

MATHEMATICS 300 — FALL 2018

Introduction to Mathematical Reasoning

H. J. Sussmann

INSTRUCTOR'S NOTES

Contents

Part VI	2
1 Relations and functions	2
1.1 The definition of “relation”	2
1.1.1 The domain and range of a relation	2
1.2 Functions	3
1.2.1 The unique output property	3
1.2.2 The definition of “function”	3
1.2.3 The definition of “value” of a function at a member of its domain	4
1.2.4 When are two functions equal?	4
1.2.5 The definition of “function from a set to a set”	5
1.2.6 Composition of functions	5
1.2.7 The definition of “one-to-one function”	6
1.2.8 The composite of two one-to-one functions	6
1.2.9 The definition of “function onto a set”	6
1.2.10 The composite of two onto functions	7
1.3 The definition of “bijection”	7
1.3.1 The exchange lemma	9
1.3.2 The composite of two bijections	10
1.3.3 The identity function of a set	10
1.3.4 The inverse of a relation	11
1.3.5 The inverse of the inverse	13
1.3.6 The inverse of a bijection	13
1.3.7 Some problems	14
2 Cardinality of sets	15
2.1 Sets with the same cardinality	15
2.2 Finite sets	16
2.2.1 An important notational convention: the sets \mathbb{N}_k	16
2.2.2 Finite lists	17

2.2.3	Finite sets and their cardinality	19
2.2.4	Can we talk about <i>the</i> cardinality of a finite set? The fundamental theorem of finite set cardinality theory	20
2.2.5	Definition of “cardinality” of a finite set	24
2.2.6	A trivial but important lemma	24
2.2.7	Subsets of a finite set	25
2.2.8	The Dirichlet pigeonhole principle	26
2.2.9	Unions of finite sets	28
2.2.10	Sets of subsets	29
2.2.11	Cartesian products of finite sets	29
2.3	Infinite sets	29
2.3.1	Countable sets	32
2.3.2	Do all infinite sets have the same cardinality?	33
2.3.3	Consequences of Cantor’s Theorem	34
2.3.4	Comparing sizes of sets. The Cantor-Schroeder-Bernstein Theorem	35
2.3.5	Infinitely many infinite cardinals	41
3	The paradoxes of set theory: Russell’s paradox and others	42
3.0.6	The Russell paradox	42
3.0.7	The need for Axiomatic Set Theory	45
4	Some more problems	45

Part VI

1 Relations and functions

1.1 The definition of “relation”

Definition 1. A relation is a set of ordered pairs. □

That is:

- a relation is a set R such that every member of R is an ordered pair,
- equivalently, a relation is a set R such that

$$(\forall x \in R)(\exists u)(\exists v) x = (u, v). \quad (1.1)$$

You should picture a relation as set of arrows: for each u and each v , you draw an arrow from u to v to indicate that the pair (u, v) belongs to R .

Another way to think of a relation R is as some kind of device that takes in “inputs” and produces “outputs”. The pairs belonging to R are the ***input-output pairs*** produced by R .

For example, for the relation

$$R = \{ (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (3, 1), (3, 5), (4, 5) \}, \quad (1.2)$$

- the input 1 will produce the outputs 2, 3, and 4,
- the input 2 will produce the outputs 1 and 2,
- the input 3 will produce the outputs 1 and 5,
- the input 4 will produce the output 5.

1.1.1 The domain and range of a relation

Definition 2. If R is a relation, an input of R is an object x such that $(x, y) \in R$ for some y . □

(That is, x is an input of R if $(\exists y)(x, y) \in R$.)

Definition 3. If R is a relation, an output of R is an object y such that $(x, y) \in R$ for some x . \square

(That is, y is an output of R if $(\exists x)(x, y) \in R$.)

Definition 4. If R is a relation and x is an input for R , an output of R for the input x is an object y such that $(x, y) \in R$.

(That is, y is an output of R for x if $(x, y) \in R$.) \square

Definition 5. If R is a relation, the domain of R is the set of all inputs of R . We use $\text{Dom}(R)$ to denote the domain of R . Therefore

$$\text{Dom}(R) = \{x : (\exists y)(x, y) \in R\}. \quad (1.3)$$

Definition 6. If R is a relation, the range of R is the set of all outputs of R . We use $\text{Ran}(R)$ to denote the range of R . Therefore

$$\text{Ran}(R) = \{y : (\exists x)(x, y) \in R\}. \quad (1.4)$$

1.2 Functions

1.2.1 The unique output property

Definition 7. If R is a relation, and x is an input of R (i.e., $x \in \text{Dom}(R)$), we say that x has the unique output property if there is only one output of R for x .

That is, x has the unique output property if

$$(\forall y)(\forall z) \left(((x, y) \in R \wedge (x, z) \in R) \implies y = z. \right) \quad (1.5)$$

Example 1. For the relation R given by (1.2), the input 4 has the unique output property. The inputs 1, 2 and 3 do not. \square

1.2.2 The definition of “function”

Definition 8. A function is a relation f such that every input of f has the unique output property. \square

That is, a set f is a function if

- (i) f is a set of ordered pairs.
- (ii) for all x, y, z , if $(x, y) \in f$ and $(x, z) \in f$ then $y = z$.

In purely formal language, f is a function if

- (I) $(\forall u \in f)(\exists x)(\exists y)u = (x, y)$,
- (II) $(\forall x)(\forall y)(\forall z)\left(\left((x, y) \in f \wedge (x, z) \in f\right) \implies y = z\right)$.

1.2.3 The definition of “value” of a function at a member of its domain

Definition 9. If f is a function and $x \in \text{Dom}(f)$ (that is, x is an input of f), then the value of f at x is the object $f(x)$ such that $f(x)$ is the unique output of f for x . (Recall that the definition of function tells us that every input of f has the unique output property, so the output for x is indeed unique.) \square

1.2.4 When are two functions equal?

The following theorem is the analogue for functions of the theorem on equality of sets: two sets are the same set if they have the same members. Here, the result says: two functions are the same function if they have the same domain and, for each member x of this domain, have the same values at x .

Theorem 1. *Let f, g be functions. Then $f = g$ if and only if $\text{Dom}(f) = \text{Dom}(g)$ and $f(x) = g(x)$ for every $x \in \text{Dom}(f)$.*

Proof. It is clear that if $f = g$ then $\text{Dom}(f) = \text{Dom}(g)$ and $f(x) = g(x)$ for every $x \in \text{Dom}(f)$.

Now assume that $\text{Dom}(f) = \text{Dom}(g)$ and $f(x) = g(x)$ for every $x \in \text{Dom}(f)$.

We want to prove that $f = g$. Since f and g are sets, it suffices to prove that $f \subseteq g$ and $g \subseteq f$. Both proofs are the same, so I will do only one of them.

Let us prove that $f \subseteq g$. Let u be an arbitrary member of f . Since f is a set of ordered pairs, u is an ordered pair, so we may pick x, y such that $u = (x, y)$. Since $(x, y) \in f$, x is an input of f , so $x \in \text{Dom}(f)$, and y is an output of f for x , so $y = f(x)$. Since $\text{Dom}(f) = \text{Dom}(g)$, and $x \in \text{Dom}(f)$,

we see that $x \in \text{Dom}(g)$. Since $g(x) = f(x)$, it follows that $(x, y) \in g$. So $u \in g$

So we have proved that every $u \in f$ is in g . Hence $f \subseteq g$. **Q.E.D.**

1.2.5 The definition of “function from a set to a set”

Definition 10. If f , A , B are sets, we say that f is a function from A to B , and write

$$f : A \rightarrow B,$$

if

1. f is a function,
2. A is the domain of f ,
3. The range of f is a subset of B . □

In other words, “ $f : A \rightarrow B$ ” means “ f is a function, A is the domain of f , and $f(x) \in B$ for every $x \in A$ ”.

1.2.6 Composition of functions

Definition 11. If A , B , C are sets, $f : A \rightarrow B$ and $g : B \rightarrow C$, the composite function of f and g is the function $g \circ f : A \rightarrow C$ given by

$$g \circ f(x) = g(f(x)) \quad \text{for every } x \in A.$$

The following theorem says that the operation of composition of functions satisfies the associative law in the sense that when both $h \circ (g \circ f)$ and $(h \circ g) \circ f$ are defined, they are equal

Theorem 2. Let A , B , C , D be sets, and assume that $f : A \rightarrow B$, $g : B \rightarrow C$, and $h : C \rightarrow D$. Then

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{1.6}$$

Proof. **YOU DO IT.**

Problem 1. *Prove* Theorem 2. □

1.2.7 The definition of “one-to-one function”

Definition 12. A function f is one-to-one (or injective) if whenever two inputs are different, the values of f at those inputs are different as well.

In other words: f is one-to-one if

$$(\forall u \in \text{Dom}(f))(\forall v \in \text{Dom}(f))(u \neq v \implies f(u) \neq f(v)). \quad (1.7)$$

Equivalently, f is one-to-one if

$$(\forall u \in \text{Dom}(f))(\forall v \in \text{Dom}(f))(f(u) = f(v) \implies u = v). \quad (1.8)$$

1.2.8 The composite of two one-to-one functions

Theorem 3. Let A, B, C be sets, and assume that $f : A \rightarrow B$, $g : B \rightarrow C$, and both f and g are one-to-one. Then $g \circ f$ is one-to-one.

Proof. Let $h = g \circ f$. We want to prove that h is one-to-one.

For that purpose, we prove that if $x_1 \in A$, $x_2 \in A$, and $x_1 \neq x_2$, it follows that $h(x_1) \neq h(x_2)$.

Let $y_1 = f(x_1)$, $y_2 = f(x_2)$. Since f is one-to-one and $x_1 \neq x_2$, we can conclude that $y_1 \neq y_2$.

Then, since g is one-to-one and $y_1 \neq y_2$, we can conclude that $g(y_1) \neq g(y_2)$.

But $g(y_1) = g(f(x_1)) = h(x_1)$, and $g(y_2) = g(f(x_2)) = h(x_2)$. So $h(x_1) \neq h(x_2)$, as desired. **Q.E.D.**

1.2.9 The definition of “function onto a set”

Definition 13. A function $f : A \rightarrow B$ is onto B if $B = \text{Ran}(f)$.

In other words, f is onto B if

$$(\forall b \in B)(\exists a \in A)f(a) = b. \quad (1.9)$$

Example 2. Let f be the “squaring a real number” function.

The domain $\text{Dom}(f)$ is \mathbb{R} , the set of all real numbers. And, for $x \in \mathbb{R}$, the value $f(x)$ is given by

$$f(x) = x^2.$$

Then the range $\text{Ran}(f)$ is \mathbb{R}_+ , the set of all nonnegative real numbers. (Reason: every nonnegative real number has a square root.)

Then both statements “ $f : \mathbb{R} \rightarrow \mathbb{R}$ ” and “ $f : \mathbb{R} \rightarrow \mathbb{R}_+$ ” are true. And f is onto \mathbb{R}_+ but f is not onto \mathbb{R} . \square

Remark 1. The previous example shows that the sentence “ f is onto” is meaningless, in the same way as the sentence “ a is divisible” is meaningless.

