10 Functions and Measurements

Definition 10.1: Ordered Pair.
For any a and b, we use the notation (a, b) to denote the *ordered pair* whose first coordinate is a and second coordinate is b. This is defined formally as follows.

(a, b)

$$(a, b) \coloneqq \{\{a\}, \{a, b\}\}$$

This is a *"good"* implementation of the idea behind ***ordered** pairs* in set theory because, using this definition, we have $a \neq b \Leftrightarrow (a, b) \neq (b, a)$. ⌟

Theorem 10.1: Characteristic Property of Ordered Pairs.
$\forall a \forall b \forall x \forall y \big((a = x \ \wedge \ b = y) \ \Leftrightarrow \ (a, b) = (x, y) \big)$. ⌟

Lemma 10.1: Extraction of Coordinates.
Let x and y be arbitrary sets. The first coordinate x of the ordered pair (x, y) satisfies the formula $(\forall z \in (x, y))(\text{◌} \in z)$ and can therefore be extracted as follows.

$$x = \cup\cap(x, y)$$

The second coordinate y satisfies the formula $(\exists z \in (x, y))(\text{◌} \in z)$, making the situation more nuanced—notice that $\{x, y\} = \{x\}$ when $x = y$. If we know $x \neq y$, then the second coordinate can be extracted by the following computation.

$$y = \cup\Big(\big(\cup(x, y)\big) \setminus \big(\cap(x, y)\big) \Big)$$

In the case when $x = y$, we can simply extract the first coordinate instead. ⌟

Definition 10.2: Cartesian Product.
For any sets A and B, the *Cartesian product of A with B*, denoted $A \times B$, is defined to be the set of all ordered pairs (a, b) whose first coordinate a is an element of A and whose second coordinate b is an element of B.

$A \times B$

$$\begin{aligned} A \times B &\coloneqq \Big\{ z \ \Big| \ \exists a \exists b \big(z = (a, b) \wedge a \in A \wedge b \in B \big) \Big\} \\ &= \Big\{ (a, b) \ \Big| \ a \in A \wedge b \in B \Big\} \end{aligned}$$

⌟

Theorem 10.2: Existence of Cartesian Products.
$\forall A \forall B \exists P (P = A \times B)$. ⌟

10.1 Functions

Definition 10.3: Unique Existential Quantification.
For any well-formed formula φ, we define what it means to say that *exactly one* object in the universe of discourse satisfies the formula φ.

$$\exists! x \big(\varphi(x) \big) \ \ :\Leftrightarrow \ \ \exists x \Big(\varphi(x) \wedge \forall y \big(\varphi(y) \Rightarrow y = x \big) \Big)$$

⌟

Definition 10.4: Function.
For any X and Y, we define what it means for f to be a *function from X to Y* below.

$f : X \to Y$

$$f : X \to Y \quad :\Leftrightarrow \quad f \subseteq X \times Y \ \wedge \ (\forall x \in X)(\exists! y \in Y) \big((x, y) \in f \big)$$

The notation $f : X \to Y$ is used to indicate that *"f is a function from X to Y."* The set X, where f takes its inputs from, is called the *domain of f*, which we sometimes denote by $\mathfrak{dom}(f)$. The set Y, containing all of the outputs of f (and possibly more), is called the *codomain of f*, which we sometimes denote by $\mathfrak{cod}(f)$. The domain and codomain of a function are two of the key components in the definition of a function; changing either the domain or the codomain changes *who the function is.*

$\mathfrak{dom}(f)$

$\mathfrak{cod}(f)$

An ordered pair $(x, y) \in f$ corresponds to an *"input-output pair,"* where x is the *input* and y is the *output.* Since the definition above guarantees that every input $x \in X$ has exactly one corresponding output $y \in Y$, we use the *functional notation* $f(x) = y$ to indicate that f maps x to y.

$f(x) = y$

$$f(x) = y \quad :\Leftrightarrow \quad (x, y) \in f$$

We refer to y as *the image of x* under f, and we say x is *a preimage of y* under f. The *range of f* is the set of all images under f—*i.e.,* the set of actual outputs of f.

$\mathfrak{ran}(f)$

$$\mathfrak{ran}(f) := \Big\{ y \;\Big|\; y \in Y \wedge (\exists x \in X)\big(f(x) = y\big) \Big\}$$

Notice $\mathfrak{ran}(f) \subseteq \mathfrak{cod}(f)$ always holds, but $\mathfrak{ran}(f) \stackrel{?}{=} \mathfrak{cod}(f)$ might go either way.

