

3.2.5 Given that the divisors of $2^{n-1}p$ are those just listed, show that $2^{n-1}p$ is perfect when $p = 2^n - 1$ is prime.

3.3 The Euclidean Algorithm

This algorithm is named after Euclid because its earliest known appearance is in Book VII of the *Elements*. However, in the opinion of many historians [for example, Heath (1921), p. 399] the algorithm and some of its consequences were probably known earlier. At the very least, Euclid deserves credit for a masterly presentation of the fundamentals of number theory, based on this algorithm.

The Euclidean algorithm is used to find the greatest common divisor (gcd) of two positive integers a, b . The first step is to construct the pair (a_1, b_1) where

$$\begin{aligned} a_1 &= \max(a, b) - \min(a, b), \\ b_1 &= \min(a, b), \end{aligned}$$

and then one simply repeats this operation of subtracting the smaller number from the larger. That is, if the pair constructed at step i is (a_i, b_i) , then the pair constructed at step $i + 1$ is

$$\begin{aligned} a_{i+1} &= \max(a_i, b_i) - \min(a_i, b_i), \\ b_{i+1} &= \min(a_i, b_i). \end{aligned}$$

The algorithm terminates at the first stage when $a_{i+1} = b_{i+1}$, and this common value is $\gcd(a, b)$. This is because taking differences preserves any common divisors, hence when $a_{i+1} = b_{i+1}$ we have

$$\gcd(a, b) = \gcd(a_1, b_1) = \cdots = \gcd(a_{i+1}, b_{i+1}) = a_{i+1} = b_{i+1}.$$

The sheer simplicity of the algorithm makes it easy to draw some important consequences. Euclid of course did not use our notation, but nevertheless he had results close to the following.

1. If $\gcd(a, b) = 1$, then there are integers m, n such that $ma + nb = 1$.

The equations

$$\begin{aligned}
 a_1 &= \max(a, b) - \min(a, b), \\
 b_1 &= \min(a, b), \\
 &\vdots \\
 a_{i+1} &= \max(a_i, b_i) - \min(a_i, b_i), \\
 b_{i+1} &= \min(a_i, b_i)
 \end{aligned}$$

show successively that a_1, b_1 are integral linear combinations, $ma + nb$, of a and b , hence so are a_2, b_2 , hence so are a_3, b_3, \dots , and finally this is true of $a_{i+1} = b_{i+1}$. But $a_{i+1} = b_{i+1} = 1$, since $\gcd(a, b) = 1$; hence $1 = ma + nb$ for some integers m, n .

2. If p is a prime number that divides ab , then p divides a or b (the *prime divisor property*).

To see this, suppose p does *not* divide a . Then since p has no other divisors except 1, we have $\gcd(p, a) = 1$. Hence by the previous result we get integers m, n such that

$$ma + np = 1.$$

Multiplying each side by b gives

$$mab + nbp = b.$$

By hypothesis, p divides ab , hence p divides *both* terms on the left-hand side, and therefore p divides the right-hand side b .

3. Each positive integer has a unique factorization into primes (the *fundamental theorem of arithmetic*).

Suppose on the contrary that some integer n has two different prime factorizations:

$$n = p_1 p_2 \dots p_j = q_1 q_2 \dots q_k.$$

By dividing out common factors, if necessary, we can assume there is a p_i that is not among the q 's. But this contradicts the previous result, because p_i divides $n = q_1 q_2 \dots q_k$, yet it does not divide any of q_1, q_2, \dots, q_k individually, since these are prime numbers $\neq p_i$.

EXERCISES

We can now fill the gap in the proof of Euclid's theorem on perfect numbers (previous exercise set), using the prime divisor property.

3.3.1 Use the prime divisor property to show that the proper divisors of $2^{n-1}p$, for any odd prime p , are $1, 2, 2^2, \dots, 2^{n-1}$ and $p, 2p, 2^2p, \dots, 2^{n-2}p$.

The result that if $\gcd(a, b) = 1$ then $1 = ma + nb$ for some integers m and n is a special case of the following way to represent the gcd.