There is no such thing as “being divisible”. What makes sense is “being divisible by some number”. “Divisible” is a 2-argument predicate: we say things like “ a is divisible by b ”, and we do not say things like “ a is divisible”. Similarly, “is onto” is a 2-argument predicate: we say things like “ f is onto B ”, and we do not say things like “ f is onto”. \square

1.2.10 The composite of two onto functions

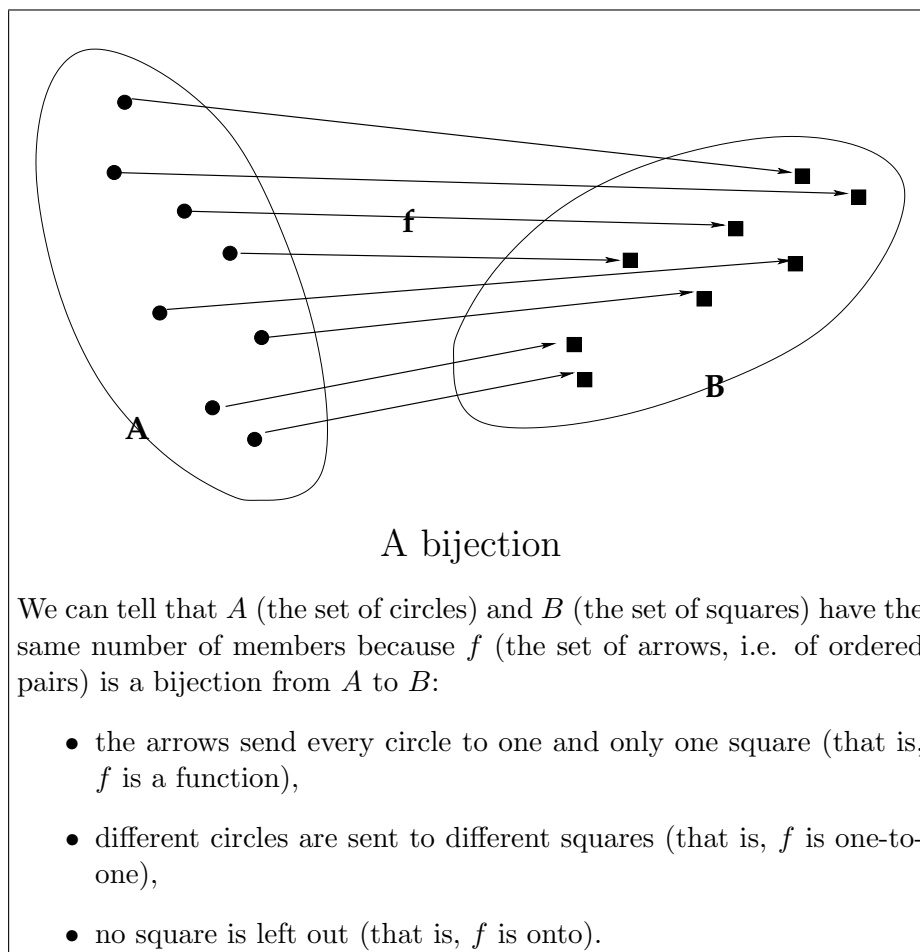
Theorem 4. *Let A, B, C be sets, and assume that $f : A \rightarrow B$, $g : B \rightarrow C$, f is onto B and g is onto C . Then $g \circ f$ is onto C .*

Proof. **YOU DO IT.**

Problem 2. *Prove* Theorem 4. \square

1.3 The definition of “bijection”

Definition 14. A function $f : A \rightarrow B$ is a bijection from A to B if it is one-to-one and onto B . \square



Theorem 5. *The empty set is a bijection from the empty set to the empty set.*

Proof. **YOU PROVE THIS.**

Problem 3. *Prove* Theorem 5.

You have to prove several things:

1. that \emptyset is a relation, i.e., a set of ordered pairs. (So you have to prove that every member of \emptyset is an ordered pair.)
2. that \emptyset is in fact a function. (So you have to prove that

$$(\forall x)(\forall y)(\forall z)\left(\left((x, y) \in \emptyset \wedge (x, z) \in \emptyset\right) \implies y = z\right),$$

3. that the domain of \emptyset is \emptyset ,
4. that $\emptyset : \emptyset \rightarrow \emptyset$,
5. that \emptyset is one-to-one,
6. that \emptyset is onto \emptyset .

□

1.3.1 The exchange lemma

The following theorem is rather simple, but very important. To understand what it says, think of an example. Suppose you have several couples dancing in a large ballroom: every man is dancing with one and only one woman, every woman is dancing with one and only one man. That means that we have in front of our eyes a bijection f from M to W , if M is the set of all the men that are dancing, and W is the set of all the women. The bijection $f : M \rightarrow W$ is defined as follows: for $m \in M$, $f(m)$ is the woman with whom m is dancing.

Theorem 6 then says that, if we are interested in a particular man a and a particular woman b , we can rearrange things so that a will be dancing with b . And the way we do that is as follows: if a is already dancing with b then we do not have to do anything. And if a is not dancing with b , that means that a is dancing with some other woman b' , and b is dancing with some other man a' . So in this case we just have the two couples (a, b') (a', b) **exchange partners**: we remove the pairs¹ (a, b') , (a', b) from f and put, instead, the pairs (a, b) and (a', b') , this gives rise to a new bijection g . And $g(a)$ is b , as we wanted.

Theorem 6. *Suppose A, B are sets, $f : A \rightarrow B$, and f is a bijection from A to B . Then for every $a \in A$ and every $b \in B$ there exists a bijection g from A to B such that $g(a) = b$.*

Proof. **YOU PROVE THIS.**

Problem 4. Prove Theorem 6. Make sure you write a complete, detailed proof of the fact that the function g that you define is a bijection.

¹Do not forget that f is a set of ordered pairs. The fact that a is dancing with b' and a' is dancing with b means that $b' = f(a)$ and $b = f(a')$, and this in turn means that the pairs (a, b') and (a', b) are members of f .

HINT: Read carefully the explanation before the statement of the theorem. All you have to do is repeat the same construction of g in general, for any sets A , B , and bijection f , without mentioning “men”, “women”, and “dancing”.

□

1.3.2 The composite of two bijections

Theorem 7. *Let A , B , C be sets, and assume that f is a bijection from A to B and g is a bijection from B to C . Then $g \circ f$ is a bijection from A to C .*

Proof. Let $h = g \circ f$. We want to prove that h is a bijection.

Since f and g are bijections, they are one-to-one. Hence h is one-to-one by Theorem 3.

Since f and g are bijections, f is onto B and g is onto C . Hence h is onto C by Theorem 4.

So h is one-to-one and onto C . Therefore h is a bijection from A to C .
Q.E.D.

1.3.3 The identity function of a set

Definition 15. If A is a set, the identity function of A , or identity map of A , is the function $1_A : A \rightarrow A$ such that

$$1_A(x) = x \quad \text{for every } x \in A. \quad (1.10)$$

The reason 1_A is called the “identity function” is that it behaves, with respect to the operation of function composition, very much like the number 1 behaves with respect to the operation of multiplication of numbers:

- Multiplying a number by 1 yields the same number: $x \cdot 1 = x$ for every number x .
- Composing a function with 1_A yields the same function, except only for the detail that now we have to be careful about domains: $f \circ 1_A$ makes sense for functions $f : A \rightarrow B$ (for any B), whereas $1_A \circ f$ makes sense for functions $f : B \rightarrow A$ (for any B). The precise result is Theorem 8 below.

Theorem 8. *If A, B are sets, then*

$$f \circ 1_A = f \quad \text{if} \quad f : A \rightarrow B, \quad (1.11)$$

$$1_A \circ f = f \quad \text{if} \quad f : B \rightarrow A. \quad (1.12)$$

Proof. Suppose $f : A \rightarrow B$. Since $1_A : A \rightarrow A$, it follows that $f \circ 1_A : A \rightarrow B$. So both f and $f \circ 1_A$ have domain A .

To prove that the functions $f \circ 1_A$ and f are equal it suffices, according to Theorem 1, to prove that they have the same domain and that they have the same value for every x in that domain.

We already know that $f \circ 1_A$ and f have the same domain, because both have domain A . If $x \in A$, then

$$(f \circ 1_A)(x) = f(1_A(x)) = f(x),$$

So $(f \circ 1_A)(x) = f(x)$. This completes the proof that $\boxed{f \circ 1_A = f}$.

Now let us suppose that $f : B \rightarrow A$. Since $1_A : A \rightarrow A$, it follows that $1_A \circ f : B \rightarrow A$. So both f and $1_A \circ f$ have domain B .

To prove that the functions $1_A \circ f$ and f are equal it suffices, according to Theorem 1, to prove that they have the same domain and that they have the same value for every x in that domain.

We already know that $1_A \circ f$ and f have the same domain, because both have domain B . If $x \in B$, then

$$(1_A \circ f)(x) = 1_A(f(x)) = f(x),$$

So $(1_A \circ f)(x) = f(x)$, and this completes the proof that $\boxed{1_A \circ f = f}$. **Q.E.D.**

Theorem 9. *If A is a set, then 1_A is a bijection from A to A .*

Proof. **YOU DO IT.**

Problem 5. *Prove* Theorem 9. □

1.3.4 The inverse of a relation

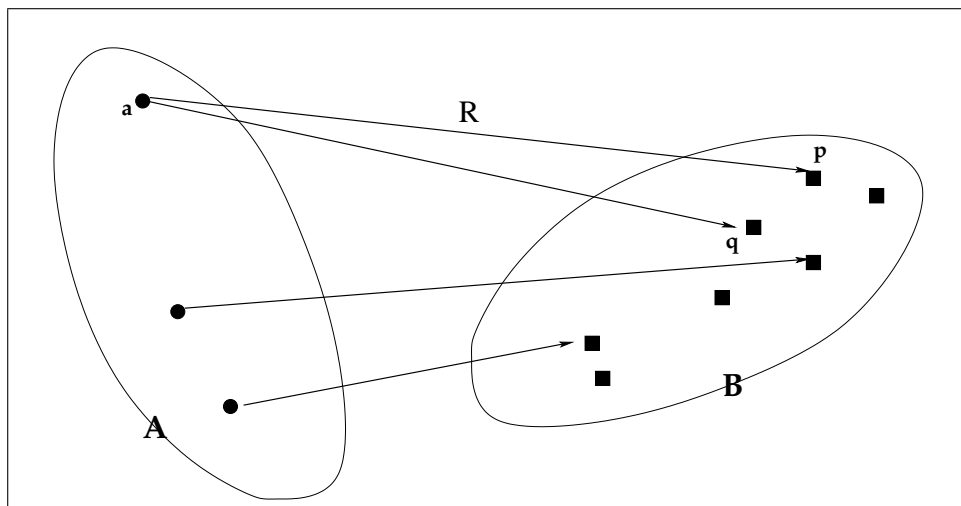
If you think of a relation R as a set of arrows (or of ordered pairs, which amounts to the same thing), then it is clear that can define another relation R^1 by just reversing the arrows of R : for every arrow of R going from a point x to a point y , we put in R^1 an arrow going from y to x . (Or, in terms of ordered pairs: R^1 consists of all the pairs (u, v) such that $(v, u) \in R$.)

