L10

Finite and infinite sets

One major reason for the development of set theory is to deal with issues involving infinite sets. Whenever one talks about infinite objects, some apparent paradoxes arise almost immediately, and it is necessary to make sure that these do not lead to serious problems with the reliability of mathematical material. This is particularly true when one discusses the standard number systems, all of which are infinite. For example, the introduction of zero and negative numbers quickly leads to questions about the logical soundness of this concept, and their use was not fully accepted until the late 17^{th} century, about 1000 years after the ideas were first formulated. Similarly, the basic idea of expressing real numbers using infinite decimal expansions began to develop some time in the 11^{th} or 12^{th} century, and by the end of the 16^{th} century it had become fairly widely accepted, but some form of infinite set theory was needed to make everything logically rigorous and the latter was not completed until later in the 19^{th} century. Especially for the real numbers, there was the following key question: *Even though infinite decimals are relatively easy to understand and have turned out to be highly reliable for practical purposes, are there some hidden flaws which can lead to incorrect or even absurd conclusions*?

The answer is mixed. Monumental work of K. Gödel, beginning around 1930, showed that one can never be totally sure that a theory of infinite sets is logically consistent. This may seem very discouraging, but there are two important facts which are considerably more positive:

1. Further results of Gödel and work of G. Peano show that if there is a problem with axioms for a workable theory of infinite sets (which contains the natural numbers $\mathbb{N}$), then there is a problem with a very short list of axioms which characterize $\mathbb{N}$. Furthermore, a workable theory containing an object like $\mathbb{N}$ will always contain other standard number systems, including the real numbers (this is due to R. Dedekind and G. Cantor).

2. Mathematics has made dramatic advances on many fronts since 1930, but none of this work on "genuine" mathematical problems has indicated that there are any inconsistencies. Statements like this are of course subject to future discoveries; there are some speculations about where difficulties may be found, but no such problems have surfaced to date.

For our purposes, one of the crucial properties of $\mathbb{N}$ is given by the following *Well – Ordering Property*: ***Every nonempty subset of*** $\mathbb{N}$ ***has a least element.***

Here is a simple plausibility argument for understanding this property: Suppose that A is a nonempty subset, and let $n \in A$. Since A has no least element, there must be some $x_1 \in A$ so that $x_1 < n$. Similarly, there must be an $x_2 \in A$ so that $x_2 < x_1$, and so forth until we have a sequence of elements in A of the form

$$x_{n+1} < x_n < \ldots < x_1 < n.$$

It follows that $x_k \leq n - k$ for all k, and therefore we also have $x_{n+1} \leq -1$. Since -1 does not belong to $\mathbb{N}$, we have a contradiction. The source of the contradiction is our assumption that A has no least element, and therefore this must be false, meaning that A indeed has a minimal element.

Finite sets

Before going into detail about infinite sets, it is instructive to look at some properties of finite sets. The first issue is to define this concept formally.

Definition. A set A is ***finite*** if there is a $1-1$ mapping $f : A \to \{1, \ldots, n\}$ for some $n \in \mathbb{N}_+$ (the positive integers).

Given a finite set A, consider the nonempty family $\mathcal{F}$ of all $1-1$ mappings $f : A \to \{1, \ldots, k\}$ where $k \in \mathbb{N}_+$. The set of all k which can be realized in this way is nonempty, and if k can be realized then so can $k+1$; we need only compose f with the inclusion of $\{1, \ldots, k\}$ in $\{1, \ldots, k, k+1\}$. In any case, there is a least value k_0 which can be realized. We shall say that k_0 is the ***cardinality*** or ***cardinal number*** of X and write $|X| = k_0$.

Proposition. *If B is a proper subset of the finite set A, then $|B| < |A|$.*

Corollary. *There are no $1-1$ onto maps from A to a proper subset of itself.* ■

PROOF OF THE PROPOSITION. If there were such a $1-1$ correspondence then we would have A <-> B with $|B| < k_0$, so that $|A| \leq k_0 - 1$. Since k_0 is the least value of n such that there is a $1-1$ mapping from A to $\{1, \ldots, n\}$, we have a contradiction. Therefore no such $1-1$ map from A to B can exist. ■

Proposition. *If $|A| = n$ and $f : A \to \{1, \ldots, n\}$ is $1-1$, then f is onto.*