3.3.2 Show that, for any integers a and b , there are integers m and n such that $\gcd(a, b) = ma + nb$.

This in turn gives a general way to find integer solutions of linear equations.

3.3.3 Deduce from Exercise 3.3.2 that the equation $ax + by = c$ with integer coefficients a , b , and c has an integer solution x , y if $\gcd(a, b)$ divides c .

The converse of this result is also valid, as one discovers when considering a *necessary* condition for $ax + by = c$ to have an integer solution.

3.3.4 The equation $12x + 15y = 1$ has no integer solution. Why?

3.3.5 (Solution of linear Diophantine equations) Give a test to decide, for any given integers a , b , c , whether there are integers x , y such that

$$ax + by = c.$$

3.4 Pell's Equation

The Diophantine equation $x^2 - Dy^2 = 1$, where D is a nonsquare integer, is known as Pell's equation because Euler mistakenly attributed a solution of it to the seventeenth-century English mathematician Pell (it should have been attributed to Brouncker). Pell's equation is probably the best-known Diophantine equation after the equation $a^2 + b^2 = c^2$ for Pythagorean triples, and in some ways it is more important. Solution of Pell's equation is the main step in the solution of the general quadratic Diophantine equation in two variables [see, for example, Gelfond (1961)] and also a key tool in proving the theorem of Matiyasevich mentioned in Section 1.3 that there is no algorithm for solving all Diophantine equations [see, for example, Davis (1973) or Jones and Matiyasevich (1991)]. In view of this, it is fitting that Pell's equation should make its first appearance in the foundations of Greek mathematics, and it is impressive to see how well the Greeks understood it.

The simplest instance of Pell's equation,

$$x^2 - 2y^2 = 1,$$

was studied by the Pythagoreans in connection with $\sqrt{2}$. If x, y are large solutions to this equation, then $x/y \simeq \sqrt{2}$ and in fact the Pythagoreans found a way of generating larger and larger solutions by means of the recurrence relations

$$x_{n+1} = x_n + 2y_n,$$

$$y_{n+1} = x_n + y_n.$$

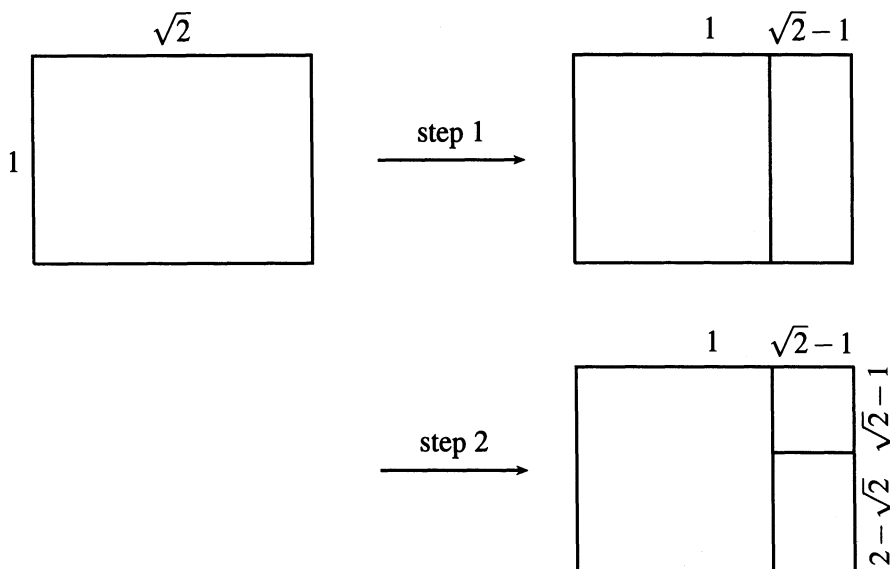
A short calculation shows that

$$x_{n+1}^2 - 2y_{n+1}^2 = -(x_n^2 - 2y_n^2),$$

so if (x_n, y_n) satisfies $x^2 - 2y^2 = \pm 1$, then (x_{n+1}, y_{n+1}) satisfies $x^2 - 2y^2 = \mp 1$. Starting with the trivial solution $(x_0, y_0) = (1, 0)$ of $x^2 - 2y^2 = 1$, we get successively larger solutions $(x_2, y_2), (x_4, y_4), \dots$ of the equation $x^2 - 2y^2 = 1$. [The pairs (x_n, y_n) were known as *side and diagonal numbers* because the ratio y_n/x_n tends to that of the side and diagonal in a square.]

