1.6 Epilogue

The relationship of having the same cardinality is an *equivalence relation* (see Exercise 1.4.9), meaning, roughly, that all of the sets in the universe can be organized into disjoint groups according to their size. Two sets appear in the same group, or *equivalence class*, if and only if they have the same cardinality. Thus, $\mathbf{N}$, $\mathbf{Z}$, and $\mathbf{Q}$ are grouped together in one class with all of the other countable sets, whereas $\mathbf{R}$ is in another class that includes the interval $(0, 1)$ among other uncountable sets. One implication of Cantor's Theorem is that $P(\mathbf{R})$—the set of all subsets of $\mathbf{R}$—is in a different class from $\mathbf{R}$, and there is no reason to stop here. The set of subsets of $P(\mathbf{R})$—namely $P(P(\mathbf{R}))$—is in yet another class, and this process continues indefinitely.

Having divided the universe of sets into disjoint groups, it would be convenient to attach a "number" to each collection which could be used the way natural numbers are used to refer to the sizes of finite sets. Given a set X, there exists something called the *cardinal number* of X, denoted $\operatorname{card} X$, which behaves very much in this fashion. For instance, two sets X and Y satisfy $\operatorname{card} X = \operatorname{card} Y$ if and only if $X \sim Y$. (Rigorously defining $\operatorname{card} X$ requires some significant set theory. One way this is done is to define $\operatorname{card} X$ to be a very particular set that can always be uniquely found in the same equivalence class as X.)

Looking back at Cantor's Theorem, we get the strong sense that there is an *order* on the sizes of infinite sets that should be reflected in our new cardinal number system. Specifically, if it is possible to map a set X *into* Y in a 1–1 fashion, then we want $\operatorname{card} X \leq \operatorname{card} Y$. Writing the strict inequality $\operatorname{card} X < \operatorname{card} Y$ should indicate that it is possible to map X into Y but that it is impossible to show $X \sim Y$. Restated in this notation, Cantor's Theorem states that for every set A, $\operatorname{card} A < \operatorname{card} P(A)$.

There are some significant details to work out. A kind of metaphysical problem arises when we realize that an implication of Cantor's Theorem is that there can be no "largest" set. A declaration such as, "Let U be the set of all possible things," is paradoxical because we immediately get that $\operatorname{card} U < \operatorname{card} P(U)$ and thus the set U does not contain everything it was advertised to hold. Issues such as this one are ultimately resolved by imposing some restrictions on what can qualify as a set. As set theory was formalized, the axioms had to be crafted so that objects such as U are simply not allowed. A more down-to-earth problem in need of attention is demonstrating that our definition of "$\leq$" between cardinal numbers really is an ordering. This involves showing that cardinal numbers possess a property analogous to real numbers, which states that if $\operatorname{card} X \leq \operatorname{card} Y$ and $\operatorname{card} Y \leq \operatorname{card} X$, then $\operatorname{card} X = \operatorname{card} Y$. In the end, this boils down to proving that if there exists $f : X \to Y$ that is 1–1, and if there exists $g : Y \to X$ that is 1–1, then it is possible to find a function $h : X \to Y$ that is both 1–1 and onto. A proof of this fact eluded Cantor but was eventually supplied independently by Ernst Schröder (in 1896) and Felix Bernstein (in 1898). An argument for the Schröder–Bernstein Theorem is outlined in Exercise 1.4.13.

There was another deep problem stemming from the budding theory of cardinal numbers that occupied Cantor and which was not resolved during his lifetime. Because of the importance of countable sets, the symbol $\aleph_0$ ("aleph naught") is frequently used for card $\mathbf{N}$. The subscript "0" is appropriate when we remember that countable sets are the smallest type of infinite set. In terms of cardinal numbers, if card $X < \aleph_0$, then X is finite. Thus, $\aleph_0$ is the smallest infinite cardinal number. The cardinality of $\mathbf{R}$ is also significant enough to deserve the special designation $\boldsymbol{c} = \operatorname{card}\mathbf{R} = \operatorname{card}(0,1)$. The content of Theorems 1.4.11 and 1.5.1 is that $\aleph_0 < \boldsymbol{c}$. The question that plagued Cantor was whether there were any cardinal numbers strictly in between these two. Put another way, does there exist a set $A \subseteq \mathbf{R}$ with $\operatorname{card}\mathbf{N} < \operatorname{card}A < \operatorname{card}\mathbf{R}$? Cantor was of the opinion that no such set existed. In the ordering of cardinal numbers, he conjectured, $\boldsymbol{c}$ was the immediate successor of $\aleph_0$.

Cantor's "continuum hypothesis," as it came to be called, was one of the most famous mathematical challenges of the past century. Its unexpected resolution came in two parts. In 1940, the German logician and mathematician Kurt Gödel demonstrated that, using only the agreed-upon set of axioms of set theory, there was no way to disprove the continuum hypothesis. In 1963, Paul Cohen successfully showed that, under the same rules, it was also impossible to prove this conjecture. Taken together, what these two discoveries imply is that the continuum hypothesis is undecidable. It can be accepted or rejected as a statement about the nature of infinite sets, and in neither case will any logical contradictions arise.

The mention of Kurt Gödel brings to mind a final comment about the significance of Cantor's work. Gödel is best known for his "Incompleteness Theorems," which pertain to the strength of axiomatic systems in general. What Gödel showed was that any consistent axiomatic system created to study arithmetic was necessarily destined to be "incomplete" in the sense that there would always be true statements that the system of axioms would be too weak to prove. At the heart of Gödel's very complicated proof is a type of manipulation closely related to what is happening in the proofs of Theorems 1.5.1 and 1.5.2. Variations of Cantor's proof methods can also be found in the limitative results of computer science. The "halting problem" asks, loosely, whether some general algorithm exists that can look at every program and decide if that program eventually terminates. The proof that no such algorithm exists uses a diagonalization-type construction at the core of the argument. The main point to make is that not only are the implications of Cantor's theorems profound but the argumentative techniques are as well. As a more immediate example of this phenomenon, the diagonalization method is used again in Chapter 6—in a constructive way—as a crucial step in the proof of the Arzela–Ascoli Theorem.