Cardinality of infinite sets

The cardinality $|A|$ of a finite set A is simply the number of elements in it. When it comes to infinite sets, we no longer can speak of the number of elements in such a set. We can, however, try to match up the elements of two infinite sets A and B one by one. If this is possible, i.e. if there is a bijective function $h : A \to B$, we say that A and B are of the *same cardinality* and denote this fact by $|A| = |B|$.

If two (finite or infinite) sets A and B are not of the same cardinality, we can try to compare which one of the two sets *has at least as many elements as* the other. There are basically two ways of doing that: if we can pair up every element a of the set A with one and only one element b of the set B so that no two elements of B are paired with the same element of A (i.e. if there is an injective function $f : A \to B$), then B must have at least as many elements as A. Alternatively, one could detect this by exhibiting a surjective function $g : B \to A$, because that would mean that there are enough elements in B to cover all elements in A.

It is a consequence of the following two theorems that these two notions of "has at least as many elements as" agree. The first one we proved in class, the second we will prove here.

Theorem 1. *Let $f : A \to B$ be a function. The following four statements are equivalent:*

(a) $f : A \to B$ *is injective.*

(b) $f(S \cap T) = f(S) \cap f(T)$ *for all* $S, T \subseteq A$.

(c) $f^{-1}(f(S)) = S$ *for all* $S \subseteq A$.

(d) *There is a function* $g : B \to A$ *such that* $g \circ f = id_A$.

Remark. Recall that the identity function id_A is defined by $id_A(a) = a$ for all $a \in A$.

Theorem 2. *Let $g : B \to A$ be a function. The following three statements are equivalent:*

(a) $g : B \to A$ *is surjective.*

(b) $g(g^{-1}(S)) = S$ *for all* $S \subseteq A$.

(c) *There is a function* $f : A \to B$ *such that* $g \circ f = id_A$.

Remark. If you expected a forth statement here, recall that $g(S \cup T) = g(S) \cup g(T)$ is *always* true for all $S, T \subseteq B$, whether g is surjective or not. (See #32 §1.8.)

PROOF.
"$(a) \to (b)$": Suppose that $g : B \to A$ is surjective. Let $S \subseteq A$ be any subset. We wish to show that $g(g^{-1}(S)) = S$. We start with proving that $g(g^{-1}(S)) \subseteq S$: let $a \in g(g^{-1}(S))$, then $a = g(b)$ for some $b \in g^{-1}(S)$. So, $a = g(b) \in S$. (Notice that this is always true.) Conversely, let $s \in S$. Since $g : B \to A$ is surjective and $s \in A$, then $s = g(b)$ for some $b \in B$. Hence, $b \in g^{-1}(S)$, and consequently $s = g(b) \in g(g^{-1}(S))$. This proves that $S \subseteq g(g^{-1}(S))$.
"$(b) \to (c)$": Let us now assume that statement (b) holds. We construct the desired function f as follows: for every $a \in A$ we can form the set

$$G_a = \{(a, b) \in A \times B \mid a = g(b)\}$$

of all pairs whose second coordinate b maps to its first coordinate a under g. Since, by assumption, $g(g^{-1}(\{a\})) = \{a\}$ for all $a \in A$, the set $g^{-1}(\{a\})$ is not empty, guaranteeing the existence of at least one element $b \in B$ with $g(b) = a$. Consequently, none of the sets G_a is empty. Moreover, if $a_1 \neq a_2$ then $G_{a_1} \cap G_{a_2} = \emptyset$, as can be seen by looking at the first coordinates. We are now in a position to apply the axiom of choice: let f be a set which contains exactly one element from each of the sets G_a with $a \in A$. Then $f \subseteq A \times B$. In fact, since by the very choice of f no two elements of f have the same first coordinate, $f : A \to B$ is a *function*. Also, $g(f(a)) = a$ for all $a \in A$ by definition of the sets G_a.
"$(c) \to (a)$": Suppose there is a function $f : A \to B$ such that $g \circ f = id_A$. That allows us to show that $g : B \to A$ is surjective. For if $a \in A$, we can define $b = f(a) \in B$. Then $g(b) = g(f(a)) = a$. Hence, $a \in g(B)$. This proves that $A \subseteq g(B)$ so that, in fact, $g(B) = A$. □

Corollary. For any two sets A and B the following two statements are equivalent:

(i) There is an injective function $f : A \to B$.

(ii) There is a surjective function $g : B \to A$.