PROOF. Suppose that f is not onto and $j \notin f[A]$. Define a new $1-1$ function g as follows: $g(a) = f(a)$ if $g(a) < j$, and $g(a) = f(a)-1$ if $g(a) > j$. This is still a $1-1$ function, but its codomain is $\{1, \ldots, n-1\}$ and hence $|A| \leq n-1$, which contradicts our assumption. The source of the contradiction is our assumption that f is not onto, so that must be false and hence f must be onto.■

We can express all this in a unified form as follows:

Theorem. *Let A and B be finite sets with $|A| = |B| = n$. Then a $1-1$ function from A to B is onto, and an onto function from A to B is $1-1$.*

PROOF. The first conclusion is given by the immediately preceding result, and the second may be proved as follows: If f were not onto, then $f[A]$ would be a proper subset of B and f would define a $1-1$ correspondence from A to the proper subset $f[A] \subset B$. This leads to the chain of equalities and inequalities $|B| = |A| = |f[A]| < |B|$. The resulting inequality $|B| < |B|$ yields a contradiction, which arises from the assumption that f was not onto. It follows that f is onto.■

Counting elements in finite sets

There are a few important principles for counting elements in a standard construction on finite sets.

Theorem. *Let A and B be finite sets.*

- **(a)** *The cardinalities of $A, B, A \cup B$ and $A \cap B$ satisfy the formula $|A \cup B| = |A| + |B| - |A \cap B|$.*
- **(b)** *The cardinality of the Cartesian product $A \times B$ is equal to $|A| \cdot |B|$.*
- **(c)** *If A and B are nonempty, then the cardinality of the set of functions $\mathbf{F}(A, B)$ is equal to $|B|^{|A|}$.*

The third formula explains the reason for Cunningham's use of ${}^{A}B$ to denote $\mathbf{F}(A, B)$.

PROOF. **(a)** Consider the sequence of length $m + n$ elements $a_1, \ldots, a_n, b_1, \ldots, b_m$. The only repetitions in this sequence occur when there are j and k such that $a_j = b_k$ and there is one such repetition for every element of $A \cap B$. If we eliminate these duplications we obtain a sequence in which every element of $A \cup B$ occurs exactly once, and the number of terms in this sequence, which is $|A \cup B|$, is also given by $|A| + |B| - |A \cap B|$.■

Before proceeding, we note one consequence of this: *If* $C_1, \ldots, C_n$ *are pairwise disjoint sets and* C *is their union, then* $|C| = |C_1| + \ldots + |C_n|$. This can be proved recursively: By the preceding paragraph we have $|C_1 \cup C_2| = |C_1| + |C_2|$. The distributive law for sets implies that $C_1 \cup C_2$ and C_3 are disjoint and therefore we have $|C_1 \cup C_2 \cup C_3| = |C_1 \cup C_2| + |C_3| = |C_1| + |C_2| + |C_3|$, and similarly for unions of 4 or more subsets satisfying the given condition.

(b) For each $a \in A$ there is a $1-1$ onto map $f_a: B \to \{a\} \times B$ given by $f_a(b) = (a, b)$; the inverse mapping sends (a, b) back to b. The sets $f_a[B]$ are a pairwise disjoint family of subsets whose union is all of $A \times B$, so by the preceding paragraph we know that $|A \times B| = |B| + \ldots + |B|$ (there are $|A|$ summands, one for each element of A). The formula follows because the right hand side Is merely $|A| \cdot |B|$.■

(c) We shall use the following product principle for counting: *Suppose we are given* n *choices* $C_1, \ldots, C_n$ *such that the number* m_k *of alternatives at each step* k *does not depend upon the previous choices. Then the total number of choices is equal to the product* $m_1 \cdot \ldots \cdot m_n$.

In our situation we want to look at the number of ways to define a function f:A\to B. Write the elements of A as $a_1, \ldots, a_m$. For each j there are exactly $|B|$ ways to define $f(a_j)$, one for each element of B. Therefore the total number of ways for defining a function is equal to $|B| \cdot \ldots \cdot |B|$ (there are $|A|$ factors, one for each element of A). By the definition of positive integer exponents, this number is equal to $|B|^{|A|}$.■

Here are two more counting rules that are very useful.