Definition 16. If R is a relation, then the inverse of R is the relation R^{-1} given by

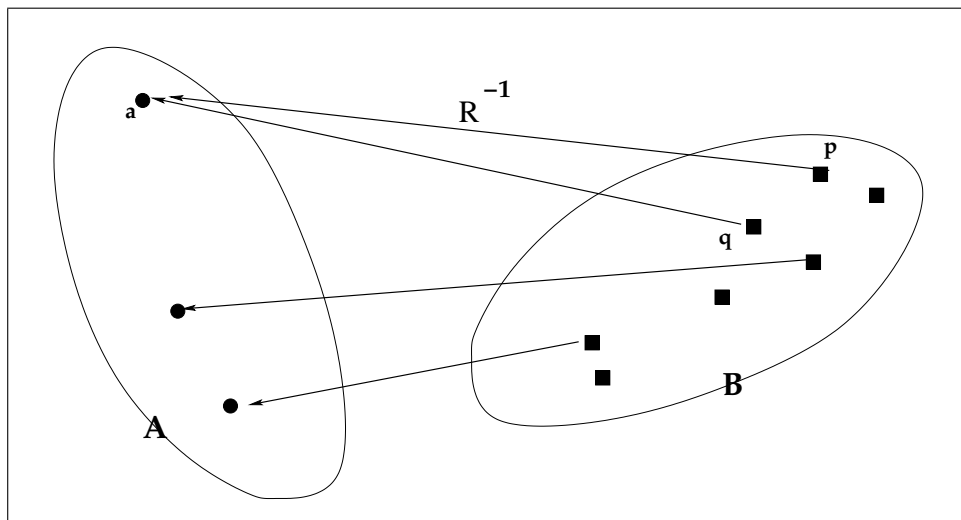
$$R^{-1} = \{(u, v) : (v, u) \in R\}. \quad (1.13)$$

Now that we know what the inverse of a relation is, it is clear that various properties of R correspond to properties of R^{-1} .

For example, consider the following relation (i.e., set of ordered pairs, or set of arrows) R :



and its inverse R^{-1} :



We see that R **is not a function**, because there is one point a which is the source of two different arrows ending at two different points p, q . (So R violates the unique output property: the input u gives rise to two different outputs, p and q .)

What does this tell us about R^{-1} ? It tells us that R^{-1} is not one-to-one, because there are two different squares (i.e, inputs for R^{-1}) that are sources of arrows going to the same point. (Reason: both p and q are sent by R^{-1} to the same point a .)

This tells us that if R is a function then R^{-1} should be one-to-one, and if R is one-to-one then R^{-1} should be a function. In particular, we have:

Theorem 10. *If f is a function, then*

- *the relation f^{-1} is a function if and only if f is one-to-one,*
- *if f is one-to-one, then the domain of f^{-1} is the range of f .*

Proof. **YOU DO THIS.**

Problem 6. *Prove* Theorem 10. □

1.3.5 The inverse of the inverse

Theorem 11. *If R is a relation, then the inverse of the inverse of R is R . That is,*

$$R = (R^{-1})^{-1}.$$

Proof. **YOU DO THIS.**

1.3.6 The inverse of a bijection

Theorem 12. *If A, B are sets, and f is a bijection from A to B , then f^{-1} is a bijection from B to A .*

Proof. We use Theorem 10 repeatedly.

Since f is a bijection, f is one-to-one, so by Theorem 10, f^{-1} is a function. Since the inverse of f^{-1} is f , and f is a function it follows from Theorem 10 that f^{-1} is one-to-one.

Since f is a function from A onto B , the range of f is B , so by Theorem 10 the domain of f^{-1} is B .

Since f^{-1} is a one-to-one, Theorem 10 tells that the domain of $(f^{-1})^{-1}$ is the range of f^{-1} . But $(f^{-1})^{-1} = f$, so the domain of F is the range of f^{-1} . But the domain of f is A , so the range of f^{-1} is A , which means that f^{-1} is a function onto A .

So $f^{-1} : B \rightarrow A$, f^{-1} is one-to-one, and f^{-1} is onto A . So f^{-1} is a bijection from B to A . **Q.E.D.**

1.3.7 Some problems

Problem 7. *Prove* Theorem 11.

HINT: This is trivial. The proof should be no more than one or two lines. \square

Problem 8. *Prove* that if $f : A \rightarrow B$, then f is a bijection from A to B if and only if the following is true:

(#) There exists a function $g : B \rightarrow A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$.
 \square

Problem 9. *Prove or disprove* each of the following statements. (NOTE; two of the statements are true; the other four are false.)

1. If A, B, C are sets, $f : A \rightarrow B$ and $g : B \rightarrow C$ are functions, and $g \circ f$ is one-to-one, then f is one-to-one.
2. If A, B, C are sets, $f : A \rightarrow B$ and $g : B \rightarrow C$ are functions, and $g \circ f$ is one-to-one, then g is one-to-one.
3. If A, B, C are sets, $f : A \rightarrow B$ and $g : B \rightarrow C$ are functions, and $g \circ f$ is onto C then f is onto B .
4. If A, B, C are sets, $f : A \rightarrow B$ and $g : B \rightarrow C$ are functions, and $g \circ f$ is onto C , then g is onto to C .
5. If A, B, C are sets, $f : A \rightarrow B$ and $g : B \rightarrow C$ are functions, and $g \circ f$ is a bijection from A to C , then g is a bijection from B to C .
6. If A, B, C are sets, $f : A \rightarrow B$ and $g : B \rightarrow C$ are functions, and $g \circ f$ is a bijection from A to C , then f is a bijection from A to B .

2 Cardinality of sets

2.1 Sets with the same cardinality

If you look at the picture of a bijection on page 8, you can see right away, *without having to count them*, that the number of squares is exactly the same as the number of circles. This is the crucial insight that leads to the following definition:

Definition 17. Let A, B be sets. We say that

- B has the same number of members as A ,

or that

- B has the same cardinality as A ,

and write

$$\text{card}(A) = \text{card}(B), \quad (2.14)$$

if there exists a bijection from A to B ,

Some authors call two sets “equivalent” if they have the same cardinality in the sense of our definition. I do not like this because in mathematics there are already too many different meanings of the word “equivalent” and I do not want to add one more meaning to the list.

Other authors use the word “equipotent”, and you are welcome to use it if you wish. I just do not like it so I will not use it.

But it is true that “having the same number of members” is an equivalence relation, in the sense of the following theorem:

Theorem 13. *Let A, B, C be sets. Then:*

1. *A has the same cardinality as A ;*
2. *if B has the same cardinality as A , then A has the same cardinality as B ;*

3. if B has the same cardinality as A , and C has the same cardinality as B , then C has the same cardinality as A ,

Proof. To prove that A has the same cardinality as A , we need a bijection from A to A . But we already know such a bijection, namely, the identity function 1_A . We proved in Theorem 9 that 1_A is a bijection from A to A , so A has the same cardinality as A .

Now assume that B has the same cardinality as A . Then there exists a bijection f from A to B . And Theorem 12 tells us that f^{-1} is a bijection from B to A , so A has the same cardinality as B .

Finally, assume that B has the same cardinality as A and C has the same cardinality as B . Then there exist a bijection f from A to B and a bijection g from B to C . Theorem 7 then tells us that $g \circ f$ is a bijection from A to C . So C has the same cardinality as A . **Q.E.D.**

2.2 Finite sets

2.2.1 An important notational convention: the sets \mathbb{N}_k

In what follows we will be making lots of statements about “the natural numbers $1, 2, \dots, k$ ”, that is “all the natural numbers j such that $j \leq k$ ”. So it will be convenient to give a name to the set of all such j s.

THE SETS \mathbb{N}_k (A.K.A. $\{1, 2, \dots, k\}$)

The expression “ \mathbb{N}_k ” stands for the set of all natural numbers that are less than or equal to k . That is,

$$\mathbb{N}_k = \{n \in \mathbb{N} : n \leq k\}. \quad (2.15)$$

Another notation often used for this set is “ $\{1, \dots, k\}$ ”, or “ $\{1, 2, \dots, k\}$ ”. We will use “ \mathbb{N}_k ” when k is a natural number, and also when $k = 0$. (So \mathbb{N}_k makes sense when $k \in \mathbb{N} \cup \{0\}$.)

Naturally, for $n = 0$ the set defined by (2.15) has no members, because there are no natural numbers k such that $k \leq 0$. So $\mathbb{N}_0 = \emptyset$.

Here are other examples:

$$\begin{aligned} \mathbb{N}_1 &= \{1\}, & \mathbb{N}_2 &= \{1, 2\}, & \mathbb{N}_3 &= \{1, 2, 3\}, \\ \mathbb{N}_4 &= \{1, 2, 3, 4\}, & \mathbb{N}_5 &= \{1, 2, 3, 4, 5\}, & \mathbb{N}_6 &= \{1, 2, 3, 4, 5, 6\}, \end{aligned}$$

Then

$j \in \mathbb{N}_k$

is just another way of saying “ $j \in \mathbb{N}$ and $j \leq k$ ”.

2.2.2 Finite lists

Definition 18.

- A function whose domain is the set \mathbb{N}_n for some nonnegative integer² is called a finite list of length n .
- If f is a finite list of length n and $k \in \mathbb{N}_n$, then $f(k)$ is the k -th entry of the list.
- If $f : \mathbb{N}_n \rightarrow A$ (so that every entry $f(k)$ of the finite list f is in A), then f is said to be a list of members of A .
- If $f : \mathbb{N}_n \rightarrow A$ and f is onto A (so that every entry $f(k)$ of the finite list f is in A and every member a of A occurs (in the list, in the sense that $a = f(k)$ for some $k \in \mathbb{N}_n$) then f is said to be a list of **all** the members of A .
- If $f : \mathbb{N}_n \rightarrow A$ and f is one-to-one (so that f has “no repeated entries”, that is, $f(j)$ and $f(k)$ are never equal if $j \neq k$) then f is said to be a list of members of A without repetition. □

Example 3. Let A be the set of all U.S. presidents from George Washington to Donald Trump. Since Donald Trump is the 45-th president, we can define a function $f : \mathbb{N}_{45} \rightarrow A$ by letting

$$f(k) = \text{the } k\text{-th U.S. president, for } k \in \mathbb{N}_{45}.$$

²“Nonnegative integer” means “natural number or zero”. So the empty set \emptyset is a finite list of length 0, because 0 is a nonnegative integer, and $\mathbb{N}_0 = \emptyset$.

So, for example,

$$\begin{aligned}
 f(1) &= \text{George Washington}, \\
 f(2) &= \text{John Adams}, \\
 f(3) &= \text{Thomas Jefferson}, \\
 &\dots \\
 f(16) &= \text{Abraham Lincoln}, \\
 &\dots \\
 f(44) &= \text{Barack Obama}, \\
 f(45) &= \text{Donald Trump}.
 \end{aligned}$$

Then f is a finite list of all the U.S. Presidents. (That is, $f : \mathbb{N}_{45} \rightarrow A$ and f is onto A .)

