

the nonterminating Euclidean algorithm far enough to see that it becomes *periodic* in certain cases; for example, when the two line segments have lengths 1 and  $\sqrt{2}$ .

Independently of these developments, the first form of the Euclidean algorithm arose in China in the Han dynasty, between 200 BCE and 200 CE. It was used by the Chinese to simplify fractions—dividing numerator and denominator by their gcd—and also to find integer solutions of linear equations.

A typical “application” of such an equation is the following. Suppose there are  $365\frac{1}{4}$  days in a year and  $29\frac{1}{2}$  days in a lunar month. If we go to units of  $1/4$  day, the year and lunar month are then measured by the integers 1461 and 118. Now suppose there is a full moon on the first day of the year. How long will it be before there is a full moon on the *second* day of the year? This will happen in  $x$  years (and  $y$  months), where

$$1461x = 118y - 4.$$

We therefore seek the least integer solution of this equation and, as we saw in Section 3.3, this depends on expressing  $1 = \gcd(1461, 118)$  as a combination of the form  $118y - 1461x$ , which can be done with the help of the Euclidean algorithm. In the equation, of course, we are only interested in part of the solution—the number  $x$ —because we only want to know *a* multiple of 1461 that is 4 less than *some* multiple of 118 (we don’t care which one). Such a problem would later be described as a *congruence* problem: we seek an  $x$  such that  $1461x$  is *congruent to*  $-4$ , *mod* 118. The Chinese became highly skilled in such problems, extending their methods to multiple congruences, as the next section explains. This led to an important theorem, known today as the *Chinese remainder theorem*.

Around the fifth and sixth centuries CE, similar linear Diophantine equations were solved in India, and perhaps with similar calendar problems in mind. However, the Indians took the idea in a different direction. They independently discovered the Pell equation  $x^2 - Ny^2 = 1$ , found by the Greeks in trying to understand  $\sqrt{N}$ , and also rediscovered the periodicity in it. Most remarkable of all, they did this without any split between rational and irrational. Their treatment of the Pell equation is completely based on integer operations, and it blends smoothly with their treatment of linear equations.

## 5.2 The Chinese Remainder Theorem

The origin of this theorem is the following problem, occurring in the *Mathematical Manual* of Sun Zi, late in the third century CE. It is required to find a number that leaves remainder 2 on division by 3, remainder 3 on division by 5, and remainder 2 on division by 7. The answer can easily be found by experiment to be 23, but Sun Zi offers the following explanation, presumably to indicate a general method.

If we count by threes and there is a remainder 2, put down 140.

If we count by fives and there is a remainder 3, put down 63.

If we count by sevens and there is a remainder 2, put down 30.

Add them to obtain 233 and subtract 210 to get the answer.

[From the translation of Sun Zi's *Mathematical Manual* in Lam and Ang (1992), p. 178.]

The numbers 140, 63, and 30 were chosen because of the following properties:

- $140 = 4 \times (5 \times 7)$   
leaves remainder 2 on division by 3,  
and remainder 0 on division by 5, 7.
- $63 = 3 \times (3 \times 7)$   
leaves remainder 3 on division by 5,  
and remainder 0 on division by 3, 7.
- $30 = 2 \times (3 \times 5)$   
leaves remainder 2 on division by 7,  
and remainder 0 on division by 3, 5.

Hence their sum 233 necessarily leaves remainders 2, 3, 2 on division by 3, 5, 7, respectively. Since  $3 \times 5 \times 7 = 105$  leaves remainder 0 on division by 3, 5, and 7, we can subtract 105 from 233 and obtain a smaller number that leaves the same remainders on division by 3, 5, and 7. Subtracting 105 twice gives the smallest solution, 23.

But why choose 140, 63, and 30, in particular? It would be simpler to choose 35 in place of 140, because

- $35 = 5 \times 7$   
leaves remainder 2 on division by 3,  
and remainder 0 on division by 5, 7.

Sun Zi's explanation continues:

If we count by threes and there is a remainder 1, put down 70.

If we count by fives and there is a remainder 1, put down 21.

If we count by sevens and there is a remainder 1, put down 15.

Apparently he began with  $70 = 2 \times (5 \times 7)$  because it is the smallest multiple of 5 and 7 leaving remainder 1 on division by 3, then multiplied by 2 to get a number leaving remainder 2 on division by 3.

