ to show that

$$\text{For all } A, \text{ if there is an injection } g : A \to m \text{ then A is finite}$$

When $m = 1$, *domain*(g) must have only one element, so g witnesses $A \approx m$ therefore A is finite. For $m \implies m+1$, let g be an injection from A into $m+1$. If $m \in \text{range}(g)$ then in fact g maps A into m and m is finite by induction. Otherwise, let $a_0 \in A$ with $g(a_0) = m$, note this a_0 is the only such element, and then g is an injection from $A - \{a_0\}$ into m. By induction, $A - \{a_0\}$ is finite, say $f : n \to A - \{a_0\}$ where f is a bijection and $n \in \omega$. Then $f \cup \{\langle n, a_0\rangle\}$ is a bijection from $n + 1$ onto A, so A is finite.

Note, by the way, that for $A = \emptyset$ these are all true.

Cor 10.2 *If A is finite then there is a* unique *$m \in \omega$ with $A \approx m$*

Proof: Suppose $m \approx A \approx n$ for some $m, n \in \omega$. Suppose (for a contradiction) that $m \neq n$, say (by trichotomy) $m < n$. Of course, m is a proper subset of n, so this means that n is Dedekind-infinite, a contradiction.

Remark: For A finite we can now define card(A) to be the unique natural number n such that $A \approx n$. Similarly, if A is any countable set then we write card(A) $= \aleph_0$, and we write card(A) $\leq \aleph_0$ to mean that card(A) $\leq$ card(ω).

10.2 Countability

Definition 10.4 *A is* countable *(or* enumerable*) if either A is finite or $A \approx \omega$.*

Theorem 10.6 *Suppose $A \neq \emptyset$. The following are equivalent:*

1. *A is countable*
2. $\mathrm{card}(A) \leq \aleph_0$
3. *There is a surjection f from ω onto A*

Proof: (1) $\implies$ (2): If A is finite then $A \approx n \subseteq \omega$ for some $n \in \omega$, and if A is countable and infinite then $A \approx \omega$ by definition; either way, $\mathrm{card}(A) \leq \mathrm{card}(\omega) = \aleph_0$

(2) $\implies$ (3): This follows immediately from Theorem 10.2.

(3) $\implies$ (1): If A is finite then (1) holds by definition. If A is infinite and there is a surjection f from ω onto A, then by Theorem 10.2 and the definition of "A is infinite" we have $\mathrm{card}(\omega) \leq \mathrm{card}(A) \leq \mathrm{card}(\omega)$, which suffices.

Remark: We will call a set A *countably infinite* if it is countable and infinite. This is of course, just another way of saying that $A \approx \omega$.

Theorem 10.7 *If A and B are countable then (1) $A \times B$ and (2) $A \cup B$ are countable*

Proof: class and text.

Cor 10.3 *If A_1,...,A_n are countable then $A_1 \times A_2 \times \cdots \times A_n$ and $A_0 \cup \cdots \cup A_n$ are countable*

Proof: Class (induction). The union result is also a consequence of the next theorem.

Theorem 10.8 *A countable union of countable sets is countable.*

Proof: Let $\mathcal{A}$ be such a countable collection, WOLG $A \neq \emptyset$ for every $A \in \mathcal{A}$. There is a function f from ω onto $\mathcal{A}$ with $f(n)$ countable for all n, so for each $n \in \omega$ there is a function g_n from ω onto $f(n)$. Define $\phi : \omega^2 \to \cup\mathcal{A}$ by $\phi(m, n) = g_n(m)$; it is easy to verify that ϕ is onto.

Remark: Is ϕ actually a function? Let $G = \{(a, g) \in \mathcal{A} \times {}^{\omega}(\cup\mathcal{A}) : \mathit{range}(g) = a\}$ This is a set (by subset selection). Then by AC there is a function $H \subseteq G$ with $\mathit{domain}(G) = \mathit{domain}(H)$. Our function ϕ can now be defined by $\phi(m, n) = H(f(n))(m)$

An alternate approach to countability:

Definition 10.5 *If Σ is any set (which we think of as a set of "letters", let Σ^* be the set of finite "words" on Σ; formally,* $\Sigma^* := \bigcup_{n<\omega} {}^{n}\Sigma$

Theorem 10.9 *If Σ is countable then so is Σ^**

Proof 1: From earlier in the semester, ${}^n\Sigma \approx \Sigma \times \cdots \times \Sigma$ (n times) and the latter is countable, so Σ^* is the countable union of countable sets, hence countable

Proof 2: Without loss of generality $\Sigma \subseteq \omega - \{0\}$. Define $\phi : \Sigma^* \to \omega$ by $phi(\tau) = p_0^{\tau(0)} p_1^{\tau(1)} \cdots p_{n-1}^{\tau(n-1)}$, where $n = domain(\tau)$ and $p_0, p_1, \ldots$ enumerates the prime numbers. It is easy to see that ϕ is one-to-one, and that suffices.

Remark: The nice thing about the latter proofs is that it doesn't require the earlier results about cartesian products etc. In fact, these earlier results can be proved as applications of the theorem:

Cor 10.4 $\mathbb{Q}$ *is countable*

Proof: Every rational can be written as a word on the finite alphabet $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, /, -\}$; eg, $\frac{-37}{4}$ is the 5 character *word* "-37/4".