But f is not a one-to-one function (that is, f is not a list without repetitions), because Grover Cleveland is both the 22nd and the 24th U.S. president, so $f(22) = f(24)$. \square

Theorem 14. *If A is a set, n is a nonnegative integer, and f is a list of length n of all the members of A (that is, $f : \mathbb{N}_n \rightarrow A$ and f is onto A), then there exist a nonnegative integer m such that $m \leq n$, and a list g of length m of all the members of A such that g is without repetition (that is, $g : \mathbb{N}_m \rightarrow A$, g is onto A , and g is one-to-one).*

Proof. **YOU DO THIS.**

Problem 10. *Prove* Theorem 14.

HINT: Eliminate the repetitions that occur in f , one at a time, until there are none left. Read Example 4 below to see how to eliminate one repetition. Then, in your proof, you have to eliminate all the repetitions, one at a time. This means that you will have to do a proof by induction or by well-ordering. \square

Example 4. In Example 3 we saw how to write a list f of length 45 of all the U.S. presidents. But this list is not without repetition (that is, f is not a one-to-one function) because $f(22) = f(24)$.

How can we create a list g of all the U.S. presidents without repetitions?

The trick is to eliminate the repetition. Here is how:

Define $g : \mathbb{N}_{44} \rightarrow A$ by letting

$$g(k) = \begin{cases} f(k) & \text{if } k \leq 23 \\ f(k+1) & \text{if } k > 23 \end{cases}.$$

(In other words: from $k = 1$ up to $k = 23$ we don't change anything, so $g(k)$ is the same as $f(k)$. But for $k = 24$ we start changing things: we do not let $g(24)$ be $f(24)$, because if we did that we would get $g(24) = g(22)$, and there would be a repetition. What we do instead is **skip over** Grover Cleveland, and let $g(24)$ be $f(25)$ (that is we let $g(24) = \text{William McKinley}$). And then we go on: $g(25)$ is $f(26)$, $g(26)$ is $f(27)$, and so on.)

Then g is a list without repetitions of all the U.S. presidents. (That is, g is a one-to-one function from \mathbb{N}_{44} onto A . So g is a bijection from \mathbb{N}_n to A .) \square

2.2.3 Finite sets and their cardinality

Definition 19. If n is a nonnegative integer, then a set A is a set with n members (or a set of cardinality n) if there exists a list of length n of all the members of A without repetitions.

Using function language, we can say the same thing as follows:

if $n \in \mathbb{N} \cup \{0\}$, then A is a set with n members (or “ A is a set of cardinality n ”) if there exists a bijection from \mathbb{N}_n to A .

And then we can define the concepts of “finite set” and “infinite set”.

Definition 20. A set A is finite if it is a set of cardinality n for some nonnegative integer n .
Equivalently, a set A is finite if for some nonnegative integer n there exists a bijection from \mathbb{N}_n to A .

Definition 21. A set A is infinite if it is not a finite set.

Example 5.

1. The empty set is a set of cardinality 0. (Proof: as shown in Theorem 5, the empty set is a bijection from \mathbb{N}_0 to the empty set.)
2. Any singleton $\{a\}$ is a set of cardinality 1. (Proof: if $A = \{a\}$, define $f : \mathbb{N}_1 \rightarrow A$ by letting $f(1) = a$. Then f is a bijection from \mathbb{N}_1 to A .)
3. An unordered pair $\{a, b\}$ is a set of cardinality 2 if and only if $a \neq b$. (Proof: **YOU DO IT.**)
4. An unordered triple $\{a, b, c\}$ is a set of cardinality 3 if and only if $a \neq b$, $a \neq c$, and $b \neq c$. (Proof: **YOU DO IT.**)
5. If $n \in \mathbb{N} \cup \{0\}$, then the set \mathbb{N}_n is a set of cardinality n . (Proof: the function $1_{\mathbb{N}_n}$ is a bijection from \mathbb{N}_n to \mathbb{N}_n .)

Problem 11. Do the two proofs of parts 3 and 4 of Example 5. □

2.2.4 Can we talk about *the* cardinality of a finite set? The fundamental theorem of finite set cardinality theory

Now that we know what a finite set is, we would like to go further and give the following definition: *If A is a finite set and f is a bijection from \mathbb{N}_n to A for some nonnegative integer n , then the number n is said to be the number of members of A , or the cardinality of A .*

But there is a problem. Suppose the following was possible, for some finite set A :

1. m and n are nonnegative integers,
2. there exists a bijection f from \mathbb{N}_m to A ,
3. there exists a bijection g from \mathbb{N}_n onto A ,
4. $m \neq n$

(For example, there could exist bijections from \mathbb{N}_{10} to A and from \mathbb{N}_{12} to A .)

If this happened, then we would not know which number should be called “*the* cardinality of A ”. We would have to talk about “cardinalities of A ”, accepting that a finite set can have several different cardinalities, in the same

way as, for example, an integer can have several factors, and several multiples, so we do not say “6 is *the* factor of 30”, or “30 is *the* multiple of 6”; we say “6 is *a* factor of 30” and “30 is *a* multiple of 6”.

Fortunately, this problem does not occur. The cardinality of a finite set is unique, so we *can* talk about “the” cardinality of a finite set. This is so because of the following theorem:

Theorem 15. *Assume that A is a set and m and n are nonnegative integers such that there exists a bijection f from \mathbb{N}_m to A , and there exists a bijection g from \mathbb{N}_n to A . Then $m = n$.*

Thanks to Theorem 15, if I know that a set A is of cardinality n , then I can say that n is *the* cardinality (or *the* number of members) of A .

Proof of Theorem 15. The key point of the proof is the exchange lemma that we proved earlier as Theorem 6. We will use the exchange lemma to prove Lemma 2, which is essentially the inductive step in the proof by induction of Lemma 1, which easily implies our result.

Let us assume that f is a bijection from \mathbb{N}_m to A , g is a bijection from \mathbb{N}_n to A . We want to prove that $m = n$.

Since g is a bijection from \mathbb{N}_n to A , Theorem 12 tells us that g^{-1} is a bijection from A to \mathbb{N}_n .

Now we have bijections

$$f : \mathbb{N}_m \rightarrow A, \quad g^{-1} : A \rightarrow \mathbb{N}_n.$$

And then Theorem 7 tells us that the composite function $h : \mathbb{N}_m \rightarrow \mathbb{N}_n$, defined by $h = g^{-1} \circ f$, is a bijection from \mathbb{N}_m to \mathbb{N}_n .

So all we need to do prove the following lemma:

Lemma 1. *If $m \in \mathbb{N} \cup \{0\}$, $n \in \mathbb{N} \cup \{0\}$, and there exists a bijection from \mathbb{N}_m to \mathbb{N}_n , then $m = n$.*

To prove Lemma 1, we use the exchange lemma, i.e., Theorem 6.

Suppose there exists a bijection f from \mathbb{N}_m to \mathbb{N}_n . Then by Theorem 6, there exists a bijection k from \mathbb{N}_m to \mathbb{N}_n that has the additional property that $f(m) = n$.

But then, if we remove the pair (m, n) from k , we get a bijection from \mathbb{N}_{m-1} to \mathbb{N}_{n-1} . So we have proved³

Lemma 2. *If $m \in \mathbb{N}$, $n \in \mathbb{N}$, and there exists a bijection from \mathbb{N}_m to \mathbb{N}_n , then there exists a bijection from \mathbb{N}_{m-1} to \mathbb{N}_{n-1} .*

Once we have Lemma 2, we can do a proof of Lemma 1 by induction or by well-ordering.

I will give you the proof using well-ordering.

Call a nonnegative integer n “bad” if there exist a nonnegative integer m such that $m \neq n$ and there exists a bijection from \mathbb{N}_m to \mathbb{N}_n .

We want to prove that there are no bad nonnegative integers. In other words, if we let B be the set of all bad nonnegative integers, we want to prove that B is empty.

We first prove that 0 is not bad.

Proof that 0 is not bad.

- Assume that $\boxed{0 \text{ is bad}}$.
- Since 0 is bad, there exist a nonnegative integer m such that $m \neq 0$, and a bijection f from \mathbb{N}_m to \mathbb{N}_0 .
- Since $m \neq 0$, $m \in \mathbb{Z}$, and $m \geq 0$, it follows that $m \in \mathbb{N}$, so $1 \in \mathbb{N}_m$.
- On the other hand, since f is bijection from \mathbb{N}_m to \mathbb{N}_0 , the domain of f is the set \mathbb{N}_m , so $1 \in \text{Dom}(f)$.
- Then $f(1) \in \mathbb{N}_0$, so $\boxed{\mathbb{N}_0 \neq \emptyset}$.
- But $\boxed{\mathbb{N}_0 = \emptyset}$.

So the assumption that 0 is bad has led us to a contradiction. Hence $\boxed{0 \text{ is not bad}}$.

We now prove that B is empty, and do it by contradiction.

- Assume that $\boxed{B \neq \emptyset}$.

³Why have I suddenly switched from “ $m \in \mathbb{N} \cup \{0\}$ and $n \in \mathbb{N} \cup \{0\}$ ” to “ $m \in \mathbb{N}$ and $n \in \mathbb{N}$ ”? That’s because if $m = 0$ or $n = 0$ then k is empty so the pair (m, n) is not in k and cannot be removed from k .

- Then B is a nonempty subset of \mathbb{Z} , and B is bounded below because every member of B is ≥ 0 .
- So by the well-ordering principle, B has a smallest member b .
- Then $b \in \mathbb{Z}$, $b \geq 0$, and b is bad.
- So in particular $b \neq 0$, because we already know that 0 is not bad.
- Then $b \in \mathbb{N}$.
- Since b is bad, there exist a nonnegative integer m such that $m \neq b$, and a bijection f from \mathbb{N}_m to \mathbb{N}_b ,
- Since $b \in \mathbb{N}$, $b \geq 1$, so $1 \in \mathbb{N}_b$.
- Since f is onto \mathbb{N}_b , and 1 belongs to \mathbb{N}_b , we may pick $x \in \text{Dom}(f)$ such that $f(x) = 1$.
- Then $\text{Dom}(f) \neq \emptyset$, so $\mathbb{N}_m \neq \emptyset$, and then $m \neq 0$.
- Since $m \in \mathbb{Z}$, $m \geq 0$, and $m \neq 0$, it follows that $m \in \mathbb{N}$.
- Since f is a bijection from \mathbb{N}_m to \mathbb{N}_b , and both m and b are natural numbers, we can apply Lemma 2 and conclude that there exists a bijection g from \mathbb{N}_{m-1} to \mathbb{N}_{b-1} .
- But $m \neq b$, so $m - 1 \neq b - 1$.
- So $\boxed{b - 1 \text{ is bad}}$.
- But $\boxed{b - 1 \text{ is not bad}}$, because we are assuming that b is the smallest bad integer.

So the assumption that B is nonempty has led us to a contradiction.

Hence B is empty, and our proof of Lemma 1 is complete.

End of the proof of Theorem 15. As explained before, if f is a bijection from \mathbb{N}_m to A and g is a bijection from \mathbb{N}_n to A , then $g^{-1} \circ f$ is a bijection from \mathbb{N}_m to \mathbb{N}_n . Then Lemma 1 tells that $m = n$, and proof of Theorem 15 is complete. **Q.E.D.**

Problem 12. We have given a proof of Lemma 1 using well-ordering. **Prove** Lemma 1 using induction.