Equality between functions is determined based on whether or not they have all of the same properties outlined in the definitions above. Most importantly, they must have the same domain and codomain, and they must map in the same way as each other. To be more precise: given two arbitrary functions f and g, we say that $f = g$ precisely when each of the three clauses below hold simultaneously.

$$\mathfrak{dom}(f) = \mathfrak{dom}(g) \;\wedge\; \mathfrak{cod}(f) = \mathfrak{cod}(g) \;\wedge\; \forall x \forall y \big(f(x) = y \Leftrightarrow g(x) = y\big) \qquad \lrcorner$$

Some Important Examples of Functions

Many of the examples of functions below have domains and codomains that can be restricted or extended to get different *"versions"* of these functions. For example, the absolute value function described below often has its domain and codomain restricted to $|\circlearrowleft| : \mathbb{Z} \times \mathbb{Z} \to \mathbb{N}$ for convenience without making any modification to the *mapping* itself. While these variants would all technically be different functions, those distinctions are often blurred in practice because the properties of the *functions themselves* are usually less important than the *input-to-output mapping.*[1] While we won't go through the trouble of describing every possible variant of each function below, we will make some comments about some of the notable ones.

1. › The *min* function is the map $\mathtt{min} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given below for $(x, y) \in \mathbb{R} \times \mathbb{R}$.

$$\mathtt{min}(x, y) := \begin{cases} x & \text{if } x \leqslant y \\ y & \text{if } x > y \end{cases}$$

› The *infimum* map $\mathtt{inf} : \{R \in \mathbb{P}(\mathbb{R}) \mid R \neq \varnothing \wedge (\exists \ell \in \mathbb{R})(\forall r \in R)(\ell \leqslant r)\} \to \mathbb{R}$ is defined, for every subset $R \subseteq \mathbb{R}$ that is *bounded from below,* as follows.

$$\mathtt{inf}(R) := \ell^*, \text{where } \ell^* \text{ is the greatest } \ell \in \mathbb{R} \text{ such that } (\forall r \in R)(\ell \leqslant r)$$

$\mathbb{R}$ in the domain and codomain can *both* be replaced by $\mathbb{Z}$ or $\mathbb{N}$, but not $\mathbb{Q}$.[2][3]

[1] For our purposes, if there is some doubt, it's best to just specify the domain, codomain, and mapping that we want to talk about whenever there is ambiguity.

[2] Every nonempty subset $\varnothing \neq R \subseteq \mathbb{R}$ that is *bounded from below* has a *greatest lower bound* in $\mathbb{R}$. The same applies to nonempty bounded-from-below subsets $Z \subseteq \mathbb{Z}$ and $N \subseteq \mathbb{N}$. However, there are nonempty subsets $\varnothing \neq Q \subseteq \mathbb{Q}$ who have *no greatest lower bound in the rationals.* For example, the subset $\{q \mid q \in \mathbb{Q} \wedge q^2 < 2\} \subseteq \mathbb{Q}$ has no greatest lower bound within $\mathbb{Q}$.

[3] The fact that `inf` is a function when $\mathbb{R}$ in its domain and codomain are both replaced by either $\mathbb{Z}$ or $\mathbb{N}$ can be easily proven by induction. This is an important property of $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{R}$ called the *"greatest lower bound"* property.

2. › The *max* function is the map $\mathtt{max} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given below for $(x, y) \in \mathbb{R} \times \mathbb{R}$.

$$\mathtt{max}(x, y) \coloneqq \begin{cases} x & \text{if } x \geqslant y \\ y & \text{if } x < y \end{cases}$$

› The *supremum* map $\mathtt{sup} : \{R \in \mathbb{P}(\mathbb{R}) \mid R \neq \varnothing \wedge (\exists u \in \mathbb{R})(\forall r \in R)(r \leqslant u)\} \to \mathbb{R}$ is defined, for every subset $R \subseteq \mathbb{R}$ that is bounded from *above*, as follows.

$$\mathtt{sup}(R) \coloneqq u^*, \text{where } u^* \text{ is the least } u \in \mathbb{R} \text{ such that } (\forall r \in R)(r \leqslant u)$$