Cor 10.5 *if A and B are countable then $A \times B$ is countable*

Proof: Every pair in $A \times B$ can be writen as a word on the alphabet $A \cup B \cup \{(,),,\}$ (note the comma in the last set).

11 Ordinals

Before we can proceed with cardinality, it makes sense to finally *define* some cardinals other than the natural numbers and $\aleph_0$. the first step is to define the *ordinal numbers.*

Just as cardinality is meant to be a measure of a set's size, ordinality is meant to be a measure of the order structure of a set. When discussing a queue of people, for example, we might say that a person is 5th in line. This has cardinality implications - the set of people from the front to/including him has 5 people in it - but also an ordinality implication: 5 indicates his place in the line's order.

Intuitively, an ordinal is a well-ordered set which represents all wosets with a given 'order type' (that is, it represents all wosets which are order-isomorphic to it). Since ordinals are meant to generalize the natural numbers, the order relation on ordinals will just be set membership $\in$.

The collection of all ordinals will be too large to be a set; while occasionally we might write "$\alpha \in \mathbf{ON}$", this is really just shorthand for "α is an ordinal" and shouldn't be taken as implying the existence of a set **ON**. However, we will show that the ordinals 'line up' just like a linearly-ordered set, and that any nonempty set of ordinals has a least one, just like a woset. In general, it really won't hurt to much to act as if the collection of ordinals form a set **ON**.

Definition 11.1 *An* ordinal number *is a woset* $(\alpha, <)$ *satisfying:*

$$\forall a \in \alpha \ a = \{x \in \alpha \ : \ x < a\}$$

Note that this means that for any $x, a \in \alpha$, $x \in a \iff x < a$, that is, the well-order on α is really just set membership.

Lemma 11.1 *The following are equivalent:*

1. *α is an ordinal*
2. *α is a transitive set, and $\in$ is a linear order on α*

Proof: class.

Lemma 11.2 *If α is an ordinal, then:*

1. *Every $a \in \alpha$ is an ordinal*
2. *$\alpha^+ = \alpha \cup \{\alpha\}$ is an ordinal*

Proof: class.

Note that every $a \in \alpha$ is an 'initial segment' (or cut) of α; it is not obvious, however, that every initial segment is an element of α. We will prove this as Corollary 11.3, but you should keep this in mind for the next couple of lemmas.

Lemma 11.3 *If α and β are ordinals, then $\alpha \cap \beta$ is an ordinal.*

Proof: class.

Lemma 11.4 *(Trichotomy) If α and β are ordinals then either $\alpha \in \beta$ or $\alpha = \beta$ or $\beta \in \alpha$*

(Note that this means that every set of ordinals is well-ordered by $\in$.) **Proof:** I'll prove something stronger, namely

$$\forall x \in \alpha^+ \forall y \in \beta^+ (x \in y) \wedge (x = y) \wedge (y \in x)]$$

Suppose not, that is, $\exists x \in \alpha^+ \exists y \in \beta^+ [x$ and y are incomparable]. Let a be the least element of α^+ such that a is incomparable to *some* element of β^+; note that every element of a is smaller, so is comparable to every element of β^+. Now, let b be the least element of β^+ such that a and b are not comparable; note that every element of b is comparable to a.

Let x be any element of a; x is comparable to b. If $b \in x$ then $b \in x \in a$, so $b \in a$ (by transitivity of a), but this can't happen since a and b are incomparable. Similarly, $b \neq x$. It follows that $x \in b$. Since x was arbitrary in a, $a \subseteq b$. A similar argument shows that $b \subseteq a$, which proves that $a = b$. However, this contradicts the assumption that a and b are incomparable, proving the Lemma.

Cor 11.1 *If α and β are ordinals, and $\alpha \subseteq \beta$, then $\alpha = \beta$ or $\alpha \in \beta$*

Proof: If neither $\alpha = \beta$ nor $\alpha \in \beta$, then by trichotomy, $\beta \in \alpha$. But then $\beta \in \beta$, which violates the foundation axiom.

Cor 11.2 *If A is a nonempty set of ordinals then $\cup A$ is an ordinal*

Proof: We'll use the criterion from Lemma 11.1. First, note that $\cup A$ is itself a set of ordinals (why?), so trichotomy applies to $\cup A$, proving $\cup A$ is linearly ordered by $\in$. For transitivity, let $x \in a \in \cup A$, then for some ordinal $b \in A$, $x \in a \in b$. Then $x \in b$, so $x \in \cup A$, as desired.

Cor 11.3 *Suppose A is a nonempty set of ordinals such that whenever $x < \alpha \in A$, $x \in A$. Then A is an ordinal.*

Proof: Since $<$ is just $\in$, A is a transitive set. By the comment above, A is well-ordered by $\in$.

11.1 Well-orderings and ordinals

Suppose $(X, <)$ is a woset. In this section we show how to find an order-isomorphisn $E : X \to \alpha$ for some ordinal α.

FIRST APPROXIMATION

Put

$$E(x) := \{E(y) \ : \ y < x\}$$

This is a recursive definition; note, for example, that if x_0 is the smallest element of X, then $E(x_0) = \emptyset$.