You are allowed to use Lemma 2.

□

2.2.5 Definition of “cardinality” of a finite set

Now that we have proved Theorem 15, we can talk about *the* cardinality (or *the* number of members) of a finite set A .

Definition 22. Let A be a finite set. Then the nonnegative integer n such that

(*) *there exists a bijection from \mathbb{N}_n to A ,*

is called the cardinality of A , or the number of members of A . (The number n exists because A is finite, and is unique thanks to Theorem 15.)

We write “ $\text{card}(A)$ ” to denote the cardinality of A . □

Problem 13. *Prove* that:

1. If A is a finite set, B is a set, and there exists a bijection from B to A , then B is finite and $\text{card}(B) = \text{card}(A)$.
2. If A and B are finite sets, and $\text{card}(A) = \text{card}(B)$, then there exists a bijection from B to A . □

2.2.6 A trivial but important lemma

The following lemma is very obvious but, as all obvious things in this game, needs proof. The lemma says that if you have a set with n members and remove one member then you are left with a set with $n - 1$ members.

Lemma 3. *If A is a finite set, and a is a member of A , then the set $A - \{a\}$ is finite and has cardinality $\text{card}(A) - 1$.*

Proof. Let $n = \text{card}(A)$. Then we may pick a bijection f from \mathbb{N}_n to A . Thanks to the exchange theorem (i.e., Theorem 6) there exists a bijection g from \mathbb{N}_n to A such that $g(n) = a$.

Let $h = g - \{(n, a)\}$. (That is, h is the set of ordered pairs obtained from g by removing the pair (n, a) .)

Then h is a bijection from \mathbb{N}_{n-1} to $A - \{a\}$. (This is easy to prove. **YOU SHOULD PROVE IT.**)

So $A - \{a\}$ is a finite set and $\text{card}(A - \{a\}) = \text{card}(A) - 1$. **Q.E.D.**

2.2.7 Subsets of a finite set

Definition 23. A proper subset of a set A is a subset B of A such that $B \neq A$. \square

Theorem 16. Let A be a finite set, and let B be a proper subset of A . Then B is finite and $\text{card}(B) < \text{card}(A)$.

Proof. We will use induction.

Let $P(n)$ be the predicate

$P(n)$ If A is a finite set with cardinality n , then every proper subset B of A is finite and has cardinality $< n$.

We prove that $P(n)$ is true for every nonnegative integer n , by induction on n starting at $n = 0$.

Basis step. $P(0)$ is true because, if A is a finite set with cardinality 0, then A must be the empty set, so A has no proper subsets, so the statement “if B is a proper subset of A then B is finite and $\text{card}(B) < 0$ ” is vacuously true.

Inductive step. Assume $P(n)$ is true.

We want to prove $P(n + 1)$. That is, we want to prove that

(#) If A is a finite set with cardinality $n + 1$, then every proper subset of A is finite and has cardinality $< n$.

Let A be a finite set with cardinality $n + 1$, and let B be a proper subset of A . We want to prove that

(##) B is finite and $\text{card}(B) < n + 1$.

Since B is a proper subset of A , $B \neq A$, so there exists $a \in A$ such that $a \notin B$. (Reason: if every member of A was in B , it would follow that $A \subseteq B$. Since $B \subseteq A$, this would imply $B = A$, contradicting our assumption that $B \neq A$.) Let $A' = A - \{a\}$. Then $B \subseteq A'$.

By Lemma 3, A' is a finite set with cardinality n .

The set B is either equal to A' or a proper subset of A' .

If $B = A'$, then B is finite and $\text{card}(B) = n$, so (##) holds.

If B is a proper subset of A' then, by the inductive hypothesis $P(n)$, B is finite and $\text{card}(B) < n$, so *a fortiori* (##) holds.

So (##) holds in both cases. Therefore $P(n + 1)$ is true.

This completes the induction, and then the proof of Theorem 16. **Q.E.D.**

The following is a slightly stronger version of Theorem 16, in which the subset B is not required to be proper.

Theorem 17. *Let A be a finite set, and let B be a subset of A . Then*

1. B is finite,
2. $\text{card}(B) \leq \text{card}(A)$,
3. if B is a proper subset of A then $\text{card}(B) < \text{card}(A)$.

Proof. Let A be a finite set with cardinality n , and let B be a subset of A . If B is proper, then Theorem 16 tells us that B is finite and $\text{card}(B) < n$.

If B is not proper, then $B = A$, so B is finite and $\text{card}(B) = n$.

So our desired conclusion holds in both cases.

Q.E.D.

2.2.8 The Dirichlet pigeonhole principle

We are going to use another very important result.

Theorem 18. *Assume that A, B are finite sets and $f : A \rightarrow B$ is a one-to-one function. Then*

1. $\text{card}(A) \leq \text{card}(B)$,
2. $\text{card}(A) < \text{card}(B)$ if and only if f is not onto B .

Theorem 18 is known as the Dirichlet pigeonhole principle (DPHP), for the following reason.

Think of A as a set of pigeons, and B as a set of holes. The function f assigns a hole $f(p)$ to each pigeon p . The fact that f is one-to-one says that different pigeons go to different holes, i.e., that it does not happen that two different pigeons are assigned to the same hole. The theorem then says that

1. there are at least as many holes as there are pigeons,
2. the number of pigeons is equal to the number of holes if and only if every hole is occupied by a pigeon.

Here is another example: suppose that in a classroom there are m students and n seats, and each student is seated, and no seat has more than one student in it. Then the DPHP says that there are at least as many seats as there are students.

Proof of Theorem 18. **YOU DO IT.**

Problem 14. *Prove* Theorem 18.

HINT: Let C be the range of f , so C is a subset of B . Prove first that f is a bijection from A to C . Then use Theorems 16 and 17. \square

Here is another version of the pigeonhole principle. This time, we look at the case when every hole is occupied by at least one pigeon. The conclusion in this case is that

1. there are at least as many pigeons as there are holes.
2. the number of pigeons is equal to the number of holes if and only if no hole is occupied by more than one pigeon.

Theorem 19. *Assume that A, B are finite sets and $f : A \rightarrow B$ is a function onto B . Then*

1. $\text{card}(A) \geq \text{card}(B)$,
2. $\text{card}(A) > \text{card}(B)$ if and only if f is not one-to-one.

Proof. **YOU DO IT.**

Problem 15. *Prove* Theorem 19.

HINT: Construct a one-to-one “hole-to-pigeon” function by picking for each hole h a pigeon p that occupies hole h . Show that this function is one-to-one and then use Theorem 18. \square

Theorem 20. *Let A, B be finite sets. Then*

1. $\text{card}(A) \leq \text{card}(B)$ if and only if there exists a one-to-one function from A to B ,
2. $\text{card}(A) < \text{card}(B)$ if and only if there exists a one-to-one function from A to B but there does not exist a one-to-one function from B to A .

Proof. **YOU DO IT.**

Theorem 21. *Let A, B be finite sets. Then $\text{card}(A) \leq \text{card}(B)$ if and only if there exists a function from B onto A .*

1. $\text{card}(A) \leq \text{card}(B)$ if and only if there exists a function from B onto A ,
2. $\text{card}(A) < \text{card}(B)$ if and only if there exists a function from B onto A but there does not exist a function from A onto B .

Proof. **YOU DO IT.**

Problem 16. *Prove* Theorems 20 and 21. □

2.2.9 Unions of finite sets

Theorem 22. *Let A, B be disjoint⁴ finite sets. Then $A \cup B$ is finite, and*

$$\text{card}(A \cup B) = \text{card}(A) + \text{card}(B). \quad (2.16)$$

Proof. **YOU DO IT.**

Problem 17. *Prove* Theorem 22.

HINT: You can do this by induction with respect to $\text{card}(A)$ or $\text{card}(B)$. Or you can do it directly, by combining a bijection f from \mathbb{N}_m to A and a bijection g from \mathbb{N}_n to B to construct a bijection h from \mathbb{N}_{m+n} to $A \cup B$. □

Theorem 23. *Let A, B be finite sets. Then $A \cup B$ is finite, and*

$$\text{card}(A \cup B) = \text{card}(A) + \text{card}(B) - \text{card}(A \cap B). \quad (2.17)$$

Proof. **YOU DO IT.**

Problem 18. *Prove* Theorem 23.

HINT: Divide $A \cup B$ into three sets C, D, E as follows: $C = \{x : x \in A \wedge x \notin B\}$, $D = \{x : x \in B \wedge x \notin A\}$, and $E = A \cap B$. Then C, D, E are finite by Theorem 17.

Also, $C \cap E = \emptyset$, $C \cup E = A$, $D \cap E = \emptyset$, $D \cup E = B$, $A \cap D = \emptyset$, $A \cup D = A \cup B$. □

⁴Two sets A, B are disjoint if $A \cap B = \emptyset$.

2.2.10 Sets of subsets

Theorem 24. *Let A be a finite set. Then the power set $\mathcal{P}(A)$ is finite, and*

$$\text{card}(\mathcal{P}(A)) = 2^{\text{card}(A)}. \quad (2.18)$$

Proof. **YOU DO IT.**

Problem 19. *Prove* Theorem 24.

HINT: Do it by induction on $n = \text{card}(B)$.

Fix $a \in A$, and let $A' = A - \{a\}$. Then the subsets of A are of two kinds: those that do not contain a as a member and those that do. The subsets of the first kind are exactly the subsets of A' , so by the inductive hypothesis there are 2^{n-1} such sets. If \mathcal{B} is the set of all the subsets of the second kind, then you should construct a bijection from \mathcal{B} to $\mathcal{P}(A')$. Then $\mathcal{P}(A) = \mathcal{P}(A') \cup \mathcal{B}$, $\mathcal{P}(A') \cap \mathcal{B} = \emptyset$, and $\text{card}(\mathcal{B}) = \text{card}(\mathcal{P}(A'))$. \square

2.2.11 Cartesian products of finite sets

Theorem 25. *Let A, B be finite sets. Then $A \times B$ is finite, and*

$$\text{card}(A \times B) = \text{card}(A) \cdot \text{card}(B). \quad (2.19)$$

Proof. **YOU DO IT.**

Problem 20. *Prove* Theorem 25.

HINT: Do it by induction on $m = \text{card}(B)$.

Prove that if $b \in B$, and $B' = B - \{b\}$, then

$$A \times B = (A \times B') \cup (A \times \{b\}), \text{ and } (A \times B') \cap (A \times \{b\}) = \emptyset.$$

and use this in your inductive argument. \square

2.3 Infinite sets

We recall the definition of “infinite set”: *A set is infinite if it is not a finite set.*

Theorem 26. *The set \mathbb{N} of all natural numbers is infinite.*

Idea of the proof. If \mathbb{N} was finite, then it would have some cardinality $n \in \mathbb{N} \cup \{-\}$, and this means that we can put a pigeon for each $k \in \mathbb{N}$ and fit all these pigeons in n holes. But \mathbb{N} has at least m members, for every m . So for every m we can fit m pigeons in n holes. So $m \geq n$ by the Dirichlet pigeonhole principle. So the number n is greater than every natural number. But such an n cannot exist.

Now let us write this down in mathematical language.

Proof. Suppose \mathbb{N} was finite.