PROOF. The above theorems imply that being injective is equivalent with having a "left inverse" and being surjective is equivalent with having a "right inverse". To prove the corollary one only has to observe that a function with a "right inverse" is the "left inverse" of that function and vice versa. □

This allows us finally to make the following

Definition. If either one of the two equivalent statements in the above corollary holds, then we write $|A| \leq |B|$.

It is a highly non-trivial fact that for any two sets A and B one of the three relationships $|A| = |B|$, $|A| \leq |B|$, or $|B| \leq |A|$ must hold. We will not attempt to prove this fact here. What we can prove, however, is the following (also not obvious)

Theorem (Schröder-Bernstein). If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.

PROOF. Suppose there are two injective functions $f : A \to B$ and $g : B \to A$. Then their composition $h : A \to A$ given by $h = g \circ f$ is also injective (Problem #25 of Section 1.8). Consider the set $C = A \setminus g(B)$. We claim that none of the sets $h(C), h(h(C)), h(h(h(C))), \cdots$ (which are all subsets of A) shares any elements with the set C. To see this, notice that each one of the sets $h(C), h(h(C)), h(h(h(C))), \cdots$ is a subset of $h(A) = g(f(A))$ which, in turn, is a subset of $g(B) = A \setminus C$. We then form the union $D = C \cup h(C) \cup h(h(C)) \cup \cdots$. As in Problem #32 of Section 1.8 one shows that $h(D) = h(C) \cup h(h(C)) \cup h(h(h(C))) \cup \cdots$. Therefore, $h(D) = D \setminus C$ (due to the above claim). Since h is injective, then the function $k : D \to h(D)$ defined by $k(x) = h(x)$ is bijective. Hence, defining $k(x) = x$ for $x \in A \setminus D$, extends k to a bijective function $k : A \to A \setminus C$. (Verify that!) Consequently, $|A| = |A \setminus C|$. We also have $|B| = |A \setminus C|$, because $g : B \to g(B) = A \setminus C$ is a bijection. Thus, $|A| = |B|$. $\square$

Introducing the convenient notation $|A| < |B|$ for $(|A| \leq |B| \wedge |A| \neq |B|)$ we summarize some results from class:

(i) $|\mathbb{Q}| = |\mathbb{N}|$

(ii) $|\mathbb{R}| > |\mathbb{N}|$

(iii) $|\mathbb{R}| = |I|$ for every non-trivial interval I of real numbers.

(iv) $|\mathbb{R} \times \mathbb{R}| = |\mathbb{R}|$

(v) $|\mathcal{P}(A)| > |A|$ for every set A.

[**Proof:** Since $\{a\} \in \mathcal{P}(A)$ for every $a \in A$, clearly $|\mathcal{P}(A)| \geq |A|$. So, we only have to show that $|\mathcal{P}(A)| \neq |A|$. We do this by way of contradiction. Suppose that there were a surjective function $g : A \to \mathcal{P}(A)$. Then $g(a)$ is a subset of A for every $a \in A$. Consider the set $B = \{a \in A \mid a \notin g(a)\}$. Since $B \subseteq A$, i.e. $B \in \mathcal{P}(A)$, and g is surjective, then $B = g(b)$ for some $b \in A$. Now, either $b \in B$ or $b \notin B$. We will show that *neither* is possible, thus establishing the desired contradiction. If $b \in B$, then (by the very definition of the set B) $b \notin g(b) = B$; but this is impossible! On the other hand, if $b \notin B$, then (again by the definition of B) $b \in g(b) = B$; another impossibility. This completes the proof.]

(vi) $|\mathcal{P}(I\!\!N)| = |I\!\!R|$

[**Sketch of proof:** every subset $A \subseteq I\!\!N$ of the natural numbers gives rise to the binary expansion $x = .a_0a_1a_2\cdots$ of a real number $x \in [0, 1]$ by the rule $a_i = 0$ if $i \notin A$ and $a_i = 1$ if $i \in A$. Clearly, for every number in $x \in [0, 1]$ there is a set $A \subseteq I\!\!N$ whose assigned expansion equals x. The only numbers in $[0, 1]$ that have been represented more than once by this assignment are those positive numbers that have a terminating binary expansion (because such numbers have exactly two different expansions – one terminating, like .101, and one non-terminating, like $.1001111\cdots$). Since the set of all numbers in $[0, 1]$ with terminating binary expansion is countable (Problem #39 of Section 3.2), we can "fix" this assignment and make it a bijection. (How?) That will prove that $|\mathcal{P}(I\!\!N)| = |[0, 1]| = |I\!\!R|$.]