Introduction

In this short article, we will describe some basic notions on cardinality of sets. Given two sets $A$ and $B$ it is natural to ask “how big are they” in terms of the number of elements and how do their sizes compare with each other. Cardinality is the measure of the number of elements in a set. When the sets are finite the notion is simple enough and we simply “count” the number of elements. For example, we say that the cardinality of the set $A = \{2, 5, 7, 9\}$ is 4 and that the cardinality of the set $B = \{\text{my large suitcase, my small suitcase, my backpack}\}$ is 3. However, when the sets are infinite then we do not have a clear definition for cardinality in terms of “the number of elements”. We might simply assign a symbol $\infty$ to denote that the cardinality of the set is “infinite”. But then this is not sufficient to compare all infinite sets as it turns out that there are “several levels of infinity”.

In order to distinguish various levels of infinity, we attempt to redefine cardinality in terms of “bijections between sets”. It is reasonable to say that two sets are of the same size if we could associate each member in one of the sets with a corresponding member in the other set and in this way build a one-to-one correspondence between the members of the sets. From our knowledge of injective and surjective functions between sets, we have also developed the following intuition.

Let $A$ and $B$ be sets. Suppose there is an injective function, $f : A \to B$. Then this information gives us a faint picture that in some sense the set $A$ is “smaller than (or of the same size as) the set $B$.” This is because the injective function is a bijection onto its image, $\text{im } f$ which is a subset of $B$. (So, in a sense, $A$ can not have more number of elements than $B$, or equivalently $B$ has at least as many elements as $A$). In the same way, now suppose that there is a surjective function $g : A \to B$. Then this information tells us that in some sense $B$ can not have more elements than $A$, or equivalently $A$ has atleast as many elements as $B$ because, for each element in the set $B$ there is at least one corresponding element in set $A$ that maps to it under the surjection $g$.

With these intuitions we proceed with the following definitions. While we hope that these intuitions formalize in the expected way, we must also note carefully that the Axiom of Choice is needed at certain steps. The material that I am about to present is taken from [Fol99] along with [Gan09].

The Theory

**Definition 0.1.** Let $A$ and $B$ be nonempty sets. We say

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


to mean that there exists \( f : A \rightarrow B \) which is injective, bijective, or surjective respectively. We also say that

\[
\text{card}(A) < \text{card}(B), \quad \text{card}(A) > \text{card}(B)
\]
to mean there exists an injection but no bijection, or a surjection but no bijection from \( A \) to \( B \), respectively.

**Example 0.2.** \( \text{card}([2, 3]) \leq \text{card}([4, 5, 6]) \) (since, for example, \( f : [2, 3] \rightarrow [4, 5, 6] \) defined by \( f(x) = x + 2 \) is an injection), and in fact, \( \text{card}([2, 3]) < \text{card}([4, 5, 6]) \).

Similarly, \( \text{card}([21, 23, 25, 27]) \geq \text{card}([11, 13]) \) (can you think of a surjection?). And in fact, \( \text{card}([21, 23, 25, 27]) > \text{card}([11, 13]) \).

And observe that \( \text{card}([2, 3]) < \text{card}([4, 5]) \), for example.

**Example 0.3.** For a nontrivial example, let \( A \) be the set of all positive integers and let \( B \) be the set of all odd positive integers. First observe that \( \text{card}(B) \leq \text{card}(A) \) (the obvious inclusion map \( i : B \rightarrow A \) defined by \( i(x) = x \) is an injection). And also observe that in fact, \( \text{card}(A) = \text{card}(B) \) due to the bijection, \( f : A \rightarrow B \) given by \( f(n) = 2n - 1 \).

**Example 0.4.** The empty set, \( \emptyset \) is special due to the funny/silly logical arguments leading to *vacuous truths*. For example, observe that for any given set, \( A \), there is a unique function from the empty set into \( A \), which is called the *empty function*. (Please read page 9 in section 0.2 in [Gau09].) And observe that the empty function is injective vacuously. Also, the empty function is surjective if and only if the codomain \( A = \emptyset \).

On the other hand, observe that there is no function from any nonempty set to the empty set (why?). However if (hypothetically), there were a function from a nonempty set into the empty set, then it would have been surjective vacuously. Hence, we may extend the definitions given in Definition 0.1 to the empty set as follows.