Then there exist a natural number n and a bijection f from \mathbb{N}_n to \mathbb{N} .

The inverse function $g = f^{-1}$ is then a bijection from \mathbb{N} to \mathbb{N}_n .

Now let m be an arbitrary natural number.

Define a function g_m , with domain \mathbb{N}_m , by letting

$$g_m(k) = g(x) \text{ for } k \in \mathbb{N}_m.$$

Then $g_m : \mathbb{N}_m \rightarrow \mathbb{N}$ and g_m is one-to-one. (Proof: if $k_1, k_2 \in \mathbb{N}_m$ and $g_m(k_1) = g_m(k_2)$ then $g(k_1) = g(k_2)$, so $k_1 = k_2$ because g is one-to-one.)

Furthermore, $g : \mathbb{N}_m \rightarrow \mathbb{N}_n$. It follows from the Dirichlet pigeonhole principle (Theorem 18) that $n \geq m$.

Since n was an arbitrary real number, we have proved that

$$(\forall m \in \mathbb{N}) n \geq m. \quad (2.20)$$

So we have found a natural number n which is larger than every natural number.

But we know that such a number cannot exist. (That is, (2.20) is impossible. This is easily seen as follows: if (2.20) was true for some $n \in \mathbb{N}$, then we could specialize to $m = n + 1$ and conclude that $n \geq n + 1$. But $n < n + 1$, so we have arrived at a contradiction.)

So the assumption that \mathbb{N} is finite has led us to a contradiction. **Q.E.D.**

Theorem 27. *If A, B are sets, f is a bijection from A to B , and A is infinite, then B is infinite.*

Proof. **YOU DO THIS.**

Problem 21. *Prove* theorem 27. □

Theorem 28. *If A is a set, B is a subset of A , and B is infinite, then A is infinite.* □

Proof. **YOU DO THIS.**

Problem 22. *Prove* theorem 28. □

Theorem 29. *Define a function f from \mathbb{N} to \mathbb{Z} as follows:*

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \in \mathbb{N} \text{ and } n \text{ is even} \\ \frac{1-n}{2} & \text{if } n \in \mathbb{N} \text{ and } n \text{ is odd} \end{cases} . \quad (2.21)$$

Then f is a bijection from \mathbb{N} to \mathbb{Z} .

Proof. **YOU DO THIS.**

Problem 23. *Prove* Theorem 29. □

Problem 24. Let $\mathcal{E}_{>0}$ be the set of all even natural numbers. That is,

$$\mathcal{E}_{>0} = \{n \in \mathbb{N} : 2|n\} , \quad (2.22)$$

and let \mathcal{E} be the set of all even integers, so

$$\mathcal{E} = \{n \in \mathbb{Z} : 2|n\} , \quad (2.23)$$

Construct bijections from \mathbb{N} to $\mathcal{E}_{>0}$ and from \mathbb{N} to \mathcal{E} . □

NOTE: If f is a function whose domain in $\mathbb{N} \times \mathbb{N}$, the set of all order pairs (m, n) of natural numbers, and u is a member of the domain of f , we have to write $f(u)$ for the value of f at u . But u is itself an ordered pair (m, n) , so we should write $f((m, n))$ for the value of f at u . i.e., at (m, n) . But we are going to follow the standard practice of omitting one pair of parentheses and just write “ $f(m, n)$ ”, instead of “ $f((m, n))$ ”.

Theorem 30. *Define a function f from $\mathbb{N} \times \mathbb{N}$ to \mathbb{N} as follows:*

$$f(m, n) = 2^{m-1}(2n - 1) \text{ for } \quad (2.24)$$

Then f is a bijection from $\mathbb{N} \times \mathbb{N}$ to \mathbb{N} .

Proof. **YOU DO THIS.**

Problem 25. *Prove* Theorem 30. □

Problem 26. *Construct* a partition of \mathbb{N} into infinitely many infinite sets. (The definition of “partition” is given in a previous homework.)

HINT: Using the result of Theorem 30, this should just require two or three lines. □

Theorem 31. *There exists a bijection from \mathbb{N} , the set of all natural numbers, to \mathbb{Q} , the set of all rational numbers.*

Proof. This proof is done in the book. You should look it up there.

Problem 27. *Prove* that there exists a bijection from \mathbb{N} to $\mathbb{Q} \times \mathbb{Q}$. □

2.3.1 Countable sets

Definition 24.

- A set A is countable if there exists a one-to-one function $f : A \rightarrow \mathbb{N}$.
- A set A is countably infinite if it is countable and infinite. □

It is clear that

1. \mathbb{N} is countable. (Proof: the function $1_{\mathbb{N}}$ is a bijection from \mathbb{N} to \mathbb{N} , so in particular it is a one-to-one function from \mathbb{N} to \mathbb{N} .)
2. Every subset of a countable set is countable. (Proof: Let A be countable, and let B be a subset of A . Let $f : A \rightarrow \mathbb{N}$ be a one-to-one function. Define $g : B \rightarrow \mathbb{N}$ by letting $g(x) = f(x)$ for each $x \in B$. Then g is one-to-one.)
3. Every finite set is countable.
4. \mathbb{N} is countably infinite. (Proof: \mathbb{N} is countable, and \mathbb{N} is infinite because of Theorem 26.)
5. If A, B are sets and f is a bijection from A to B , then A is countably infinite if and only if B is.
6. \mathbb{Z} , $\mathbb{N} \times \mathbb{N}$, and \mathbb{Q} are countably infinite. (Reason: we have already constructed bijections from \mathbb{N} to \mathbb{Z} and from \mathbb{N} to $\mathbb{N} \times \mathbb{N}$, and Theorem 31 tells us that there exists a bijection from \mathbb{N} to \mathbb{Q} .)

2.3.2 Do all infinite sets have the same cardinality?

The results of Theorems 29, 30 and problems 24 show that sets as diverse as \mathbb{Z} , $\mathcal{E}_{>0}$, \mathcal{E} , and $\mathbb{N} \times \mathbb{N}$, some of which (for example, \mathbb{Z} and $\mathbb{N} \times \mathbb{N}$) appear to be much larger than \mathbb{N} , have in fact the same cardinality as \mathbb{N} .

Could it be that all infinite sets have the same cardinality as \mathbb{N} ?

The answer is a resounding NO!!!. Here is a result which is very easy to prove but has momentous consequences.

Theorem 32 (Cantor) *If X is a set, then there does not exist a function from X onto the power set $\mathcal{P}(X)$.*

Proof. Suppose $f : X \rightarrow \mathcal{P}(X)$ is onto $\mathcal{P}(X)$.

Let

$$S = \{x \in X : x \notin f(x)\}.$$

Then S is a subset of X , so $S \in \mathcal{P}(X)$. Since f is onto $\mathcal{P}(X)$, we can pick $s \in X$ such that $f(s) = S$.

Let us prove that $s \in S$. Suppose $s \notin S$. Then $s \notin f(s)$, so s satisfies the membership criterion for S (" $x \notin f(x)$ "). Therefore $s \in S$. But we are assuming that $s \notin S$, so we have reached a contradiction. So the assumption that $s \notin S$ had led us to a contradiction, and then it follows that $s \in S$.

Now let us prove that $s \notin S$. Suppose $s \in S$. Then $s \in f(s)$, so s does not satisfy the membership criterion for S (" $x \notin f(x)$ "). Therefore $s \notin S$. But we are assuming that $s \in S$, so we have reached a contradiction. So the assumption that $s \in S$ had led us to a contradiction, and then it follows that $s \notin S$.

We have proved that $s \in S \wedge s \notin S$. So we have arrived at a contradiction. The contradiction resulted from assuming that there exists a function f from X onto $\mathcal{P}(X)$. Hence a function f from X onto $\mathcal{P}(X)$ does not exist. **Q.E.D.**

2.3.3 Consequences of Cantor's Theorem

Theorem 32 says that if X is any set then there does not exist a function from X onto $\mathcal{P}(X)$. So in particular there cannot exist a bijection from X to $\mathcal{P}(X)$.

This means that the sets X and $\mathcal{P}(X)$ do not have the same cardinality. Can we say that one of them is “larger” than the other one? The answer is “yes”, $\mathcal{P}(X)$ definitely has a larger cardinality than X . But before we say that, we have to know what it means: ***what does it mean for a set A to have a larger cardinality than that of a set B ?***

In order to answer that question, we first look at what happens for finite sets:

- (I) First of all, if A and B are finite sets, then the cardinalities $\text{card}(A)$ and $\text{card}(B)$ are nonnegative integers, and we know what it means for a nonnegative integer to be larger than another nonnegative integer.
- (II) Second, Theorem 20 tells us that, if A and B are finite sets, then
 1. $\text{card}(A) \leq \text{card}(B)$ if and only if there exists a one-to-one function from A to B ,
 2. $\text{card}(A) < \text{card}(B)$ if and only if there exists a one-to-one function from A to B but there does not exist a one-to-one function from B to A .

The conditions of (II) make perfect sense for infinite sets as well, even though we do not know what “ $\text{card}(A)$ ” means. (We have only defined what it means for two sets A , B to “have the same cardinality”. This does not say that there is some object called “the cardinality of a set”, and that two sets have the same cardinality if and only if that object is the same for both. For finite sets, we were able to define such an object, and it turned out to be a nonnegative integer. For infinite sets, this can be done too, but we have not done it yet, so at this point we do not know what “ $\text{card}(A)$ ” means. All we know is what “ $\text{card}(A) = \text{card}(B)$ ” means. And in the next section we are going to assign a meaning to “ $\text{card}(A) \leq \text{card}(B)$ ”, “ $\text{card}(A) < \text{card}(B)$ ”, “ $\text{card}(A) \geq \text{card}(B)$ ”, and “ $\text{card}(A) > \text{card}(B)$ ”, but we will still not know what “ $\text{card}(A)$ ” means.

2.3.4 Comparing sizes of sets. The Cantor-Schroeder-Bernstein Theorem

We follow the idea of Theorem 20. (We could also follow the second one, and it works, but there are some complications.)

Definition 25. Suppose that Let A, B are sets.

- We say that A has cardinality smaller than or equal to that of B , or that B has cardinality larger than or equal to that of A , and write

$$\text{card}(A) \leq \text{card}(B),$$

or

$$\text{card}(B) \geq \text{card}(A),$$

if there exists a one-to-one function from A to B .

- We say that A has cardinality strictly smaller than that of B , or that B has cardinality strictly larger than that of A , or that and write

$$\text{card}(A) < \text{card}(B)$$

or

$$\text{card}(B) > \text{card}(A),$$

if there exists a one-to-one function from A to B , but there does not exist a bijection from A to B . \square

As an important example of the use of these definitions, we prove the following theorem.

Theorem 33 (*Cantor*) *If X is a set, then the cardinality of X is strictly smaller than the cardinality of the power set $\mathcal{P}(X)$. That is:*

$$\text{card}(X) < \text{card}(\mathcal{P}(X)) . \quad (2.25)$$

Proof. First, we show that $\text{card}(X) \leq \text{card}(\mathcal{P}(X))$, by constructing a one-to-one function $f : X \rightarrow \mathcal{P}(X)$.