Suppose for the time being that this definition of E is OK.

Lemma 11.5 *1. E is order-preserving from $(X, <)$ onto* $(\mathrm{range}(E), \in)$

2. $\mathrm{range}(E) \in \mathbf{ON}$ *(that is,* $\mathrm{range}(E)$ *is an ordinal)*

Proof. (1) If $a < b$ (in X) then $E(a) \in \{E(y) \ : \ y < b\} = E(b)$, so E is order preserving. (2) Put $\gamma = \mathrm{range}(E)$. To show: $\gamma \in \mathbf{ON}$, we need to show that (a) γ is linearly ordered by $\in$, and (b) γ is transitive. (a) is automatic, since γ is the image of an order-preserving function from a linearly ordered set, namely, $(X, <)$. For (b), let $a \in b \in \gamma$, then there is an $x \in X$ such that $b = E(x)$; since $E(x) = \{E(y) \ : \ y < x\}$, there is a $y < x$ such that $a = E(y)$, so $a \in \mathrm{range}(E) = \gamma$, done.

Cor 11.4 *For any woset $(X, <)$ there exists a unique ordinal γ such that $(X, <)$ is order-isomorphic to γ.*

Proof: The last lemma shows that *some* ordinal γ exists which is isomorphic to $(X, <)$. For uniqueness, suppose (for a contradiction) that there was second ordinal α isomorphic to $(X, <)$. By trichotomy we may assume WOLG that $\alpha \in \gamma$. Then $(\gamma, \in) \cong (X, <) \cong (\alpha, \in)$. Then there is an order isomorphism $\phi : \gamma \to \alpha$. Consider $\phi(\alpha), \phi(\phi(\alpha)), \phi(\phi(\phi(\alpha))), \cdots$. Note $\phi(\alpha) \in \mathrm{range}(\phi) \subseteq \alpha$, so $\phi(\alpha) \in \alpha$. Similarly, $\phi(\phi(\alpha)) \in \phi(\alpha)$, etc.

More precisely, we can recursively define $h : \omega \to \gamma$ by $h(0) := \alpha, \quad h(n^+) := \phi(h(n))$, and show by induction that $h(n^+) \in h(n)$ for all $n \in \omega$. This is a decreasing sequence in γ, contradicting the fact that $\in$ is a well-ordering on γ, done.

Cor 11.5 *No two distinct ordinals are order-isomorphic.*

Definition 11.2 *1. If α is an ordinal,* $\mathrm{card}(\alpha)$:= *the least element of* $\{y \leq \alpha : y \approx \alpha\}$

2. An ordinal κ is a cardinal number *provided* $\mathrm{card}(\kappa) = \kappa$

3. If $(X, <)$ is a woset, $ot(X)$:= the "order type" of X:= the unique ordinal γ such that $(X, <) \cong (\gamma, \in)$, and $\mathrm{card}(X) := \mathrm{card}(ot(X))$

4. More generally, if X is any *set, κ is a cardinal, and $X \approx \kappa$, put* $\mathrm{card}(X) := \kappa$

Remarks:

1. Note that if α is an ordinal, then $\mathrm{card}(\alpha)$ must really exist, since the set $\{y \leq \alpha : y \approx \alpha\}$ is nonempty (it contains α!) and α is well ordered. Of course, the least element of $\{y \leq \alpha : y \approx \alpha\}$ is also the least element of $\{y \in \mathbf{ON} : y \approx \alpha\}$ (why?).

2. This means that cardinals actually exist. You should verify that every $n \in \omega$ is a cardinal, and that ω itself is a cardinal. However, ω^+ is *not* a cardinal, since $\omega^+ \approx \omega$; $\text{card}(\omega^+) = \omega$. (We usually write $\aleph_0$ for ω - we'll see why later.) We do *not* yet know that there are any cardinals *bigger* than $\aleph_0$

3. This definition of $\text{card}(X)$ (for X well ordered) does *not* depend on the wellorder on X; it depends on X alone (see class notes).

4. For the last part of the definition, we need do show that it is well-defined It suffices to observe that if $\kappa \approx \lambda$ are both cardinals, then $\kappa = \lambda$ (why is this true?)

5. We have finally defined cardinal, and assigned cardinals to all well-ordered sets, including ordinals, and possibly many other sets. Exercise: Show that for any X and Y if $\text{card}(X)$ and $\text{card}(Y)$ are defined, then $\text{card}(X) \leq \text{card}(Y) \iff \text{card}(X)$ (as an ordinal) is less than or equal to $\text{card}(Y)$. At least convince yourself that you *understand* this exercise!

Salvaging the argument.

Everything we've just done works provided we know that E is a well-defined function. There are two ways to demonstrate that in fact it is:

(A) Use a different version of the recursion theorem (i.e., 8.5) to get E. This is how many texts do it. The problem is that *we* never proved this!

(B) Find a suitable set Z to serve as our range, and allow us to use the version of the recursion theorem we did (or will) prove. This is how we will proceed.

The Z we need must be a 'big' set; we digress for a bit and go hunting for big sets. Begin with the:

Theorem 11.1 *(Burali-Forti)* **ON** *is not a set.*

Proof: (two proofs in class)

Most of the rest of this section will be devoted to finding an ordinal α into which the woset $(X, <)$ embeds.