Proposition (Counting subsets of a set). *If* A *is a finite set and* $|A| = n$, *then the set of all subsets* $\mathcal{P}(A)$ *is also finite and* $|\mathcal{P}(A)| = 2^{|A|}$.

PROOF. The main idea is to find a $1-1$ correspondence from $\mathcal{P}(A)$ to the set of all functions from A to $\{0, 1\}$ and then to apply the formula in **(c)** from the previous; this construction also works for infinite sets.

Given $B \subset A$, define its ***characteristic function*** $\chi_B: A \to \{0, 1\}$ by $\chi_B(a) = 1$ if $a \in B$ and by $\chi_B(a) = 0$ if $a \notin B$. By construction the inverse image of $\{1\}$ is equal to B, and therefore $\chi_B = \chi_C$ implies $B = C$. Therefore the characteristic function map $\mathcal{P}(A) \to \mathbf{F}(A, \{0, 1\})$ is $1-1$. But this mapping is also onto, for if f is any function from A to $\{0, 1\}$ then $f = \chi_B$ where B is the inverse image of $\{1\}$. Finally, apply formula **(c)** to conclude that $|\mathcal{P}(A)| = 2^{|A|}$.■

Proposition (Pigeonhole Principle). *Suppose that* A *and* B *are finite sets such that* $|A| > |B|$ *and* $f: A \to B$ *is a function. Then* f *is not* $1-1$*; in other words, there exist distinct elements* $x, y \in A$ *such that* $f(x) = f(y)$.

PROOF. We shall prove the contrapositive: If no such pair $x, y \in A$ exists then we must have $|f[A]| \leq |B|$. If no such pair exists, then f is $1-1$. If we now let $g: B \to \{1, \ldots, n\}$ be $1-1$ and onto function, then g\tinycirc f is also $1-1$, which implies $|A| \leq n = |B|$.■

Countable sets

The Pigeonhole Principle implies that the set $\mathbb{N}$ of nonnegatie integers is not a finite set; in fact, for each n we know that a function $\{1, \ldots, n+1\}$ to $\{1, \ldots, n\}$ is never $1-1$. We shall eventually show that in some sense $\mathbb{N}$ is the smallest set which is not finite. This will be part of a more general theory of cardinal numbers for infinite sets. Our development of this material will use the following axiom for set theory (see also Subsection 3.3.4 of Cunningham):

AXIOM OF CHOICE. *Let* A *be a nonempty set and let* $\mathcal{P}_+(A)$ *denote the set of nonempty subsets of* A. *Then there is a* ***choice function*** $c: \mathcal{P}_+(A) \to A$ *such that* $c(B) \in B$ *for all nonempty subsets* $B \subset A$.

In other words, *given a family of nonempty subsets, one can choose a representative element for each one in a coherent manner*. Initially some mathematicians questioned the appropriateness of including such an axiom because it is a pure statement about existence, giving no specific method for finding such a choice function. However, it is now generally accepted, although sometimes reluctantly. One major reason for its general acceptance is a theorem of Gödel which states that if there is a logical inconsistency in set theory with the axiom of choice, then there is already a logical inconsistency if one does not assume it. Later work of P. Cohen showed that there are models for the Zermelo – Fraenkel axioms of set theory (as stated in Chapter 1 of Cunningham) for which the Axiom of Choice is true and other models for which it is false. Since the axiom yields many interesting and important mathematical objects which could not be considered otherwise and by itself leads to no further logical uncertainties, most mathematicians now have few if any problems with it.

The following result illustrates the use of this axiom in mathematics.

Proposition (Cross Section Property). *Let* $f:A \to B$ *be a function which is onto. Then there is a* $1-1$ *function* $\sigma:B \to A$ (*a* ***cross section***) *such that* $f \circ \sigma(b) = b$ *for all* $b \in B$.

PROOF. We shall use the Axiom of Choice to construct a function $\sigma:B \to A$ such that $\sigma(b) \in f^{-1}[\{b\}]$ for each $b \in B$. Note that each level subset $f^{-1}[\{b\}]$ is nonempty since f is onto. By construction we have $f(\sigma(b)) = b$. Finally, we need to verify that σ is $1-1$. This follows because $\sigma(b) = \sigma(b')$ implies $b = f(\sigma(b)) = f(\sigma(b')) = b'$.■

This discussion will be continued in the next lecture.