$\mathbb{R}$ in the domain and codomain can *both* be replaced by $\mathbb{Z}$ or $\mathbb{N}$, but not $\mathbb{Q}$.[1][2]

3. The *floor* map $\lfloor \circlearrowleft \rfloor : \mathbb{R} \to \mathbb{Z}$ is given by $\lfloor r \rfloor \coloneqq \mathtt{sup}(\{\ell \mid \ell \in \mathbb{Z} \wedge \ell \leqslant r\})$ for $r \in \mathbb{R}$.
4. The *ceiling* map $\lceil \circlearrowleft \rceil : \mathbb{R} \to \mathbb{Z}$ is given by $\lceil r \rceil \coloneqq \mathtt{sup}(\{u \mid u \in \mathbb{Z} \wedge r \leqslant u\})$ for $r \in \mathbb{R}$.
5. The *absolute value* function $|\circlearrowleft| : \mathbb{R} \to \mathbb{R}$ is given, for $r \in \mathbb{R}$, by $|x| \coloneqq \begin{cases} x & \text{if } 0 \leqslant x \\ -x & \text{if } x < 0 \end{cases}$.

[1] Every nonempty subset $\varnothing \neq R \subseteq \mathbb{R}$ that is bounded from *above* has a *least upper bound* in $\mathbb{R}$. The same applies to every nonempty bounded-from-above subset $Z \subseteq \mathbb{Z}$ and $N \subseteq \mathbb{N}$. However, there do exist nonempty subsets $\varnothing \neq Q \subseteq \mathbb{Q}$ who have *no least upper bound in the rationals*. For example, the set $\{q \mid q \in \mathbb{Q} \wedge q^2 > 2\}$ has no least upper bound within $\mathbb{Q}$.

[2] The fact that sup is a function when $\mathbb{R}$ in its domain and codomain are both replaced by either $\mathbb{Z}$ or $\mathbb{N}$ can be easily proven by induction. This is an important property of $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{R}$ called the *"least upper bound"* property.

10.2 *Measuring the Size of a Set*

Definition 10.5: ◌-jections.
Let X and Y be sets, and consider an arbitrary function $f : X \to Y$. We say that f is an *injection iff* f always maps different points to different places.

injection
$$f \text{ is injective} \;:\Leftrightarrow\; (\forall x_1, x_2 \in X)\big(x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)\big)$$
$$\Leftrightarrow\; (\forall x_1, x_2 \in X)\big(f(x_1) = f(x_2) \Rightarrow x_1 = x_2\big)$$

We say f is a *surjection iff* every element of the codomain of f has a preimage.

surjection
$$f \text{ is surjective} \;:\Leftrightarrow\; (\forall y \in Y)(\exists x \in X)\big(f(x) = y\big)$$

bijection When a function is simultaneously injective and surjective, we call it a *bijection*. ⌟

Definition 10.6: Relative Cardinality.
Given arbitrary sets A and B, we say that *the cardinality of A is no greater than the cardinality of B* when there exists an injection from A into B. This intuitively suggests that B has *"enough room"* to contain all of A.

$|A| \leqslant |B|$
$$\big|A\big| \leqslant |B| \;:\Leftrightarrow\; \exists f\big(f : A \to B \;\wedge\; f \text{ is injective}\big)$$

$|A| \geqslant |B|$
$$\big|A\big| \geqslant |B| \;:\Leftrightarrow\; |B| \leqslant \big|A\big|$$

In keeping with popular convention for the use of the $\leqslant$ symbol, we will take the notation $|A| \geqslant |B|$ to simply mean that B can be injected into A. Now, if we have surjection from A to B, that should suggest that A is *"large enough to cover"* all of B without having to duplicate elements of A. Following this idea, we will say that A *has the same cardinality as* B when there exists a bijection from A to B.

$|A| = |B|$
$$\big|A\big| = |B| \;:\Leftrightarrow\; \exists f\big(f : A \to B \;\wedge\; f \text{ is bijective}\big)$$

Finally, we define a strict version of this relation in the typical way you might expect.