**Definition 0.5.** If \( X \neq \emptyset \), then we declare \( \text{card}(\emptyset) < \text{card}(X) \) (since the empty function is injective but not bijective) and \( \text{card}(X) > \text{card}(\emptyset) \) (even though, there is no surjection (no function) from \( X \) onto \( \emptyset \)).

**Theorem 0.6.** Let \( A, B \) and \( C \) be sets (may or may not be empty). Then

(i) \( \text{card}(A) = \text{card}(A) \).

(ii) If \( \text{card}(A) = \text{card}(B) \), then \( \text{card}(B) = \text{card}(A) \).

(iii) If \( \text{card}(A) = \text{card}(B) \) and \( \text{card}(B) = \text{card}(C) \), then \( \text{card}(A) = \text{card}(C) \).

*Proof.* This is precisely Theorem 0.11 given in [Gau09]. Please refer it for the proof. \( \square \)

Observe that we do not associate any meaning to the expression \( \text{card}(A) \). However, by the definition of *finite* sets (please see [Gau09]), we have that a set \( A \) is finite if and only if \( A = \emptyset \) or \( \text{card}(A) = \text{card}([1, 2, ..., n]) \). Hence, if \( A \) is a finite set, it is convenient to interpret “\( \text{card}(A) \)” as the number of elements of \( A \). And in this case it is reasonable to write

\[
\text{card}(A) = n \text{ if and only if } \text{card}(A) = \text{card}([1, 2, ..., n])
\]

and

\[
\text{card}(A) = 0 \text{ if and only if } A = \emptyset
\]
If \( A \) is an infinite set, we do not associate any meaning to the stand-alone expression “card(\( A \))”. If we are given another set \( B \) (finite or infinite), then all that makes sense is how \( A \) compares with \( B \) in terms of the existence of an injection/surjection/bijection from \( A \) to \( B \) or vice versa.

In the rest of our discussion we assume all our sets to be nonempty (unless stated otherwise) in order to avoid special arguments for the empty set.

It is natural to ask (based on our intuition and the choice of our notation), does card(\( X \)) ≤ card(\( Y \)) imply card(\( Y \)) ≥ card(\( X \)) and conversely? (That is, does the existence of an injection from \( X \) into \( Y \) imply the existence of a surjection from \( Y \) onto \( X \) and vice versa?) Of course, we want this to be true when \( X \) and \( Y \) are finite sets in order to think of the cardinality of a set as the number of elements in it. And in fact, this is true and this can easily be proven in the case of finite sets. Also the following theorem holds true in general for all nonempty sets if we assume that the **Axiom of Choice** (stated below) holds.

**Axiom 0.7** (Axiom of Choice).

Let \( X = \{A_\lambda\}_{\lambda \in \Lambda} \) be a nonempty indexed family of nonempty sets. Then there exists a choice function \( f \) defined on \( X \). That is, there exists a function \( f: \{A_\lambda\}_{\lambda \in \Lambda} \to \bigcup_{\lambda \in \Lambda} A_\lambda \) such that for all \( \lambda \in \Lambda \), \( f(A_\lambda) \in A_\lambda \).

**Theorem 0.8.** Let \( X \) and \( Y \) be nonempty sets. Then card(\( X \)) ≤ card(\( Y \)) if and only if card(\( Y \)) ≥ card(\( X \)).

**Proof.** Recall that we proved this theorem in class. Observe that proving the converse required the use of the Axiom of Choice.

**Forward Implication:**

Let \( f: X \to Y \) be an injection. Choose \( x_0 \in X \). Then define \( g: Y \to X \) by \( g(y) = f^{-1}(y) \) if \( y \in \text{im}(f) \) and \( g(y) = x_0 \) otherwise. Then \( g \) is surjective.

**Converse Implication:**

Let \( g: Y \to X \) be a surjection. Then the family of nonempty and disjoint sets \( \{g^{-1}(\{x\})\}_{x \in X} \) is nonempty. Let \( F \) be a choice function defined on this family (such a function exists by the Axiom of Choice). Define \( f: X \to Y \) by \( f(x) = F(g^{-1}(\{x\})) \). Then \( f \) is injective.

The Axiom of Choice (or its equivalent called as Zorn’s Lemma) also implies the following theorem which we state without proof. It says that given any two sets, “their sizes are comparable” in the sense that there is always an injective function from one of them to the other.

**Theorem 0.9.** For any sets \( X \) and \( Y \), either card(\( X \)) ≤ card(\( Y \)) or card(\( Y \)) ≤ card(\( X \)).

