4 Spirals and Seashores

Sir Isaac Newton's relation to algebra was peculiarly ambivalent. As a young man, he studied Descartes's works carefully. In maturity, he so loathed Descartes that he sometimes seems to have been unwilling to utter or write Descartes's name, as if it would defile him. The reasons for this are not totally clear. Newton despised Descartes's philosophy as cloaked atheism, but there were also mathematical reasons for his displeasure. Newton, like Viète, styled himself a devotee and restorer of antiquity, whereas Descartes, he felt, had betrayed the profundity of Greek geometry for the questionable advantages of analytic geometry.

To demonstrate his claim, in his *Principia* Newton offers a purely geometric resolution of the four-line locus problem Descartes so prided himself on, noting caustically that this gives "not an [analytical] computation but a geometrical synthesis, such as the ancients required, of the classical problem of four lines, which was begun by Euclid and carried on by Apollonius." In the rest of the *Principia*, Newton also tends to avoid analytical algebra, preferring to state his propositions in the manner of Euclid. However, this appearance is deceptive, for in fact Newton does use algebraic expressions as well as a new tool of analytical mathematics that he himself has created: the calculus. Though he phrases it in terms of geometry, Newton's calculus goes beyond anything known to the ancients, even Archimedes, whom Newton considered his precursor.

Thus, though he tends to prefer geometry, Newton is steeped in algebra. As a young professor, he lectures on the topic (during 1672–1683), though his lectures on *Universal Arithmetic* are not published until 1707. Newton also makes a number of discoveries that will bear on the problem of the quintic. He was able to make simple connections between the coefficients of an equation and its roots, known as "Newton's identities." These generalize "Girard's identities": the coefficient of the next-lowest degree term (of x^4, for a quintic equation) is equal to the negative sum of all the roots (see box 4.1). Likewise, all the other coefficients are equal successively to the sum of all products of the roots taken two at a time, then three at a time, until the final, constant term is equal to the negative of the product of all the roots. Later on, it will become extremely important to note that all of the roots appear in a *symmetric* manner in each of these products and sums. That is, every root appears in exactly the same way as every other root, so that if one were to exchange two roots, the product of all of them would be unchanged, as would the products of them taken two at a time, and so on. The importance of these rules is that they let us see direct relations between coefficients and roots, without knowing the value of the roots. Newton is also able to obtain upper and lower bounds for the roots, that is, to show how large (or small) they could possibly be. Using these tools, we can look at any equation and determine the range within which its roots lie, as well as whether they are negative or positive in value.

Newton's deepest insight, however, remained hidden for many years, buried in a passage in his *Principia* that was

Box 4.1
Girard's and Newton's identities

Consider a cubic equation with the roots x_1, x_2, x_3. The equation can be written $(x - x_1)(x - x_2)(x - x_3) = 0$. Multiplying it out, we get $x^3 - (x_1 + x_2 + x_3)x^2 + (x_1x_2 + x_1x_3 + x_2x_3)x - x_1x_2x_3 = 0$. Note that the coefficient of the x^2 term is the negative sum of all three roots, $-(x_1 + x_2 + x_3)$, while the coefficient of the x term is the symmetric product of all the roots, taken two at a time: $(x_1x_2 + x_1x_3 + x_2x_3)$. Finally, the constant term of the equation is the negative product of all three roots: $-x_1x_2x_3$. We can apply the same reasoning to an equation of any degree, so that the coefficient of the next-to-highest power of the unknown must be the negative sum of all the roots, the next coefficient must be the symmetric sum of all the roots taken two at a time, and so on. We will refer to these simple relations as "Girard's identities," which Newton greatly generalized in expressions he derived for the sum of the square of all the roots, or the sum of their nth powers, "Newton's identities."