Proposition 11.1 *Let $x_0 \in X$, suppose $\forall x < x_0 \ \exists y \in$ **ON** $\exists \phi : \text{seg}(x) \to y$ which is an order-isomorphism. Then $\exists y \in$ **ON** $\exists \phi : \text{seg}(x_0) \cup \{x_0\} \to y$ which is an order-isomorphism.*

Here $\text{seg}(x) = \{y \in X : y < x\}$, analogous to an open interval *excluding* the right endpoint, whereas $\text{seg}(x_0) \cup \{x_0\}$ *includes* the right endpoint. This proposition is the key induction step in a recursive construction of the embedding we are aiming for.

Proof: The general idea is to put $\Theta(x) :=$ the least $y \in \mathbf{ON}$ such that $\text{seg}(x)$ embeds isomorphically into y; the hypothesis says that such a y exists for every $x \in \text{seg}(x_0)$. Then $\text{range}(\Theta)$ is a set of ordinals, and we can put $\alpha := \bigcup\{y^+ : y \in \text{range}(\Theta)\}$. One easily verifies Θ is an order-isomorphism from $\text{seg}(x_0)$ into α, and then $\Theta \cup \{(x_0, \alpha)\}$ is then an order-isomorphism from $\text{seg}(x_0) \cup \{x_0\}$ into α^+.

The only catch is that Θ as defined is not *a priori* an actual function. So - an annoying technicality - one needs to treat Θ as a "function-like" *formula*, and invoke the Axiom of Replacement.

Cor 11.6 *For some ordinal α there is an order isomorphism $\phi : X \to \alpha$*

Proof: First, extend $(X, <)$ by adding a right endpoint: let e be any element not already in X, put $X' := X \cup \{e\}$, $\quad <' := < \cup \{(x, e) : x \in X\}$, and consider two cases:

Case 1: $\forall x \in X\ \exists y \in \mathbf{ON}\ \ \exists \phi : \text{seg}(x) \to y$: Then by the last proposition (with $x_0 = e$) there is an ordinal α and an order isomorphism $\phi : X \cup \{e\} \to \alpha$, and the restriction of this ϕ to $X = \text{seg}(e)$ works.

Case 2: Otherwise: Then there is a least $x_0 \in X$ witnessing failure, that is, such that $\forall y \in \mathbf{ON}\ \neg\exists\phi : \text{seg}(x_0) \to y$ The hypotheses of the last proposition now hold for this x_0, and so $\exists y \in \mathbf{ON}\ \exists \phi : \text{seg}(x_0) \cup \{x_0\} \to y$ a contradiction (so this case cannot happen).

Note that in some sense we could now skip our quest to properly define the function E, as we have an order-isomorphism from $(X, <)$ into an ordinal. However, we don't yet have that this is onto (though we could have built this into the argument above), and it is nice to have the explicit function E. So, we press on.

THE CONSTRUCTION...SALVAGED!

We can now redefine our E as follows. Fix an order-preserving ϕ from X to an ordinal α. Define:

$$E(x) := \begin{cases} \{E(y) \ : \ y < x\} & \text{if } \{E(y) \ : \ y < x\} \subseteq \phi(x), \\ \alpha & \text{otherwise.} \end{cases}$$

More formally, in Theorem 8.4 put

$$Z = \mathcal{P}\,(\alpha), \quad G(\tau, x) = \begin{cases} \text{range}(\tau) & \text{if range}(\tau) \subseteq \phi(x), \\ \alpha & \text{otherwise.} \end{cases}$$

Lemma 11.6 $\forall x \in X \ \{E(y) \ : \ y < x\} \subseteq \phi(x)$

Proof: A straightforward induction (see class notes).

This means that in the (re)definition of E, the second case never happens, and our original, provisional definition really works:

Cor 11.7 $\forall x \in X \ E(x) = \{E(y) \ : \ y < x\}$

Cor 11.8 *All our earlier results about E hold*

Remark: E is really a function on *orders*, not sets; probably we really should have written $E :< \to \mathbf{ON}$ instead of $E : X \to \mathbf{ON}$.

We have now shown that every well-ordered set has a cardinality. The following theorem now implies that *every* set has a cardinality:

Theorem 11.2 *(AC) Every set can be well-ordered.*

Note the annotation (AC); that means that this theorem relies on the Axiom of Choice. In fact, it is *equivalent* to the Axiom of Choice, so its proof will be deferred to the next section.

Cor 11.9 *Every set has a cardinality; that is, for every X there is a unique cardinal number κ such that $X \approx \kappa$.*

Proof: Existence follows from the previous theorem (given X, well order it, and let $\kappa = \text{card}(X)$). For uniqueness, if κ, λ are cardinals and $\kappa \approx X \approx \lambda$ then $\kappa \approx \lambda$ and by the previous observation $\kappa = \lambda$

12 Cardinal Arithmetic

We no extend some definitions of operation on cardinals from ω to more general cardinals.