**Proof.** Please refer to [Fol99].

The next theorem is an important theorem. It states that given any two sets, if there are injective functions from each of them to the other, then there is a bijection between them. It is to be noted carefully that the theorem can be proven without using the Axiom of Choice. We will omit the proof.
Theorem 0.10 (The Schröder-Bernstein Theorem).
If \( \text{card}(X) \leq \text{card}(Y) \) and \( \text{card}(Y) \leq \text{card}(X) \), then \( \text{card}(X) = \text{card}(Y) \).

Proof. Please refer to [Fol99].

We now state another important theorem.

Theorem 0.11. For any set \( X \), \( \text{card}(X) < \text{card}(P(X)) \), where \( P(X) \) is the power set of \( X \), that is the set of all subsets of \( X \).

Proof. Exercise for you! Recall problem 37 in Chapter 0 of Gaughan’s book ([Gau09]).

Let us now recall the definitions of a countable infinite set and that of a countable set. We say that a set \( A \) is countably infinite if and only if there is a bijection \( f : A \to \mathbb{N} \) where \( \mathbb{N} \) is the set of all natural numbers. That is, a set \( A \) is countably infinite if and only if \( \text{card}(A) = \text{card}(\mathbb{N}) \). And we say that a set, \( A \) is countable if and only if either \( A \) is finite or \( A \) is countably infinite.

Now observe that if a set \( S \) is countable, there is always an injection from \( S \) into \( \mathbb{N} \). Let us state this as a proposition in order to emphasize its importance.

Proposition 0.12. Let \( S \) be a countable set. Then there exists an injection \( f : S \to \mathbb{N} \), that is, \( \text{card}(S) \leq \text{card}(\mathbb{N}) \).

Proof. If \( S \) is countable, then either \( S \) is finite or \( S \) is countably infinite.

Suppose \( S \) is finite. If \( S = \emptyset \), then as observed in Example 0.4, the empty function \( e : \emptyset \to \mathbb{N} \) is an injection. If \( S \neq \emptyset \) then there is some \( n \in \mathbb{N} \) and a bijection \( g : S \to \{1,2,...,n\} \). Let \( i : \{1,2,...,n\} \to \mathbb{N} \) be the inclusion map defined by \( i(x) = x \). Note that \( i \) is an injection. Then the composite function \( (i \circ g) : S \to \mathbb{N} \) is injective. Thus, in this case, we have \( \text{card}(S) \leq \text{card}(\mathbb{N}) \).

Suppose \( S \) is countably infinite. Then there is a bijection \( f : S \to \mathbb{N} \), by definition. \( f \) is also injective. Thus, \( \text{card}(S) \leq \text{card}(\mathbb{N}) \) in this case as well.

It is natural to ask if the converse of the above proposition holds. In fact it does hold true, and it means that the condition of existence of an injection from a set to \( \mathbb{N} \) is not only a necessary condition but also a sufficient condition for the set to be countable. And this fact is implied by the following theorem stated in Gaughan’s book ([Gau09]). The theorem says that in some sense, countably infinite sets are the smallest infinite sets.

Theorem 0.13. Any infinite subset of \( \mathbb{N} \) is countably infinite.

Proof. Please refer to [Gau09].

Corollary 0.14. Any subset of a countable set is countable.

Proof. Exercise! Recall that we proved this in class.

Corollary 0.15. Let \( S \) be a set. Then \( S \) is countable if and only if there exists an injection \( f : S \to \mathbb{N} \), that is, \( \text{card}(S) \leq \text{card}(\mathbb{N}) \).
Proof.

Forward Implication:
This was proved in Proposition 0.12

Converse Implication:
Let \( f : S \to \mathbb{N} \) be an injection. Then \( f \) is a bijection onto its image, \( \text{im}(f) = f(S) \), which is a subset of the countable set, \( \mathbb{N} \). Then by Corollary 0.14, \( f(S) \) is countable. Thus \( S \) being equivalent to a countable set is countable. \( \square \)

Corollary 0.15 not only gives an easy criterion for showing that a set is countable, but also gives us the picture that countable sets are the smallest sets in a certain sense.

We now observe that we have a hierarchy of sets where the empty set is at the lowest level. In the next level, are the singleton sets (those equivalent to the set \( \{1\} \)). And then the doubleton sets (those equivalent to the set \( \{1, 2\} \)) are in the next level and so on. There are infinitely many such levels of finite sets and the level that comes immediately above all the (infinitely many) levels of finite sets, is that level that contains all the countably infinite sets.