$|A| < |B|$
$$\big|A\big| < |B| \;:\Leftrightarrow\; \big|A\big| \leqslant |B| \;\wedge\; \big|A\big| \neq |B|$$

Notice that $|A| < |B|$ means that there exists an injection $f : A \to B$, but that *there do not exist any bijections* from A to B—while this immediately implies that the injection

f cannot be surjective, it is not immediately obvious whether there are any *other* (non-injective) surjections from A to B. We define $|A| > |B|$ to mean simply $|B| < |A|$. ⌟

$|A| > |B|$

Lemma 10.2.
$\forall A \forall B (A \subseteq B \Rightarrow |A| \leqslant |B|)$.

Proof. Let A and B be sets, and assume $A \subseteq B$. Consider the function $f : A \to B$ defined by $f(a) \coloneqq a$ for each $a \in A$.[1] We will prove that f is injective.

Let $a_1 \in A$ and $a_2 \in A$ such that $f(a_1) = f(a_2)$. Recall that, by the definition of f, $f(a_1) = a_1$ and $f(a_2) = a_2$. We therefore have $a_1 = a_2$.

Therefore, f is injective. Since we have proven existence of an injection from A to B, we conclude that $|A| \leqslant |B|$ by definition.

QED

[1] Formally, the function f is determined by the domain $\mathtt{dom}(f) = A$, the codomain $\mathtt{cod}(f) = B$, and the mapping $f = \{(a, a) \mid a \in A\} \subseteq A \times B$. We can see that, for each $a \in A$, there exists a unique $b \in B$ (namely, $b \coloneqq a$) such that $(a, b) \in f$. For each $a \in A$, the existence of $a \in B$ is given by the fact that $A \subseteq B$, and the uniqueness of a is given by the fact that a is the only thing that is equal to a. This justifies our implicit claim that f is a function.

Corollary 10.1.
$\forall A \forall B (|A| \leqslant |B| \Rightarrow \exists C (C \subseteq B \wedge |A| = |C|))$. ⌟

The result from *Corollary 10.1* tells us that an injective function $f : A \to B$ gives us a way of picking out a *"copy"* of A within B—a *substructure* of B that *"looks like"* A.[2] The statement $|A| \leqslant |B|$ tells us, intuitively, that A *"fits inside of"* B; this corollary complements that idea by suggesting that the subset $C \subseteq B$ which is in bijection with A is an example of how A could fit inside of B.

[2] This is an important idea that shows up in many areas of mathematics, and which we will encounter again soon.

The Last Two Axioms of ZFC Set Theory

Axiom 7: Schema of Replacement.
Suppose that φ is a well-formed formula with *exactly two* free variables. The following formula asserts that, if φ behaves like a *function* on some set x, where the *"input-output pairs"* are the two arguments to φ, then the collection of all *"outputs"* of φ when we apply it to elements of x forms a set.

$$\forall X \Big(\forall x \big(x \in X \Rightarrow \exists! y (\varphi(x, y)) \big) \Rightarrow \exists Y \forall y \big(y \in Y \Leftrightarrow \exists x (x \in X \wedge \varphi(x, y)) \big) \Big)$$

Slightly more simply, the sentence above is equivalent to the following assertion. If X is a set and, whenever $x \in X$, there exists *exactly one* y anywhere in the universe such that $\varphi(x, y)$ is *true,* then we can treat φ like a function whose domain is X and collect all of the results into a set.

$$\forall X \Big(\underbrace{\forall x \big(x \in X \Rightarrow \exists! y (\varphi(x, y)) \big)}_{\varphi(x,y) \text{ behaves like a function on } X} \Rightarrow \underbrace{\exists Y \Big(Y = \big\{ y \;\big|\; \exists x (x \in X \wedge \varphi(x, y)) \big\} \Big)}_{Y \text{ is the set of outputs } y \text{ of } \varphi \text{ when the input } x \text{ is in } X} \Big)$$

We frankly do not need to worry too much about this axiom for our purposes. ⌟

Axiom 8: Dichotomy.
The cardinalities of any two sets are *"comparable"* in the sense that we can always obtain an answer to the question *"which of these two sets is bigger?"*

$$\forall x \forall y \big(|x| \leqslant |y| \;\vee\; |y| \leqslant |x| \big)$$

Formally, this asserts that, between any two sets, there always exists an injection from one into the other (though not necessarily from the other into the one). ⌟

10.3 The Algebra of Functions

Definition 10.7: Function Composition.
Given sets X, Y, and Z, and given functions $f : X \to Y$ and $g : Y \to Z$, we can say that f and g are *"compatible"* in a certain sense due to the fact that $\mathfrak{cod}(f) = \mathfrak{dom}(g)$. Whenever f produces an output $f(x) \in Y$ for a given input $x \in X$, the function g can take that output $f(x)$, which lies in its domain Y, and transform it into an output $g(f(x)) \in Z$. This describes a general process for transforming points from X into points in Z, and the transformation has all of the properties of a *function* whose domain is X and codomain is Z.