Definition 12.1 *Let κ, λ be cardinals, and let K, L be disjoint sets with $\kappa = \text{card}(K), \lambda = \text{card}(L)$*

1. $\kappa + \lambda := \text{card}(K \cup L)$
2. $\kappa \cdot \lambda (\text{ or } \kappa\lambda) := \text{card}(K \times L)$
3. $\kappa^\lambda := \text{card}({}^L K)$

Remarks:

- Disjointness of K and L is only used above in for the definition of $+$. (Why is it necessary there?)
- In particular, we *could* have just defined $\kappa \cdot \lambda := \text{card}(\kappa \times \lambda), \kappa^\lambda := \text{card}({}^\lambda \kappa)$
- Similarly, for $+$ we could take $K = \{0\} \times \kappa, L = \{1\} \times \lambda$
- Note that for *finite* cardinals, the above definitions are consistent with the usual ones for natural numbers.
- WARNING: We still need to show that the above definition is sensible! In particular, we need to show that the operations don't depend on the choice of K and L (other than their disjointness). That follows from the following theorem:

Theorem 12.1 *If $K_1 \approx K_2$ and $L_1 \approx L_2$ then:*

(a) If $K_1 \cap K_2 = L_1 \cap L_2 = \emptyset$ then $K_1 \cup L_1 \approx K_2 \cup L_2$

(b) $K_1 \times L_1 \approx K_2 \times L_2$

(c) ${}^{L_1}K_1 \approx {}^{L_2}K_2$

Proof: Class.

Examples:

1. $n + \aleph_0 = \aleph_0 + n$ for $n \leq \aleph_0$.
2. For all cardinals κ, $\kappa^0 = 1$. If $\kappa \neq 0$ then $0^\kappa = 0$, and $0^0 = 1$
3. If $\text{card}(A) = \kappa,\ \text{card}(\mathcal{P}(A)) = \text{card}({}^A 2) = 2^\kappa$
4. For all cardinals κ, $\kappa < 2^\kappa$ (In particular, $\aleph_0 < 2^{\aleph_0}$)

The next result essentially repeats some properties of those operations which we proved already for natural numbers. Don't get too attached to these results, we will come up with *much* better ones soon!

Theorem 12.2 *Let κ, λ, μ be cardinals. Then:*

1) $+, \cdot$ obey the commutative, associative, and distributive properties.

2) $\kappa^{\lambda+\mu} = \kappa^\lambda \kappa^\mu$

3) $(\kappa\lambda)^\mu = \kappa^\mu \lambda^\mu$

4) $(\kappa^\lambda)^\mu = \kappa^{(\lambda\mu)}$

To prove any of these, one observes that there is a bijection between two appropriate sets. For example, for the distributive property we let K, L, M be disjoint sets and find a bijection between $K \times (L \cup M)$ and $(K \times L) \cup (K \times M)$ (in this case just the identity!) DO some others as an exercise.

Cor 12.1 *If A and B are finite, then so are $A \cup B, A \times B$, and ${}^B A$*

Like the last theorem, the next theorem is a generalization of known results for elements of ω, and is proved by considering embeddings of one set into another:

Theorem 12.3 *Let κ, λ, μ be cardinals, with $\kappa \leq \lambda$. Then:*

1) $\kappa + \mu \leq \lambda + \mu$

2) $\kappa\mu \leq \lambda\mu$

3) $\kappa^\mu \leq \lambda^\mu$

4) $\mu^\kappa \leq \mu^\lambda$ provided $\kappa \neq 0$ or $\mu \neq 0$

12.1 Absorption Laws

The next result is the key to everything that follows.

Theorem 12.4 *If κ is an infinite cardinal, then $\kappa \cdot \kappa = \kappa$*

Proof: Otherwise let κ be the least infinite cardinal such that $\kappa \cdot \kappa > \kappa$. In particular, if β is an *ordinal* less than κ, then $\mathrm{card}(\beta) < \kappa$ (why?), so either $\mathrm{card}(\beta) \cdot \mathrm{card}(\beta)$ is finite (i.e., when $\beta \in \omega$) or $\mathrm{card}(\beta) \cdot \mathrm{card}(\beta) = \mathrm{card}(\beta)$ by choice of κ. We note also that for any infinite β, $\beta \times \beta \approx \mathrm{card}(\beta) \times \mathrm{card}(\beta)$ (exercise), so if β is an ordinal less than κ then $\mathrm{card}(\beta \times \beta) < \kappa$

Put $A = \kappa \times \kappa$, and for ordinals $\beta < \kappa$ put $A_\beta := \beta \times \beta$. note $A = \bigcup_{\beta \in \kappa} A_\beta$ (why?)

Let $\prec_0$ be a well-order of A (e.g., the lexicographic order - we don't need AC here!), and define a different order $\prec$ of A by $x \prec y$ if either $x \in A_\beta, y \notin A_\beta$, or $\forall \beta < \kappa (x \in A_\beta \iff y \in A_\beta)$ and $x \prec_0 y$. In other words, $\prec := ((\bigcup_{\beta<\alpha<\kappa} A_\beta \times (A_\alpha - A_\beta)) \cup (\bigcup_{\alpha<\kappa} \prec\restriction_{(A_\alpha - \bigcup_{\beta<\alpha} A_\beta)}))$ Exercise: $\prec$ well-orders A, and if $\beta < \alpha$ then every element of A_β is less than every element of $A_\alpha \setminus A_\beta$.