And observe, that given any two countable sets, \( A \) and \( B \), there is an injection from \( A \) to \( B \) (that is, \( \text{card}(A) \leq \text{card}(B) \)) if and only if \( A \) is either at the same level or at a level lower than \( B \). Also, given any two countable sets, \( A \) and \( B \) there is an injection but no bijection from \( A \) to \( B \) (that is, \( \text{card}(A) < \text{card}(B) \)) if and only if \( A \) at a level (strictly) lower than \( B \).

This picture also extends to arbitrary sets including uncountable sets. Observe that by Theorem 0.11, there are infinitely many levels above the level containing all the countably infinite sets (\( \text{card}(\mathbb{N}) < \text{card}(\mathcal{P}(\mathbb{N})) < \text{card}(\mathcal{P}(\mathcal{P}(\mathbb{N}))) < \ldots \)).

Let us conclude our discussion with the assertions on the position of \( \mathbb{R} \), the set of all real numbers in the hierarchy of sets that we described above.

Theorem 0.16. \( \text{card}(\mathcal{P}(\mathbb{N})) = \text{card}(\mathbb{R}) \), that is, there is a bijection from \( \mathcal{P}(\mathbb{N}) \) onto \( \mathbb{R} \). (It means that the sets \( \mathcal{P}(\mathbb{N}) \) and \( \mathbb{R} \) are in the same level in the hierarchy).

Proof. (Sketch)

There are several proofs for this fact. I will give a sketch of the proof based on how we proved that \( \mathbb{R} \) is uncountable in class. Recall that the tan function (after composing with an appropriate linear function) gives a bijection from the open interval \((0, 1)\) onto \( \mathbb{R} \).

Also we showed that the set \( B_\infty \) of all binary sequences is uncountable and is in bijection with a subset of \((0, 1)\) containing all those real numbers that whose decimal expansions involve only two chosen digits (both different from 9). Let us call this special subset, \( U \). With some work, one can show that this special subset, \( U \) of \((0, 1)\) is in fact in bijection with \((0, 1)\).

Finally, it is not hard to see that the set \( B_\infty \) of all binary sequences is in bijection with \( \mathcal{P}(\mathbb{N}) \). (Given any subset, \( S \) of \( \mathbb{N} \) there is a unique binary sequence \( \alpha = (\alpha_n)_{n \in \mathbb{N}} \) defined by \( \alpha_n = 1 \) if \( n \in S \) and \( \alpha_n = 0 \) otherwise. It is easy to check that the function from \( \mathcal{P}(\mathbb{N}) \to B_\infty \) defined in this way is a bijection).

In short, we have \( \mathbb{R} \sim (0, 1) \sim U \sim B_\infty \sim \mathcal{P}(\mathbb{N}) \) where \( A \sim B \) means that there is a bijection between the sets \( A \) and \( B \). We omit the proof that \((0, 1) \sim U \).

For an alternate proof that uses Theorem 0.10 please see [Fol99]. \( \square \)

Thus we see that \( \mathbb{R} \) is at a level (strictly) higher than the level of countable sets. Clearly, we have
Corollary 0.17. If \( \text{card}(X) \geq \text{card}(\mathbb{R}) \), then \( X \) is uncountable. (In particular, \( \mathbb{R} \) is uncountable).

**Proof.** Suppose \( X \) is countable and let \( g : X \to \mathbb{R} \) be a surjection. Then by Theorem \( 0.16 \), there is a surjection \( h : X \to P(\mathbb{N}) \). Also, by Corollary \( 0.15 \) there is an injection \( i : X \to \mathbb{N} \). Then by Theorem \( 0.8 \), there is a surjection \( j : \mathbb{N} \to X \). (Note that we do not require the Axiom of Choice for using this part of Theorem \( 0.8 \)). Then the composite, \( (h \circ j) : \mathbb{N} \to P(\mathbb{N}) \) is a surjection contradicting Theorem \( 0.11 \). Hence, \( X \) must be uncountable.

The converse of the above corollary is the famous **Continuum Hypothesis**. This would mean that there are no levels between the level of countably infinite sets and the level containing the sets such as \( P(\mathbb{N}) \) and \( \mathbb{R} \).

**Conjecture 0.18.** (Continuum Hypothesis)
If \( X \) is uncountable, then \( \text{card}(X) \geq \text{card}(\mathbb{R}) \).

**References**