But how might these recurrence relations have been discovered in the first place? Van der Waerden (1976) and Fowler (1980, 1982) suggest that the key is the Euclidean algorithm applied to line segments, an operation the Greeks called *anthyphairesis*. Given any two lengths a, b , one can define the sequence $(a_1, b_1), (a_2, b_2), \dots$, as in Section 3.2, by repeated subtraction of the smaller length from the larger. If a, b are integer multiples of some unit, then the process terminates as in Section 3.3, but if b/a is irrational, it continues forever. We can well imagine that the Pythagoreans would have been interested in *anthyphairesis* applied to $a = 1, b = \sqrt{2}$. Here is what happens. We represent a, b by sides of a rectangle, and each subtraction of the smaller number from the larger is represented by cutting off the square on the shorter side (Figure 3.2). We notice that the rectangle remaining after step 2, with sides $\sqrt{2} - 1$ and $2 - \sqrt{2} = \sqrt{2}(\sqrt{2} - 1)$, is the same shape as the original, though the long side is now vertical instead of horizontal. It follows that similar steps will recur forever, which is another proof that $\sqrt{2}$ is irrational, incidentally.

Our present interest, however, is in the relation between successive similar rectangles. If we let the long and short sides of successive similar rectangles be x_{n+1}, y_{n+1} and x_n, y_n , we can derive a recurrence relations for x_{n+1}, y_{n+1} from Figure 3.3:

Figure 3.2: The Euclidean algorithm on $\sqrt{2}$ and 1

$$x_{n+1} = x_n + 2y_n,$$

$$y_{n+1} = x_n + y_n,$$

exactly the relations of the Pythagoreans! The difference is that our x_n, y_n are not integers, and they satisfy $x^2 - 2y^2 = 0$, not $x^2 - 2y^2 = 1$. Nevertheless, one feels that Figure 3.3 gives the most natural interpretation of these relations. The discovery that the same relations generate solutions of $x^2 - 2y^2 = 1$ possibly arose from wishing that the Euclidean algorithm terminated with $x_1 = y_1 = 1$. If the Pythagoreans started with $x_1 = y_1 = 1$ and applied the recurrence relations, then they could have found that (x_n, y_n) satisfies $x^2 - 2y^2 = (-1)^n$, as we did earlier.

Many other instances of the Pell equation $x^2 - Dy^2 = 1$ occur in Greek mathematics, and these can be understood in a similar way by applying anthyphairesis to the rectangle with sides 1, \sqrt{D} . In the seventh century CE the Indian mathematician Brahmagupta gave a recurrence relation for generating solutions of $x^2 - Dy^2 = 1$, as we shall see in Chapter 5. The

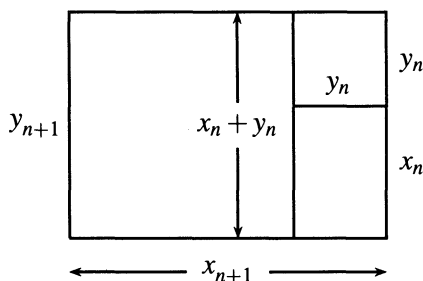


Figure 3.3: The recurrence relation

Indians called the Euclidean algorithm the “pulverizer” because it breaks numbers down to smaller and smaller pieces. To obtain a recurrence one has to know that a rectangle proportional to the original eventually recurs, a fact that was rigorously proved only in 1768 by Lagrange. The later European work on Pell’s equation, which began in the seventeenth century with Brouncker and others, was based on the continued fraction for \sqrt{D} , though this amounts to the same thing as anthyphairesis (see exercises). For a condensed but detailed history of Pell’s equation, see Dickson (1920), pp. 341–400.