Now, let $\alpha = E(A, \prec)$ be the order type of $(A, \prec)$. If $\alpha \leq \kappa$ then E maps A injectively into κ, so $\mathrm{card}(A) \leq \mathrm{card}(\kappa)$, contradiction. Alternately, by trichotomy $\kappa < \alpha$, so $\kappa = E(a)$ for some $a \in A$. Then for some $\beta < \kappa$, $a \in A_\beta$. Then $\kappa = \mathrm{card}(\{x \in A : x \prec a\}) \leq \mathrm{card}(A_\beta) < \kappa$ (from above), a contradiction.

Cor 12.2 *Suppose κ is an infinite cardinal, and $\{A_\beta\}_{\beta<\kappa}$ is a sequence of sets such that $\mathrm{card}(A_\beta) \leq \kappa$ for every β. Then $\mathrm{card}(A) \leq \kappa$, where $A = \bigcup_{\beta \in \kappa} A_\beta$*

Proof: For $\beta < \kappa$ let ϕ_β be a function from κ onto A_β. Define $\phi : \kappa \times \kappa \to A$ by $\phi(\beta, \gamma) := \phi_\beta(\gamma)$. This is obviously onto, so $\mathrm{card}(A) \leq \mathrm{card}(\kappa \times \kappa) = \kappa \cdot \kappa = \kappa$.

Cor 12.3 *Let κ, λ be cardinals, $0 \neq \kappa \leq \lambda, \lambda \geq \aleph_0$. Then $\kappa + \lambda = \kappa \cdot \lambda = \lambda$*

Proof: $\lambda \leq \kappa + \lambda \leq \lambda + \lambda = \lambda \cdot 2 \leq \lambda \cdot \lambda = \lambda$ and $\lambda \leq \kappa \cdot \lambda \leq \lambda \cdot \lambda = \lambda$

Remark: What properties of these operations are we using? Why are they true?

Cor 12.4 *$\kappa^\kappa = 2^\kappa$ for κ an infinite cardinal.*

Proof: (class)

Lemma 12.1 *Let A be a set of cardinals. Then $\bigcup A$ is a cardinal.*

Proof: (class)

Definition 12.2 *If κ is a cardinal, write κ^+ for the least cardinal larger than κ. (κ^+ is the* successor *of κ.)*

Remarks:

- WARNING: κ^+ with κ a cardinal is *different* from κ^+ when κ is viewed as an ordinal! In particular, $\text{card}(\kappa^+) > \text{card}(\kappa)$ as cardinals, but $\text{card}(\kappa^+) = \text{card}(\kappa)$ when talking about successors of ordinals. Probably it would have been prudent to use different notation for the two notions of successor, but this usage is conventional.

- Existence of κ^+ follows from the existence of *some* cardinal bigger than κ, for example $\text{card}(\mathcal{P}(\kappa))$.

Definition 12.3

- *An ordinal α is a* successor ordinal *provided $\alpha = \beta^+$ for some ordinal β (which is then called the* predecessor *of α. (Note: this is the* ordinal *successor function.)*

  An ordinal which is not a successor ordinal is called a limit *ordinal.*

- *A cardinal κ is a* successor cardinal *provided $\kappa = \mu^+$ for some cardinal μ (which is then called the* predecessor *of κ. (Note: this is the* cardinal *successor function.)*

  A cardinal which is not a successor cardinal is called a limit *cardinal.*

Lemma 12.2 *Every infinite cardinal is a limit ordinal.*

Proposition 12.1 *If $\kappa \geq \aleph_0$ then $\kappa^+ = \text{card}(\kappa^+ - \kappa)$*

Definition 12.4 *Define $\aleph : \mathbf{ON} \to \mathbf{ON}$ recursively by*

$$\aleph_0 := \omega$$

$$\aleph_{\beta^+} := (\aleph_\beta)^+$$

$$\aleph_\lambda := \bigcup_{\beta\in\lambda} \aleph_\beta, \qquad \textit{if } \lambda \textit{ is a limit ordinal}$$

Remarks:

1. The "+" in "β^+" is the ordinal successor function, while the "+" in "$\aleph_\beta^+$" is *cardinal* successor.

2. This is (of course!) not an actual function; but you should convince yourself that it is suitably fuction-like for our recursive definition theorems to apply. In particular, for any ordinal α there is one and only one cardinal β such that there is a bijection from (the set) alpha to β, that the range of this bijection hits every cardinal less than β, etc.

3. Exercise: Every cardinal is $\aleph_\beta$ for some β

4. Note that this construction proves that the class of all cardinals has the same order structure as the class of all ordinals. In particular, we can induct on cardinals (which is sometimes useful).