not much noticed until the twentieth century. In lemma 28 of Book I, in an investigation of the curved paths that bodies can take (as in the orbits of planets), he shows that "No oval figure exists whose area, cut off by straight lines at will, can in general be found by means of equations finite in the number of their terms and dimensions." That is, if we draw lines across an oval, the area between those lines cannot be expressed by a finite algebraic equation. Though he does not define the word explicitly, by "oval" Newton seems to mean any closed curve that does not cross itself (it is "simple," in the language of modern mathematicians) and is infinitely smooth (it always has finite curvature, never is "flat"). The simplest such curve is a circle, and it had long been suspected that the area of a circle is irrational with respect to its radius. But what Newton surmises goes far beyond the irrationality of π, the name later given (by Euler) to the value of a circle's perimeter divided by its diameter, and beyond the irrationality even of π^2. Newton's argument indicates that the area of a circle is not given by *any* algebraic equation, however high its degree, and thus that area (and hence π also) cannot be expressed in terms of *any* finite number of square roots, cube roots, fifth roots, and so on.

To use the term that Euler later introduced, the area of a circle is *transcendental*, meaning it cannot be expressed as the root of any equation of finite degree whose coefficients are rational numbers. At one stroke, Newton indicates that such magnitudes exist (because circles exist, and have areas), and also that there are infinitely many of them, since his proof is not restricted to circles but holds for any "oval" curve. His proof is a miracle of simplicity and power, for which he does not even bother to draw a picture or write down a line of algebra. It follows from a single brilliant contrivance. Inside the oval, pick any point whatever; let us call it the pole, P. Now let a straight line come out from that pole and rotate around it at uniform angular speed. Picture a clock hand that makes a complete revolution in one hour. Now imagine a point of light moving along that hand, starting from the pole and moving outward along the hand with speed given by the square of the distance from the pole to the point A where the hand intersects the oval (figure 4.1).

Newton has set up a way of measuring the area of the circle, for each hour the hand sweeps through that area, and the moving point keeps track of the area because it is traveling with speed proportional to the area swept out. Here Newton is implicitly using his new calculus of motion, for he knows

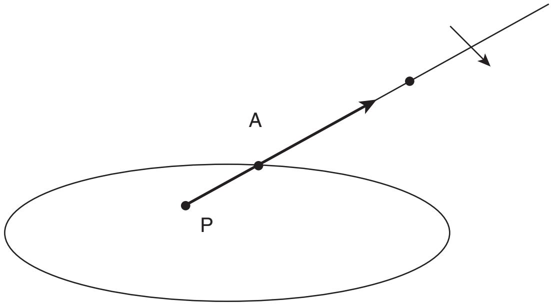

Figure 4.1
Newton's diagram for his lemma 28, which argues that no oval curve has an area expressible by a finite algebraic equation. From any point P inside the oval, draw a straight line that rotates about P at uniform angular speed.

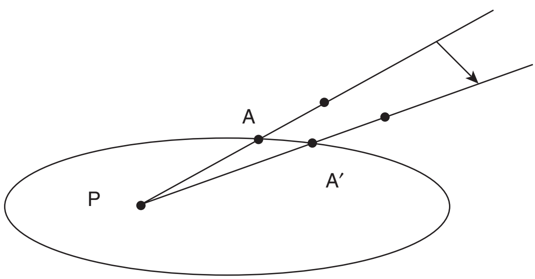

Figure 4.2
Detail of Newton's lemma 28; the speed of the moving point is proportional to the area swept out between A and A'.

that, in an infinitesimally short time, the point travels a distance from the pole equal to the area the hand has swept out in that time (figure 4.2). However, we don't need to know anything about calculus in what follows. All that matters is that since, second by second, the moving point is registering the area that the hand sweeps out, we can measure the whole area of the oval merely by waiting until an hour has elapsed and measuring the distance the moving point has traveled radially outward from the pole. For the two-dimensional problem of measuring an area, Newton has substituted a one-dimensional problem that gives the same answer: find the length traveled outward by the moving point during an hour.

