4

Infinity in Greek Mathematics

4.1 Fear of Infinity

Reasoning about infinity is one of the characteristic features of mathematics as well as its main source of conflict. We saw, in Chapter 1, the conflict that arose from the discovery of irrationals, and in this chapter we shall see that the Greeks' rejection of irrational numbers was just part of a general rejection of infinite processes. In fact, until the late nineteenth century most mathematicians were reluctant to accept infinity as more than "potential." The infinitude of a process, collection, or magnitude was understood as the possibility of its indefinite continuation, and no more—certainly not the possibility of eventual completion. For example, the natural numbers $1, 2, 3, \ldots$, can be accepted as a potential infinity—generated from 1 by the process of adding 1—without accepting that there is a completed totality $\{1, 2, 3, \ldots\}$. The same applies to any sequence $x_1, x_2, x_3, \ldots$ (of rational numbers, say), where x_{n+1} is obtained from x_n by a definite rule.

And yet a beguiling possibility arises when x_n tends to a limit x. If x is something we already accept—for geometric reasons, say—then it is very tempting to view x as somehow the "completion" of the sequence $x_1, x_2, x_3, \ldots$. It seems that the Greeks were afraid to draw such conclusions. According to tradition, they were frightened off by the paradoxes of Zeno, around 450 BCE.

We know of Zeno's arguments only through Aristotle, who quotes them in his *Physics* in order to refute them, and it is not clear what Zeno himself wished to achieve. Was there, for example, a tendency toward speculation about infinity that he disapproved of? His arguments are so extreme they

J. Stillwell, *Mathematics and Its History*

could almost be parodies of loose arguments about infinity he heard among his contemporaries. Consider his first paradox, the *dichotomy*:

> There is no motion because that which is moved must arrive at the middle (of its course) before it arrives at the end.
>
> (Aristotle, *Physics*, Book VI, Ch. 9)

The full argument presumably is that before getting anywhere one must first get half-way, and before that a quarter of the way, and before that one eighth of the way, ad infinitum. The completion of this infinite sequence of steps no longer seems impossible to most mathematicians, since it represents nothing more than an infinite set of points within a finite interval. It must have frightened the Greeks though, because in all their proofs they were very careful to avoid completed infinities and limits.

The first mathematical processes we would recognize as infinite were probably devised by the Pythagoreans, for example, the recurrence relations

$$\begin{aligned} x_{n+1} &= x_n + 2y_n, \\ y_{n+1} &= x_n + y_n \end{aligned}$$

for generating integer solutions of the equations $x^2 - 2y^2 = \pm 1$. We saw in Section 3.4 why it is likely that these relations arose from an attempt to understand $\sqrt{2}$, and it is easy for us to see that $x_n/y_n \to \sqrt{2}$ as $n \to \infty$.

However, it is unlikely that the Pythagoreans would have viewed $\sqrt{2}$ as a "limit" or seen the sequence as a meaningful object at all. The most we can say is that, by stating a recurrence, the Pythagoreans *implied* a sequence with limit $\sqrt{2}$, but only a much later generation of mathematicians could accept the infinite sequence as such and appreciate its importance in defining the limit.

In a problem where we would find it natural to reach a solution α by a limiting process, the Greeks would instead eliminate any solution *but* α. They would show that any number $<\alpha$ was too small and any number $>\alpha$ was too large to be the solution. In the following sections we shall study some examples of this style of proof and see how it ultimately bore fruit in the foundations of mathematics. As a method of finding solutions to problems, however, it was sterile: how does one guess the number α in the first place? When mathematicians returned to problems of finding limits in the seventeenth century, they found no use for the rigorous methods of the

Greeks. The dubious seventeenth-century methods of infinitesimals were criticized by the Zeno of the time, Bishop Berkeley, but little was done to meet his objections until much later, since infinitesimals did not seem to lead to incorrect results. It was Dedekind, Weierstrass, and others in the nineteenth century who eventually restored Greek standards of rigor.