This is very easy: define $f : X \rightarrow \mathcal{P}(X)$ by letting

$$f(x) = \{x\} \quad \text{for } x \in X .$$

Then f is clearly one-to-one. (Proof: suppose $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$; then $\{x_1\} = \{x_2\}$; since $x_2 \in \{x_2\}$, it follows that $x_2 \in \{x_1\}$; but x_1 is the only member of $\{x_1\}$; so $x_2 = x_1$.)

So we have constructed a one-to-one function from X to $\mathcal{P}(X)$.

On the other hand, Theorem 32 tells us that there does not exist a function from X onto $\mathcal{P}(X)$. In particular, there does not exist a bijection from X to $\mathcal{P}(X)$.

This proves that (2.25) holds.

Q.E.D.

The relations “ \leq ”, “ \geq ”, “ $<$ ”, “ $>$ ”, should behave like their homonyms⁵ the relations “ \leq ”, “ \geq ”, “ $<$ ”, “ $>$ ”, between real numbers.

⁵**Homonyms** are words or symbols that are spelled or written identically but have different meanings. For example, “crane”, a mechanical lifting machine, and “crane”, a bird, are homonyms. Mathematics is full of homonyms. For example, “ $<$ ” as a binary relation between real numbers, and “ $<$ ” as a binary relation between cardinalities, are different meanings of the symbol “ $<$ ”.

So, for example, we would expect that there is a true theorem about inequalities among cardinals corresponding to each of the following properties of inequalities between real numbers:

- (R1) If $a, b \in \mathbb{R}$, then $a \leq b$ if and only if $b \geq a$.
- (R2) If $a, b \in \mathbb{R}$, then $a < b$ if and only if $b > a$.
- (R3) If $a, b, c \in \mathbb{R}$, $a \leq b$, and $b \leq c$, then $a \leq c$.
- (R4) If $a, b \in \mathbb{R}$, then $a \leq b$ if and only if $a < b$ or $a = b$.
- (R5) If $a, b \in \mathbb{R}$ and $a < b$, then $a \neq b$.
- (R6) If $a, b \in \mathbb{R}$, $a \leq b$, and $b \leq a$, then $a = b$.
- (R7) If $a, b, c \in \mathbb{R}$, $a \leq b$, $b \leq c$, and either $a < b$ or $b < c$, then $a < c$.
- (R8) If $a, b \in \mathbb{R}$, then either $a \leq b$ or $b \leq a$.

It turns out that *the analogues of the nine properties listed above are all true; some are trivially true, but others are not at all obvious and their proofs require a lot of work.*

Let us start with the obvious ones: the analogues of (R1), (R2), (R3), (R4), (R5) for cardinalities are trivially true. We say that in the following theorem:

Theorem 34. *Let A, B, C be sets. Then*

- (C1) $\text{card}(A) \leq \text{card}(B)$ if and only if $\text{card}(A) \geq \text{card}(B)$.
- (C2) $\text{card}(A) < \text{card}(B)$ if and only if $\text{card}(B) > \text{card}(A)$.
- (C3) If $\text{card}(A) \leq \text{card}(B)$ and $\text{card}(B) \leq \text{card}(C)$ then $\text{card}(A) \leq \text{card}(C)$
- (C4) $\text{card}(A) \leq \text{card}(B)$ if and only if either $\text{card}(A) < \text{card}(B)$ or $\text{card}(A) = \text{card}(B)$.
- (C5) If $\text{card}(A) < \text{card}(B)$ then it's not true that $\text{card}(A) = \text{card}(B)$.

Proof. Statement (C1) is true because Definition 25 tells us “ $\text{card}(A) \leq \text{card}(B)$ ” and “ $\text{card}(B) \geq \text{card}(A)$ ” are two ways of writing the same thing.

Similarly, statement (C2) is true because Definition 25 tells us “ $\text{card}(A) < \text{card}(B)$ ” and “ $\text{card}(B) > \text{card}(A)$ ” are two ways of writing the same thing.

To prove (C3), assume that $\text{card}(A) \leq \text{card}(B)$ and $\text{card}(B) \leq \text{card}(C)$. This means, according to Definition 25, that there exist one-to-one functions $f : A \rightarrow B$ and $g : B \rightarrow C$. Then the composite $g \circ f$ is a function from

A to C , and Theorem 3 tells us that $g \circ f$ is one-to-one. Hence $g \circ f$ is a one-to-one function from A to C . So $\text{card}(A) \leq \text{card}(C)$.

Next, we prove (C4). We have to prove that

$$\text{card}(A) \leq \text{card}(B) \iff (\text{card}(A) < \text{card}(B) \vee \text{card}(A) = \text{card}(B)). \quad (2.26)$$

To prove (2.26) we prove the implications

$$\text{card}(A) \leq \text{card}(B) \implies (\text{card}(A) < \text{card}(B) \vee \text{card}(A) = \text{card}(B)). \quad (2.27)$$

and

$$(\text{card}(A) < \text{card}(B) \vee \text{card}(A) = \text{card}(B)) \implies \text{card}(A) \leq \text{card}(B). \quad (2.28)$$

Finally, we have to prove (C5). But (C5) is completely trivial: by definition, “ $\text{card}(A) < \text{card}(B)$ ” means “there is a one-to-one function from A to B but there is no bijection from A to B ”. So in particular if $\text{card}(A) < \text{card}(B)$ then there is no bijection from A to B , so it’s not true that $\text{card}(A) \leq \text{card}(B)$. **Q.E.D.**

To prove the analogues of (R6), (R7), and (R8), we need deeper results from set theory.

Let us start with (C6), the analogue for cardinalities of property (R6). We need to prove that

$$(\&) \text{ if } \text{card}(A) \leq \text{card}(B) \text{ and } \text{card}(B) \leq \text{card}(A) \text{ then } \text{card}(A) = \text{card}(B).$$

Translating this into English, (&) says:

(&&) *if there exist one-to-one functions $f : A \rightarrow B$, $g : B \rightarrow A$, then there exists a bijection from A to B .*

And we have to answer the question: **is (&&) true?**

For finite sets the answer is undoubtedly “yes”: if A , B have cardinalities m , n , and there exist one-to-one functions $f : A \rightarrow B$, $g : B \rightarrow A$, then it follows from the Dirichlet pigeonhole principle that $m \leq n$ and $n \leq m$, so $m = n$, and this implies that there exists a bijection from A to B .

Actually, even more can be proved:

Theorem 35. *If A , B are finite sets, and $f : A \rightarrow B$, $g : B \rightarrow A$ are one-to-one functions, then f is a bijection from A to B and g is a bijection from B to A .*

Proof. The range $\text{Ran}(f)$ of f is a subset of A , and f is a bijection from A to $\text{Ran}(f)$.

Assume that $\text{Ran}(f)$ is a proper subset of B , then $\text{card}(\text{Ran}(f)) < \text{card}(B)$ by Theorem 16, while on the other hand $\text{card}(\text{Ran}(f)) = \text{card}(A)$, so $\text{card}(A) < \text{card}(B)$, and then the Dirichlet pigeonhole principle tells us that there cannot exist a one-to-one function from B to A . Since g is a one-to-one function from B to A , the assumption that $\text{Ran}(f)$ is a proper subset of B has led us to a contradiction.

So $\text{Ran}(f) = B$, and then f is onto B . A similar argument proves that g is onto A . **Q.E.D.**

For infinite sets, the analogue of Theorem 35 is not true. Here is a simple example. Let \mathcal{E} be the set of all even natural numbers. Define $f : \mathbb{N} \rightarrow \mathcal{E}$, and $g : \mathcal{E} \rightarrow \mathbb{N}$ by letting

$$\begin{aligned} f(n) &= 4n & \text{for } n \in \mathbb{N}, \\ g(n) &= n & \text{for } n \in \mathcal{E}. \end{aligned}$$

Then $f : \mathbb{N} \rightarrow \mathcal{E}$, and $g : \mathcal{E} \rightarrow \mathbb{N}$ are one-to-one functions but neither one is a bijection.

And yet, even though f and g are not themselves bijections, a bijection from \mathbb{N} to \mathcal{E} does exist: just define $h : \mathbb{N} \rightarrow \mathcal{E}$ by letting $h(n) = 2n$, and it is clear that h is a bijection from \mathbb{N} to \mathcal{E} .

It turns out that what happened in this example actually does happen in general.

Theorem 36 *(Cantor-Schroeder-Bernstein)* If A , B are sets and $f : A \rightarrow B$, $g : B \rightarrow A$ are one-to-one functions, then there exists a bijection from A to B .

The proof of this theorem is given in the book, and I am not going to do it here.

Thanks to the Cantor-Schroeder-Bernstein theorem, statement (&&) above is true, and this implies that condition (C6), the analogue for sets of property (R6), is true. So we get the following theorem:

Theorem 37. *Let A, B be sets such that $\text{card}(A) < \text{card}(B)$. Then $\sim \text{card}(A) = \text{card}(B)$, that is, B does not have the same cardinality as A .*

(We do not give a proof, because we have already proved this.)

We now turn to (C7), the cardinality analogue of (R7).

Theorem 38. *If A, B, C are sets such that $\text{card}(A) \leq \text{card}(B)$, $\text{card}(B) \leq \text{card}(C)$, and one of the inequalities is strict (that is, either $\text{card}(A) < \text{card}(B)$ or $\text{card}(B) < \text{card}(C)$) then $\text{card}(A) < \text{card}(C)$.*

Proof. Assume that $\text{card}(A) \leq \text{card}(B)$ and $\text{card}(B) \leq \text{card}(C)$. Then we pick one-to-one functions $f : A \rightarrow B$ and $g : B \rightarrow C$.

We want to prove that if one of the inequalities is strict, then $\text{card}(A) < \text{card}(C)$.

We already know from Theorem 34, condition (C3), that $\text{card}(A) \leq \text{card}(C)$, and this means that either $\text{card}(A) < \text{card}(C)$ or $\text{card}(A) = \text{card}(C)$.

So, in order to prove that $\text{card}(A) < \text{card}(C)$, we have to exclude the possibility that $\text{card}(A) = \text{card}(C)$.

Suppose that $\text{card}(A) = \text{card}(C)$. Then there exist bijections h, k , from A to C and from C to A , respectively.

So we have functions

$$f : A \rightarrow B, \quad g : B \rightarrow C, \quad h : A \rightarrow C, \quad k : C \rightarrow A,$$

such that f and g are one-to-one and h and k are bijections.

Since the composite of two one-to-one functions is one-to-one, the composite function $k \circ g : B \rightarrow A$ is one-to-one. So we have one-to-one functions $f : A \rightarrow B$, $k \circ g : B \rightarrow A$. Then the Cantor-Schroeder-Bernstein theorem tells us that there exists a bijection from A to B so $\boxed{\text{card}(A) = \text{card}(B)}$.

Similarly, the composite function $f \circ k : C \rightarrow B$ is one-to-one. So we have one-to-one functions $g : B \rightarrow C$, $f \circ k : C \rightarrow B$. Then the Cantor-Schroeder-Bernstein theorem tells us that there exists a bijection from B to C so $\boxed{\text{card}(B) = \text{card}(C)}$.