An interesting aspect of the theory is the very irregular relationship between D and the number of steps of anthyphairesis before a rectangle proportional to the original recurs. If the number of steps is large, the smallest nontrivial solution of $x^2 - Dy^2 = 1$ is enormous. A famous example is the so-called *cattle problem* of Archimedes (287–212 BCE). This problem leads to the equation

$$x^2 - 4729494y^2 = 1,$$

the smallest solution of which was found by Krummbiegel and Amthor (1880) to have 206,545 digits!

EXERCISES

The continued fraction of a real number $\alpha > 0$ is written

$$\alpha = n_1 + \frac{1}{n_2 + \frac{1}{n_3 + \frac{1}{n_4 + \frac{1}{\ddots}}}}$$

where $n_1, n_2, n_3, n_4, \dots$ are integers obtained by the following algorithm. Let

$n_1 =$ integer part of α .

Then $\alpha - n_1 < 1$ and $\alpha_1 = 1/(\alpha - n_1) > 1$, so we can take

$n_2 =$ integer part of α_1 .

Then $\alpha_1 - n_2 < 1$ and $\alpha_2 = 1/(\alpha_1 - n_2) > 1$, so we can take

$n_3 =$ integer part of α_2 , and so on.

3.4.1 Apply the above algorithm to the number $\alpha = 157/68$, and hence show that

$$\frac{157}{68} = 2 + \frac{1}{3 + \frac{1}{4 + \frac{1}{5}}}.$$

You may notice that what happens is essentially the Euclidean algorithm applied to the pair $(157, 68)$, except that repeated operations of subtraction are replaced by division with remainder. The integers 2, 3, 4, 5 are the successive quotients obtained in these divisions: 157 divided by 68 gives quotient 2 and remainder 21, 68 divided by 21 gives quotient 3 and remainder 5, and so on.

Thus the Euclidean algorithm on integers a, b yields results that may be encoded by the (finite) continued fraction for a/b . This idea was introduced by Euler, and it became the preferred approach to the Euclidean algorithm for some mathematicians. Gauss (1801), in particular, always speaks of the Euclidean algorithm as the “continued fraction algorithm.”

The Euclidean algorithm on a pair $(\alpha, 1)$, where α is irrational, is in fact better known as the continued fraction algorithm.

3.4.2 Interpret the operations in the continued fraction algorithm—detaching the integer part and taking the reciprocal of the remainder—in terms of anthypharesis.

3.4.3 Show that

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{\ddots}}}}.$$

Notice that Exercise 3.4.3 implies $\sqrt{2} + 1$ is the *periodic* continued fraction

$$2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \frac{1}{\ddots}}}}.$$

3.4.4 Show that $\sqrt{3} + 1$ also has a periodic continued fraction, and hence derive the continued fraction for $\sqrt{3}$.

3.5 The Chord and Tangent Methods

In Section 1.3 we used a method of Diophantus to find all rational points on the circle. If $p(x, y) = 0$ is any quadratic equation in x and y with rational coefficients, and if the equation has one rational solution $x = r_1, y = s_1$, then we can find any rational solution by drawing a rational line $y = mx + c$ through the point r_1, s_1 and finding its other intersection with the curve $p(x, y) = 0$. The two intersections with the curve, $x = r_1, r_2$, say, are given by the roots r_1, r_2 of the equation

$$p(x, mx + c) = 0.$$

This means that $p(x, mx + c) = k(x - r_1)(x - r_2)$, and since all coefficients on the left-hand side are rational and r_1 is rational, then k and r_2 must also be rational. The y value when $x = r_2$, $y = s_2 = mr_2 + c$, is rational since m and c are; hence (r_2, s_2) is another rational point on $p(x, y) = 0$. Conversely, any line through two rational points is rational, and hence all rational points are found in this way.