The hand moves around uniformly, but the moving point speeds up and slows down in the course of each hour, in proportion to the square of the distance from the pole to the oval at any given moment. Each hour it returns to its initial speed of an hour before. If you were to watch the lighted point (or if you were to open the shutter of your camera and make a long exposure), you would see it move in a spiral, starting at the pole and making "an infinite number of gyrations," as Newton puts it (figure 4.3). Now Newton applies a *reductio ad absurdum:* Suppose that it is possible to describe this spiral (and hence also the area of the oval) by some polynomial equation with a finite number of terms, $f(x, y) = 0$. Then consider a straight line running across the spiral that we will call the x-axis, defined by $y = 0$. What can be said about the intersections of this line and the spiral? Each of them is a root of the equation $f(x, 0) = 0$. For instance, Descartes showed that all the conic sections can be described by equations of the second degree, and those curves can be cut by a straight line no more than two times. Now Newton relies on the fact that an equation of finite degree can have only a *finite* number of roots, no matter how large. But the spiral in its "infinite number of gyrations" crosses the line an infinite number of times. Thus, there should be an infinite number of intersection points, corresponding to an infinite number of roots of the equation. This contradicts our hypothesis that the equation has finite degree, and hence Newton's conclusion follows: there is no such equation that gives the area of the oval.

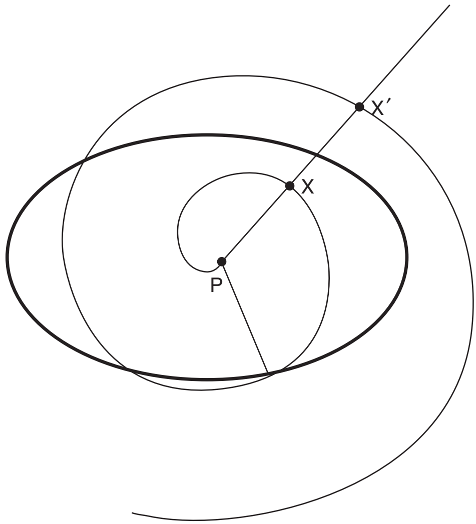

Figure 4.3
In Newton's lemma 28, the track of the moving point forms a spiral, composed of the motion of the point along the line and the uniform rotation of the line itself about P. Newton determines the area of the curve by comparing the distance from the pole P to the point X after one full revolution, which sweeps out the full area of the oval. The moving point travels an equal distance $XX' = PX$ during the next sweep, and so forth.

This brilliant argument indicates that all simple closed curves, such as the circle or the ellipse, have areas that cannot be described by finite algebraic equations. The argument seemed so simple that it made Newton's contemporaries suspicious. Daniel Bernoulli and Leibniz tried to state counterexamples, but these each involved a curve that intersects itself (for example, the lemniscate, a figure-8 on its side, ∞) or that is not closed (for example, a parabola). Later mathematicians demanded greater rigor than Newton's beautifully simple arguments (how can we *prove* that the spiral must have infinitely many intersections with the line?), but the basic thrust of his insight was sustained. There are infinitely many magnitudes that are more irrational than any radical, in the sense that no finite root is commensurable with them. In this sense, they are transcendental. Implicitly, Newton considers that his proof argues the priority of geometry over algebra by showing that a simple geometric figure includes quantities that defeat any finite amount of algebra. This makes his own preference for geometry (and his geometrically phrased calculus) more persuasive, for he thereby places his master theory beyond the limitations of algebra. In Newton's view, the ancients, with him as their modern champion, have defeated the upstart Descartes by subsuming algebra under the larger umbrella of geometry.

With this in mind, it is understandable that Newton may not have considered the solution of the quintic to be germane to his larger project, though in his younger days he did spend much energy in classifying cubic equations and contributing to the advance of algebra (for instance, "Newton's method" for the approximation of roots). If the real battle concerns geometric magnitudes not expressible in finite equations, why worry about the details of quintics? Ironically, succeeding generations of mathematicians would take up Newton's work in the simpler algebraic notation devised by his arch rival, Leibniz. His insights about transcendental magnitudes would be rediscovered centuries later. We will return to them later in light of the unfolding story of the quintic.

Despite Newton, the power and beauty of the algebraic notation ensured that interest in basic questions of algebra remained high. New hope for the solution of equations beyond the quartic was offered by the work of a Saxon nobleman, Count Ehrenfried Walter von Tschirnhaus. Tschirnhaus had varied interests. He served in the Dutch army and spent