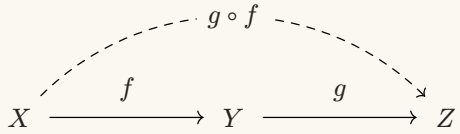

Figure 1. Three sets X, Y, and Z with two composable functions $f : X \to Y$ and $g : Y \to Z$ along with the composition $g \circ f : X \to Z$.

$g \circ f$

We call $g \circ f$ the *composition of f with g.*[1] This is a function $g \circ f : X \to Z$ whose mapping is determined by first applying f to a given input $x \in X$ and then applying g to the intermediate output $f(x)$ as described below.

[1] This is sometimes referred to as the *"composite,"* but it is much more commonly referred to as the *"composition."*

$$g \circ f(x) \coloneqq g\big(f(x)\big)$$ ⌟

If we think about function composition as an *operation* between functions, it seems like we have an algebraic structure similar to a *monoid* like $\mathbb{N}$ with addition. First of all, there is *"an" identity element* with respect to this composition operation.

Definition 10.8: Identity Function.
Given a set X, we define the *identity function on X* to be the function $\mathtt{id}_X : X \to X$ given by the mapping $\mathtt{id}_X(x) \coloneqq x$ for each $x \in X$.

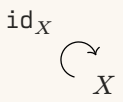

Figure 2. A set X with its identity function $\mathtt{id}_X : X \to X$.

Every set X comes equipped with an identity function $\mathtt{id}_X : X \to X$ over it. ⌟

Lemma 10.3.
For every set X, we have that $|X| = |X|$. ⌟

The function $\mathtt{id}_X : X \to X$ is called the *"identity"* function on X for several reasons, but the most important one for us right now is because $\mathtt{id}_X$ acts like an *identity* with respect to the composition operation in the following sense.

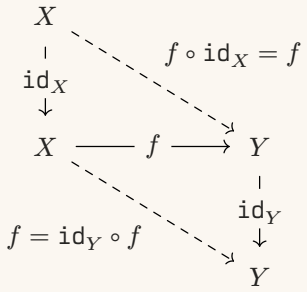

Figure 3. Given a function $f : X \to Y$, composing f with $\mathtt{id}_X$ produces a function that does the same thing that f did, with the same domain and codomain that f originally had, meaning that $f \circ \mathtt{id}_X = f$. The same thing also happens when composing with $\mathtt{id}_Y$ on the other side: $\mathtt{id}_Y \circ f = f$.

Second, the composition of functions is an *associative* operation, which intuitively means that the order in which we compose a finite sequence of composable functions doesn't matter.

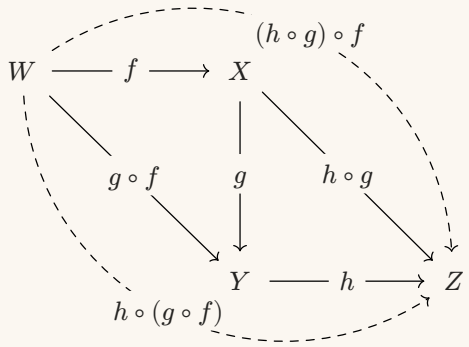

Figure 4. Given three composable functions $f : W \to X$, $g : X \to Y$, and $h : Y \to Z$, then there are two possible ways to think about the sequential composition: $h \circ (g \circ f)$ and $(h \circ g) \circ f$. They turn out to be the same function, *i.e.*, $h \circ (g \circ f) = (h \circ g) \circ f$, which is exactly what it would mean for *"function composition"* to be associative.

This makes the composition $\circ$ operator behave similarly to the addition $+$ operation on the natural numbers, an algebraic structure that we called a *"monoid."* When we were analyzing that structure, we were careful to note that we generally did not have *inverses* with respect to addition $+$ within $\mathbb{N}$. In $\mathbb{Z}$, however, we insisted on having additive inverses for every element, and the resulting structure $(\mathbb{Z}, +)$ we called a *"group."* Since we have a notion of *"identity"* for function composition, we can try to ask the same question: when, if ever, do functions have *"inverses"*?