Now if $p(x, y) = 0$ is a curve of degree 3, its intersections with a line $y = mx + c$ are given by the roots of the cubic equation $p(x, mx + c) = 0$. If we know two rational points on the curve, then the line through them will be rational, and its third intersection with the curve will also be rational, by an argument like the preceding one. This fact becomes more useful when one realizes that the two known rational points can be taken to coincide, in which case the line is the tangent through the known rational point. Thus from one rational solution we can generate another by the tangent construction, and from two we can construct a third by taking the chord between the two.

Diophantus found rational solutions to cubic equations in what seems to have been essentially this way. The surviving works of Diophantus reveal little of his methods, but a plausible reconstruction—an algebraic version of the tangent and chord constructions—has been given by Bashmakova (1981). Probably the first to understand Diophantus' methods was Fermat, in the seventeenth century, and the first to give the tangent and chord interpretation was Newton (1670s).

In contrast to the quadratic case, we have no choice in the slope of the rational line for cubics. Thus it is by no means clear that this method will give us *all* rational points on a cubic. A remarkable theorem, conjectured by Poincaré (1901) and proved by Mordell (1922), says that all rational points can be generated by tangent and chord constructions applied to finitely many points. However, it is still not known whether there is an algorithm for finding a finite set of such rational generators on each cubic curve.

EXERCISES

3.5.1 Explain the solution $x = 21/4, y = 71/8$ to $x^3 - 3x^2 + 3x + 1 = y^2$ given by Diophantus [Heath (1910), p. 242] by constructing the tangent through the obvious rational point on this curve.

3.5.2 Rederive the following rational point construction of Viète (1593), p. 145. Given the rational point (a, b) on $x^3 - y^3 = a^3 - b^3$, show that the tangent at (a, b) is

$$y = \frac{a^2}{b^2}(x - a) + b,$$

and that the other intersection of the tangent with the curve is the rational point

$$x = a \frac{a^3 - 2b^3}{a^3 + b^3}, \quad y = b \frac{b^3 - 2a^3}{a^3 + b^3}.$$

3.6 Biographical Notes: Diophantus

Diophantus lived in Alexandria during the period when Greek mathematics, along with the rest of Western civilization, was generally in decline. The catastrophes that engulfed the West with the fall of Rome and the rise of Islam, culminating in the burning of the library in Alexandria in 640 CE, buried almost all details of Diophantus' life. His dates can be placed with certainty only between 150 and 350 CE, since he mentions Hypsicles (known to be around 150) and is mentioned by Theon of Alexandria (around 350). One other scrap of evidence, a letter of Michael Psellus (eleventh century), suggests 250 CE as the most likely time when Diophantus flourished. Apart from this, the only clue to Diophantus' life is a conundrum in the *Greek Anthology* (around 600 CE):

God granted him to be a boy for the sixth part of his life, and adding a twelfth part to this, He clothed his cheeks with down.

He lit him the light of wedlock after a seventh part, and five years after his marriage He granted him a son. Alas! late-born wretched child; after attaining the measure of half his father's life, chill Fate took him. After consoling his grief by this science of numbers for four years he ended his life.

[Cohen and Drabkin (1958), p. 27]

If this information is correct, then Diophantus married at 33 and had a son who died at 42, four years before Diophantus himself died by his own hand at 84.

Diophantus' work went almost unnoticed for many centuries, and only parts of it survive. The first stirrings of interest in Diophantus occurred in the Middle Ages, but much of the credit for the eventual revival of Diophantus belongs to Rafael Bombelli (1526–1572) and Wilhelm Holtzmann (known as Xylander, 1532–1576). Bombelli discovered a copy of Diophantus' *Arithmetic* in the Vatican library and published 143 problems from it in his *Algebra* (1572). The most famous edition of the *Arithmetic* was that of Bachet de Méziriac (1621). Bachet glimpsed the possibility of general principles behind the special problems of the *Arithmetic* and, in his commentary on the book, alerted his contemporaries to the challenge of properly understanding Diophantus and carrying his ideas further. It was Fermat who took up this challenge and made the first significant advances in number theory since the classical era (see Chapter 11).