Definition 12.5 *Define* $\beth : \mathbf{ON} \to \mathbf{ON}$ *recursively by*

$$\begin{aligned} \beth_0 &:= \aleph_0 \\ \beth_{\beta+} &:= 2^{\beth_\beta} \\ \beth_\lambda &:= \bigcup_{\beta \in \lambda} \beth_\beta, \qquad \textit{if } \lambda \textit{ is a limit ordinal} \end{aligned}$$

Remarks:

1. The comments above (about the meaning of "+", and about "functionness" of $\aleph$) apply here as well.

2. Exercise: Prove that $\aleph_\alpha \leq \beth_\alpha$ for all α

3. Question: For every cardinal κ, is $2^\kappa = \beth_\beta$ for some ordinal β?

4. Cantor's *Continuum Hypthesis* (CH) can now be rephrased as: $\aleph_1 = \beth_1$

5. Cantor's *Generalized Continuum Hypthesis* (GCH) can now be rephrased as: $\forall \beta,\ \aleph_\beta = \beth_\beta$

13 Axiom of Choice

As we have seen, notions of cardinality are intimately connected to the Axiom of Choice. Here are several alternative formulations of this axiom (by no means an exhaustive list!):

(4.) If $\mathcal{A}$ is a set of disjoint nonempty sets then there is a set B such that for every $a \in \mathcal{A}$ $B \cap a$ is a singleton.

(1.) If R is a relation then there is a function $F \subseteq R$ with $\text{dom}(R) = \text{dom}(F)$

(2.) If H is a function with domain I and $\forall i \in I$ $H(i) \neq \emptyset$ then $\prod_i H(i) \neq \emptyset$

(3.) For every A there is a "choice function" $F : \mathcal{P}\ (A) - \{\emptyset\} \to A$ such that $F(a) \in a$ for every a.

(5.) (Cardinal Comparability) For every C, D there is either an injection from C into D or an injectionf from D into C. (In other words, either $\text{card}(C) \leq \text{card}(D)$ or $\text{card}(D) \leq \text{card}(C)$.)

(6.) (Zorn's Lemma) Let $\mathcal{A}$ be a set, suppose $\mathcal{A}$ is closed under unions of chains. Then $\mathcal{A}$ contains a maximal element.

(7.) (All sets have cardinalities) $\forall X$ $\exists!$ cardinal κ such that $X \approx \kappa$

(8.) Every set can be well-ordered.

(For Zorn's Lemma, note that $\mathcal{A}$ is partially ordered (but possibly not linearly ordered) by $\subseteq$; a *chain* is a subset $\mathcal{B}$ of $\mathcal{A}$ which *is* linearly ordered. The hypothesis is that for every such $\mathcal{B}$, $\cup\mathcal{B}$ is again an element of $\mathcal{A}$. *Maximal* in the conclusion means with respect to this same ordering $\subseteq$.)

13.1 Some proofs

$(1 \to 2 \to 4)$ These are all quite easy; class and/or handouts and/or exercises.

$(4 \to 3)$ The main idea is to make the nonempty subsets of A disjoint so that we can apply AC(4). So, put $\mathcal{A} := \{a \times \{a\} : \emptyset \neq a \subseteq A\}$. Observe that $a \times \{a\} \neq b \times \{b\} \iff a \times \{a\} \cap b \times \{b\} = \emptyset$. Let B consist of one element from each $a \times \{a\}$. Note that $B \cap (a \times \{a\})$ has the form (x, a) for some $x \in a$. Now we can define $F : \mathcal{P}\ (A) - \{\emptyset\} \to A$ by $F(a) :=$ the first component of $B \cap (a \times \{a\})$. (Even better: just take $F = B^{-1}$.) Then $F(a) \in a$ for all $a \neq \emptyset$, as desired.

$(3 \to 8)$ Let X be a set. By Hartog's Theorem, there is some $\alpha \in \mathbf{ON}$ such that $\text{card}(\alpha) \not\leq \text{card}(X)$. Let x_∞ some element not in X, put $X' := X \cup \{x_\infty\}$, let $f : \mathcal{P}(X) - \{\emptyset\} \to X$ be a choice function, extend f to all of $\mathcal{P}(X)$ by setting $f(\emptyset) := x_\infty$. By the recursion theorem we can define a function ϕ on α satisfying $\phi(\beta) := f(X - \text{range}(\phi \upharpoonright_\beta)),\ \beta \in \alpha$. (Convince yourself that this makes sense!) Now we confirm:

(a) ϕ is one-to-one on $\phi^{-1}(X)$: Indeed, if $a < b \in \alpha$ and $\phi(a), \phi(b) \in X$ then $\phi(b) \in X - \{\phi(x) : x < b\} \subseteq \{\phi(a)\}$, so $\phi(b) \neq \phi(a)$.

(b) ϕ is onto X: Otherwise $X - \text{range}(\phi) \neq \emptyset$, so for every $\beta \in \alpha,\ f(X - \text{range}(\phi \upharpoonright_\beta)) \in X$, so by the definition of ϕ, $\phi(\beta) \in X$. then ϕ is an injection from α into X, contradicting the choice of α.

So, ϕ is a one-to-one function from some subset of α onto X, and X can now be well-ordered by reference to α, that is, define $x \prec y$ on X provided $\phi^{-1}(x) \in \phi^{-1}(y)$.

$(8 \to 7)$ This is now Corollary 11.9.

$(7 \to 8)$ Obvious (but show it!)

$(7 \to 5)$ Let $\alpha = \text{card}(C), \beta = \text{card}(D)$; trichotomy means that either $\alpha < \beta, \alpha = \beta, or \alpha > \beta$. Say $\alpha \leq \beta$; then if $\phi : C \to \alpha, \psi : \beta \to D$ are bijections then (noting $\alpha \subseteq \beta$) $\psi \circ \phi$ is an injection from C into D.