The story of rigor lost and rigor regained took an amazing turn when a previously unknown manuscript of Archimedes, *The Method*, was discovered in 1906. In it he reveals that his deepest results were found using dubious infinitary arguments, and only later proved rigorously. Because, as he says, "It is of course easier to supply the proof when we have previously acquired some knowledge of the questions by the method, than it is to find it without any previous knowledge."

The importance of this statement goes beyond its revelation that infinity can be used to discover results that are not initially accessible to logic. Archimedes was probably the first mathematician candid enough to explain that there is a difference between the way theorems are discovered and the way they are proved.

4.2 Eudoxus' Theory of Proportions

The theory of proportions is credited to Eudoxus (around 400–350 BCE) and is expounded in Book V of Euclid's *Elements*. The purpose of the theory is to enable lengths (and other geometric quantities) to be treated as precisely as numbers, while only admitting the use of rational numbers. We saw the motivation for this in Section 1.5: the Greeks could not accept irrational numbers, but they accepted irrational geometric quantities such as the diagonal of the unit square. To simplify the exposition of the theory, let us call lengths *rational* if they are rational multiples of a fixed length.

Eudoxus' idea was to say that a length λ is determined by those rational lengths less than it and those greater than it. To be precise, he says $\lambda_1 = \lambda_2$ if any rational length $<\lambda_1$ is also $<\lambda_2$, and vice versa. Likewise $\lambda_1 < \lambda_2$ if there is a rational length $>\lambda_1$ but $<\lambda_2$. This definition uses the rationals to give an infinitely sharp notion of length while avoiding any overt use of infinity. Of course the infinite set of rational lengths $<\lambda$ is present in spirit, but Eudoxus avoids mentioning it by speaking of an arbitrary rational length $<\lambda$.

The theory of proportions was so successful that it delayed the development of a theory of real numbers for 2000 years. This was ironic, because the theory of proportions can be used to define irrational numbers just as well as lengths. It was understandable though, because the common irrational lengths, such as the diagonal of the unit square, arise from constructions that are intuitively clear and finite from the geometric point of view. Any *arithmetic* approach to $\sqrt{2}$, whether by sequences, decimals, or continued fractions, is infinite and therefore less intuitive. Until the nineteenth century this seemed a good reason for considering geometry to be a better foundation for mathematics than arithmetic. Then the problems of geometry came to a head, and mathematicians began to fear geometric intuition as much as they had previously feared infinity. There was a purge of geometric reasoning from the textbooks and industrious reconstruction of mathematics on the basis of numbers and sets of numbers. Set theory is discussed further in Chapter 23. Suffice to say, for the moment, that set theory depends on the acceptance of completed infinities.

The beauty of the theory of proportion was its adaptability to this new climate. Instead of rational lengths, take rational numbers. Instead of comparing existing irrational lengths by means of rational lengths, construct irrational numbers from scratch using sets of rationals! The length $\sqrt{2}$ is determined by the two sets of positive rationals

$$L_{\sqrt{2}} = \{r : r^2 < 2\}, \qquad U_{\sqrt{2}} = \{r : r^2 > 2\}.$$

Dedekind (1872) decided to let $\sqrt{2}$ *be* this pair of sets! In general, let any partition of the positive rationals into sets L, U such that any member of L is less than any member of U *be* a positive real number. This idea, now known as a *Dedekind cut*, is more than just a twist of Eudoxus; it gives a complete and uniform construction of all real numbers, or points on the line, using just the *discrete*, finally resolving the fundamental conflict in Greek mathematics. Dedekind was understandably pleased with his achievement. He wrote

> The statement is so frequently made that the differential calculus deals with continuous magnitude, and yet an explanation of this continuity is nowhere given.... It then only remained to discover its true origin in the elements of arithmetic and thus at the same time secure a real definition of the essence of continuity. I succeeded Nov. 24 1858.
>
> [Dedekind (1872), p. 2]