Definition 10.9: Inverse Functions.
Let X and Y be fixed sets and consider an arbitrary function $f : X \to Y$. We say that a function $g : Y \to X$ going in the opposite direction is *a left-inverse for f* if and only if $g \circ f = \mathtt{id}_X$. Alternatively, we call $h : Y \to X$ a *right-inverse for f* when $f \circ h = \mathtt{id}_Y$.

left-inverse
right-inverse

If a single function $t : Y \to X$ is both a left-inverse $t \circ f = \mathtt{id}_X$ and a right-inverse $f \circ t = \mathtt{id}_Y$ simultaneously, then we call t a *two-sided inverse for f*. In fact, when such a situation happens, there can be only one such two-sided inverse t, so we are justifying in calling it *the inverse of f*, using the notation $f^{-1} : Y \to X$ to denote it. ⌟

f^{-1}

Theorem 10.3.
Let X and Y be *nonempty* sets and consider an arbitrary function $f : X \to Y$. The following three statements govern whether or not f has inverses on either of its sides.

1. f is *injective* $\Leftrightarrow \exists g(g : Y \to X \ \wedge \ g$ is surjective $\wedge \ g \circ f = \mathtt{id}_X)$.
2. f is *surjective* $\Leftrightarrow \exists g(g : Y \to X \ \wedge \ g$ is injective $\wedge \ f \circ g = \mathtt{id}_Y)$.
3. f is *bijective* $\Leftrightarrow \exists! g(g : Y \to X \wedge \ g$ is bijective $\wedge \, g \circ f = \mathtt{id}_X \wedge f \circ g = \mathtt{id}_Y)$. ⌟

Comparing Cardinalities

Theorem 10.4: Cantor-Schröder-Bernstein.
For any X and Y, if there exists an ***injection*** $f : X \to Y$ and an ***injection*** $g : Y \to X$, then there exists a ***bijection*** $h : X \to Y$. This is summarized by the sentence below.

$$\forall x \forall y \Big(\big(|x| \leqslant |y| \ \wedge \ |y| \leqslant |x| \big) \ \Rightarrow \ |x| = |y| \Big)$$

⌟

The *Cantor-Schröder-Bernstein theorem* is an extremely powerful result.[1] It allows us to prove the existence of a bijection without every having to prove that any function is *surjective*, which is typically the more annoying of the two properties to have to

[1] Surprisingly, *Axiom 8* is not required to prove this theorem. In fact, the proof provides a method for *constructing* a bijection out of *any* two injections that point in opposite directions.

deal with. In general, coming up with a surjection between two sets is conceptually more difficult than finding an injection between two sets—this becomes more immediately obvious when we realize that surjectivity is defined by a *universally quantified existential* formula, whereas injectivity is simply a *universally quantified* formula.

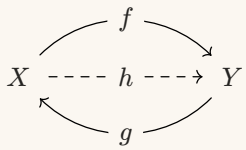

Figure 5. Given two sets X and Y, the *Cantor-Schröder-Bernstein theorem* allows us to prove $|X| = |Y|$ without having to find an explicit bijection from X to Y. Instead, we only need to find two *injective* functions $f : X \to Y$ and $g : Y \to X$ in opposite directions.

You'll also notice that, when we phrase the *Cantor-Schröder-Bernstein* theorem in terms of the relative cardinalities of sets, it seems to look just like

Corollary 10.2: Sets Are Almost Partially Ordered By Their Relative Cardinalities. If we think about the binary predicate $|A| \leqslant |B|$ as a relation, it is *almost* a partial order in the sense that it is reflexive, *almost* antisymmetric, and transitive.

1. $\forall A(|A| \leqslant |A|)$. *Reflexivity*
2. $\forall A \forall B((|A| \leqslant |B| \wedge |B| \leqslant |A|) \Rightarrow |A| = |B|)$. *Antisymmetry (almost)*
3. $\forall A \forall B \forall C((|A| \leqslant |B| \wedge |B| \leqslant |C|) \Rightarrow |A| \leqslant |C|)$. *Transitivity*

All three of these properties follow immediately from theorems we already have. ⌟