$(7 \to 6)$ Let $\alpha = \text{card}(\mathcal{A})$, then we may write $\mathcal{A} = \{A_\beta\}_{i\beta \in \alpha}$. Define chains $\mathcal{B}_\beta$ recusively by: $\mathcal{B}_0 := \{A_0\}$; given $\mathcal{B}_\gamma,\ \gamma < \beta$ put $\mathcal{B}_\beta^- := \bigcup_{\gamma < \beta} \mathcal{B}_\gamma$, note this is a chain. Now put $\mathcal{B}_\beta := \mathcal{B}_\beta^- \cup \{A_\beta\}$ if this union is a chain, otherwise just put $\mathcal{B}_\beta := \mathcal{B}_\beta^-$. Once all the chains $\mathcal{B}_\beta$ are defined, put $\mathcal{B} := \bigcup_{\beta < \alpha} \mathcal{B}_\beta$ Again, note that this is a chain. Now, put $M := \cup \mathcal{B}$, which by hypothesis is in $\mathcal{A}$. It remains to show that M is a maximal element of $\mathcal{A}$. Otherwise, there is a larger element in $\mathcal{A}$, say A_β. But $M \subseteq A_\beta$, so we must have thrown A_β into $\mathcal{B}_\beta$, so $A_\beta \subseteq M$, a contradiction.

$(6 \to 1)$ Let R be a relation, and assume Zorn's Lemma holds. Let $\mathcal{A} = \{f \subseteq R : f \text{ is a function}\}$. (In other words, $\mathcal{A}$ consists of all "partial function" approximations to the F we seek.) If $\mathcal{B} \subseteq \mathcal{A}$ is a chain in $\mathcal{A}$ then we already know - from an earlier exercise - that $\bigcup \mathcal{B}$ is a function, and obviously $\bigcup \mathcal{B} \subseteq R$, so $\mathcal{A}$ is closed under unions of chains. By Zorn we conlude $\mathcal{A}$ has a maximal element; denote it by F. It remains to show that F has the same domain as R. Let $x \in \text{dom}(R)$, then for some y, xRy. If F is not defined on x then $F \cup \{(x, y)\}$ would be a function extending F, a contradiction, so $x \in \text{dom}(F)$ as desired.

($8 \to 1$) We can also prove (1) from (8), quite easily in fact. Let $\prec$ well-order the range of R, and fpr $x \in \text{dom}(R)$ put $F(x) :=$ the least element (with respect to the order $\prec$) of $R[\![x]\!]$. In other words, $F := \{(x, y) \in R : \forall (a, b) \in R\ x = a \implies y \preceq b\}$. This F obviously works. (Why does it avoid choice?)

($5 \to 8$) Let X be a set. By Hartog's Theorem there is an ordinal α which does not embed into X. By (5), card(X) exists, and since α does not embed into X, it must not embed into card(X). By trichotomy we can then conclude that $\text{card}(X) \leq \alpha$, so there is an injection $\phi : X \to \alpha$. We can now use ϕ as usual to order X.

Cor 13.1 *If A is any set of cardinals then there is a cardinal κ such that $\forall \alpha \in A,\ \alpha < \kappa$.*

Proof: Let γ be any ordinal biggerthan every ordinal in A. Put $\kappa = \text{card}(\mathcal{P}(\kappa))$. The κ is a cardinal, we already know that it cannot be injected into γ, so it is bigger than γ hence every $\alpha \in A$.

Remark: AC is used here (when we assume that $\mathcal{P}(\gamma)$ has a cardinality). It is not strictly necessary! In fact, if $(X, <)$ is any woset then $\mathcal{P}(X)$ can be well-ordered without use of Choice. In particular, if we define $\prec$ on $\mathcal{P}(X)$ by $A \prec B$ provided the least element of $A \Delta B$ is in A, then this is a well-order (though the verification of this is not easy).

13.2 Other Equivalents

Here are some other interesting equivalents to AC:

(9.) (Zorn B) Assume $(A, \prec)$ is a poset, and assume that whenever B is a linearly ordered subset of A then B has an upper bound in A (that is, $\forall B \subseteq A$, $(B, \prec)$ a loset $\implies \exists c \in A\ \forall b \in B\ b \preceq c$).

(10.) (Hausdorff Maximality Principle) Assume $(A, \prec)$ is a poset. Every chain in A can be extended to a maximal chain. (That is, $\forall B \subseteq A$, if B is a chain with respect to $\prec$ then there is some chain B' such that $B \subseteq B' \subseteq A$ and such that if B'' is another chain with $B' \subseteq B'' \subseteq A$ then in fact $B'' = B'$.)

(11.) A set E is finite if and only if E is order-finite. (Def: E is *order-finite* provided for every well-order R of E, R^{-1} is also a well-order of E.)

13.3 Weak Forms

There are many weakenings of AC. Here is an interesting one. For $n \in \omega$ let C_n be the axiom:

If $\mathcal{A}$ is a family of n-element sets then $\mathcal{A}$ has a choice function

Obviously, (AC) implies $\forall n \in \omega \; C_n$. However, in the absence of AC all kinds of things can happen. For example, it is known that it is consistent that $C_2 \wedge \neg C_3$!

Theorem 13.1 *(Tarski)* $C_2 \implies C_4$