So we have proved that both $\text{card}(A) = \text{card}(B)$ and $\text{card}(B) = \text{card}(C)$. And this contradicts the assumption either $\text{card}(A) < \text{card}(B)$ or $\text{card}(B) < \text{card}(C)$.

This contradiction arose from assuming that $\text{card}(A) = \text{card}(C)$. Hence it is not true that $\text{card}(A) = \text{card}(C)$. **Q.E.D.**

Finally, we need the cardinality analogue of (R8). This is indeed true, and we state it as a theorem:

Theorem 39. *If A, B are sets, then $\text{card}(A) \leq \text{card}(B)$ or $\text{card}(B) \leq \text{card}(A)$. That is, one of the following is true:*

- *there exists a one-to-one function from A to B ,*
- *there exists a one-to-one function from B to A .*

The proof of this theorem requires methods that are above the level of this course. So here we just leave the theorem without proof.

2.3.5 Infinitely many infinite cardinals

It follows from Theorem 33 that

Theorem 40 *The power set $\mathcal{P}(\mathbb{N})$ is not countable.*

So there are at least two different infinite cardinals: that of \mathbb{N} (and all countably infinite sets), and that of $\mathcal{P}(\mathbb{N})$. But then it is clear that there are many more, in fact infinitely many more, because we can consider the sets

$$\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathcal{P}(\mathcal{P}(\mathbb{N})), \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N}))), \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))), \dots$$

That is, we can define, inductively,

$$\begin{aligned} \mathcal{P}_0(\mathbb{N}) &= \mathbb{N}, \\ \mathcal{P}_{n+1}(\mathbb{N}) &= \mathcal{P}(\mathcal{P}_n(\mathbb{N})) \quad \text{for } n \in \mathbb{Z}, n \geq 0, \end{aligned}$$

and in this way we obtain an infinite sequence $\left(\mathcal{P}_n(\mathbb{N})\right)_{n=1}^{\infty}$ of sets each one of which has cardinality strictly larger than the previous one.

But the story does not end there. We can construct an enormous set $\mathcal{P}_{\infty}(\mathbb{N})$, defined by

$$\mathcal{P}_{\infty}(\mathbb{N}) = \bigcup_{n=0}^{\infty} \mathcal{P}_n(\mathbb{N}),$$

and then start again, constructing the sets $\mathcal{P}_n\left(\left(\mathcal{P}_{\infty}(\mathbb{N})\right)\right)$, and so on.

The result is an infinite tower of infinite towers of infinite towers ..., of infinite cardinals.

3 The paradoxes of set theory: Russell's paradox and others

But in the story I have been telling you there are serious problems. We said that the cardinality of the power set of a set X is strictly larger than the cardinality of X . But *what if we apply this to the set of all sets?*

That is, let

$$X = \{x : x \text{ is a set}\}.$$

Then we have seen that $\boxed{\text{card}(\mathcal{P}(X)) > \text{card}(X)}$.

But $\mathcal{P}(X)$ is a set whose members are sets, so $\mathcal{P}(X)$ is a subset of X , and then $\boxed{\text{card}(\mathcal{P}(X)) \leq \text{card}(X)}$.

So we get a contradiction! And this time we have done nothing wrong!

3.0.6 The Russell paradox

Here is another way to get a contradiction in Set Theory. It is the famous ***Russell paradox***.

Let X be the set of all sets that are not members of themselves. That is, we let

$$X = \{x : x \text{ is a set} \wedge x \notin x\}.$$

Let us prove that $X \in X$. Suppose that $\boxed{X \notin X}$. Then is a set that is not a member of itself. So X satisfies the membership criterion for X . Hence $X \in X$. But we are assuming that $X \notin X$. So $\boxed{X \in X \wedge X \notin X}$, which is a contradiction.

So the assumption that $X \notin X$ has led us to a contradiction. Hence $\boxed{X \in X}$.

Now let us prove that $X \notin X$. Suppose that $\boxed{X \in X}$. Then is a set that is a member of itself. So X does not satisfy the membership criterion for X . Hence $X \notin X$. But we are assuming that $X \in X$. So $\boxed{X \in X \wedge X \notin X}$, which is a contradiction.

So the assumption that $X \in X$ has led us to a contradiction. Hence $\boxed{X \notin X}$.

So we have proved that $X \in X \wedge X \notin X$, which is a contradiction.
So

In set theory it is possible to prove a contradiction.

And, once you have proved a contradiction, then everything can be proved, because:

In a theory in which it is possible to prove a contradiction, it is possible to prove every statement, whether it is true or false.

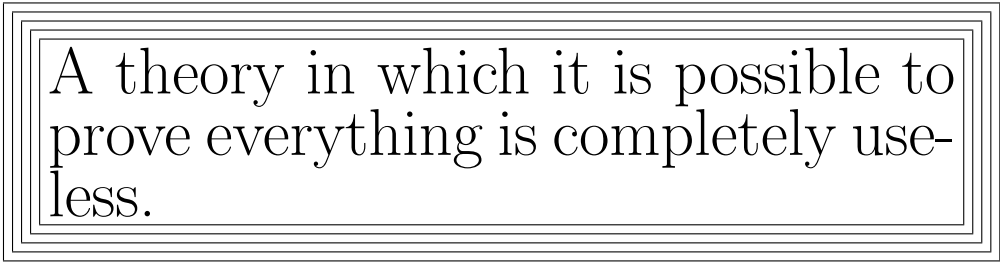
This is so for a simple reason: take any statement S you want (for example, S could be “ $2 + 2 = 5$ ”, or “6 is a prime number”, or “the 45th president of the U.S. is Hillary Clinton”, or “the Earth is flat and rests on top of a giant turtle”).

Let us prove S , assuming we know how to prove a contradiction C .

To prove S , you do it by contradiction: start with “Assume $\sim S$ ”. Then insert your proof of C , ending with C . So, you see, assuming $\sim C$ we got a contradiction. Therefore we have proved S . **Q.E.D.**

What is wrong with this. Isn't it wonderful that we can prove everything? Shouldn't we celebrate? No more having to study hard to learn how to prove things!!! We have an easy method for proving everything, without doing any work!

The trouble is,



A theory in which it is possible to prove everything is completely useless.

The purpose of writing proofs is to make sure that the mathematical statements we write are true. If we prove something, then we can be sure it is true, because the rules of logic are designed to guarantee that everything we prove is true.

The catch is: if we can prove everything, including statements that are false, then the fact that we have proved something tells us nothing: it could be true or it could be false.

Think of theorem-proving as analogous to smoke-detecting.

A naïve person might think that, since the purpose of a smoke detector is to detect smoke, a device that rings every time there is smoke is precisely what is desired of a good smoke detector.

However, it may happen that the device rings every time there is smoke because it rings all the time, whether there is smoke or not.

Such a “smoke detector” is useless. What you want is a device that rings when there is smoke and does not ring when there is no smoke. That way, when you hear the device ring, you know that there is smoke.

Similarly, theorem-proving is useful if you can prove the statements that are true, and you cannot prove those that are not true. If you can prove too much, if you can prove assertions that are false, then your theorem-proving has absolutely no value.

3.0.7 The need for Axiomatic Set Theory

The trouble with the paradoxes of the last section arose from the fact that we made indiscriminate use of the Axiom of Set Formation. The axiom says that “if we take any one-variable predicate $P(x)$, we can form the set $\{x : P(x)\}$.” Using this, we created the sets $\{x : x \text{ is a set}\}$ (the set of all sets) and $\{x : x \text{ is a set} \wedge x \notin x\}$ (the set of all sets that do not belong to themselves). And these sets led us to contradictions.

The solution that mathematicians have adopted is to develop “Axiomatic Set Theory” (AST). In AST, axioms are stated that tell us under what conditions it is possible to create a set. And the axioms are carefully chosen so that the sets that caused us trouble cannot be formed.

But this should be the subject of another course.

4 Some more problems

In the following problems, if a and b are real numbers, we write “[a, b]” to denote the closed interval $\{x \in \mathbb{R} : a \leq x \leq b\}$, and “(a, b)” to denote the open interval $\{x \in \mathbb{R} : a < x < b\}$.

Problem 28. *Prove* that if X is a set then there does not exist a one-to-one function $f : \mathcal{P}(X) \rightarrow X$. □

Problem 29.

1. Let f, g be the functions defined by

- (i) $\text{Dom}(f) = \mathbb{R}$,

- (ii) $f(x) = \frac{x}{\sqrt{1+x^2}}$ for $x \in \mathbb{R}$,

- (iii) $\text{Dom}(g) = (-1, 1)$,

- iv) $g(y) = \frac{y}{\sqrt{1-y^2}}$ for $y \in (-1, 1)$,

- (a) **Prove** that $f : \mathbb{R} \rightarrow (-1, 1)$, $g : (-1, 1) \rightarrow \mathbb{R}$, $g \circ f = I_{\mathbb{R}}$, and $f \circ g = I_{(-1, 1)}$.

- (b) **Conclude** from this that \mathbb{R} and the open interval $(-1, 1)$ have the same cardinality.

2. If $a, b \in \mathbb{R}$ and $a < b$, let $f_{a,b}$ be the function with domain $(-1, 1)$, given by $f_{a,b}(x) = a + \frac{1}{2}(b-a)(x+1)$ for $x \in (-1, 1)$. **Prove** that f is a bijection from $(-1, 1)$ to (a, b) , and **conclude** from this that \mathbb{R} and the interval (a, b) have the same cardinality.

Problem 30. **Construct** a bijection from the closed interval $[0, 1]$ to the open interval $(0, 1)$. (Recall that $[0, 1]$ is the set $\{x \in \mathbb{R} : 0 \leq x \leq 1\}$, and $(0, 1)$ is the set $\{x \in \mathbb{R} : 0 < x < 1\}$.)

HINT: Imagine an infinite hotel⁶ H , in which

- the points of the open interval $(-1, 1)$ are the numbers of the rooms of H ; each room has a number which is a member of $(0, 1)$, and for each member x of $(-1, 1)$ there is a room no. x . (So for example there is room no. 0.00001, room no. 0.3, room no. $\frac{1}{\pi}$, room no. $\frac{23}{77}$, room no. 0.9, room no. 0.99, room no. 0.999, etc. But of course there is no room 0 or room 1, because the rooms of H correspond to the members of the open interval $(0, 1)$, so 0 and 1 are not possible room numbers.)
- At some time, the hotel has a set of guests, also labeled by the members of the open interval $(0, 1)$, and each guest occupies the room with the same label (so guest no. 0.13 occupies room no. 0.13, guest no. $\frac{1}{\pi}$ occupies room no. $\frac{1}{\pi}$, and so on).
- Suddenly, two new guests, labeled 0 and 1, arrive and ask for rooms.
- And nobody, neither the old guests nor the new ones, is willing to share a room.
- So you have to accomodate these two new guests, while making sure that none of the old guests is left without a room.

If we were dealing with a finite hotel, this would be impossible. If a hotel has 100 rooms, all of which are occupied, and two new guests arrive and ask for a room, there is no way to oblige them.

But for an infinite hotel it can be done. The way you can do this is by finding within $(0, 1)$ an infinite sequence of rooms (for example, rooms $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}$, and so on), and then move the guests in those rooms to other rooms also in the sequence, making room for the two new guests.

⁶This is known as “Hilbert’s hotel”.