The numbers 63 and 30 can also be explained this way. The smallest multiple of 3 and 7 that leaves remainder 1 on division by 5 is  $21 = 3 \times 7$ . Therefore,  $63 = 3 \times (3 \times 7)$  is a multiple of 3 and 7 that leaves remainder 3 on division by 5. Similarly,  $15 = 3 \times 5$  is the smallest multiple of 3 and 5 that leaves remainder 1 on division by 7, so 30 is the smallest multiple of 3 and 5 that leaves remainder 2 on division by 7.

An interesting question arises at this point. If Sun Zi intended this to be a general method, with integers  $p, q, r$  in place of 3, 5, 7, he needed to know that there is a multiple  $m(qr)$  of  $qr$  that leaves remainder 1 on division by  $p$ . Did he know this? Such a number  $m$  is what we now call an *inverse* of  $qr$ , mod  $p$ , and Sun Zi's problem is probably the first occasion in the history of mathematics where these inverses are called for.

A method for solving Sun Zi's problem in full generality was first given in the *Mathematical Treatise in Nine Sections* by Qin Jiushao in 1247. He solved the crucial problem of finding inverses by the Euclidean algorithm. Given integers  $p$  and  $a$  with  $\gcd(p, a) = 1$ , we know from Section 2.4 that there are integers  $m$  and  $n$  such that

$$mp + na = 1.$$

But then

$$mp = 1 - na,$$

so  $mp$  leaves remainder 1 on division by  $a$ , and  $m$  is an inverse of  $p$ , mod  $a$ . Qin Jiushao found  $m$  by running the Euclidean algorithm on  $p$  and  $a$ , then substituting back to find  $m$  and  $n$  with  $mp + na = 1$ . He called it the "method of finding 1."

It is not hard to show (Exercise 5.2.1) that  $p$  has an inverse mod  $a$  *only* if  $\gcd(p, a) = 1$ . Thus in a Chinese remainder problem we generally need the divisors to be relatively prime. The method of inverses then gives the following.

**Chinese remainder theorem.** *If  $p_1, p_2, \dots, p_k$  are relatively prime integers and  $r_1 \leq p_1, r_2 \leq p_2, \dots, r_k \leq p_k$  are any integers  $\geq 0$ , then there is an integer  $n$  such that  $n$  leaves remainder  $r_i$  on division by  $p_i$  for each  $i$ .  $\square$*

This theorem has made many appearances in the history of number theory, and has often been the vehicle for important new concepts and results. Its later development in China is described in Libbrecht (1973). When it was eventually discovered in Europe, Euler and Gauss made excellent use of it.

### Exercises

**5.2.1** Prove that if  $mp$  leaves remainder 1 on division by  $a$ , then  $\gcd(p, a) = 1$ .

**5.2.2** Give a proof of the Chinese remainder theorem, using the existence of inverses mod  $p_i$  to justify Sun Zi's method.

## 5.3 Linear Diophantine Equations

We have seen how the Chinese came to use the Euclidean algorithm for remainder problems, somewhere between the third century CE and Qin Jiushao's *Mathematical Treatise* of 1247. The algorithm was also used extensively in India during the same period, beginning with the *Âryabhaṭīya* of Âryabhaṭa in 499 CE. Âryabhaṭa was born in 476 CE and is also known as Âryabhaṭa I, to distinguish him from another mathematician of the same name who lived around 950 CE.

His most important contribution was a method for finding integer solutions of equations of the form  $ax + by = c$ , where  $a$ ,  $b$ , and  $c$  are integers. Like the Chinese remainder problem, which it closely resembles, this problem cries out for the Euclidean algorithm. Both problems boil down to expressing  $\gcd(a, b)$  in the form  $ma + nb$ , and in the case of the equation  $ax + by = c$  the underlying reason is the following:

**Criterion for an integer solution of  $ax + by = c$ .** *The equation  $ax + by = c$ , where  $a$ ,  $b$ ,  $c$  are integers, has an integer solution  $\Leftrightarrow \gcd(a, b)$  divides  $c$ .*

**Proof.** If  $x$  and  $y$  are integers, then  $\gcd(a, b)$  divides  $ax + by$ ; hence if  $ax + by = c$ , then  $\gcd(a, b)$  divides  $c$ . Conversely, we know from Section 3.3 that there are integers  $m$  and  $n$  such that  $\gcd(a, b) = ma + nb$ . Hence, if  $\gcd(a, b)$  divides  $c$ , we have  $ma + nb$  divides  $c$ , say  $(ma + nb)d = c$ . Then  $x = md$ ,  $y = nd$  is a solution of the equation  $ax + by = c$ .  $\square$

As mentioned in Section 3.3,  $\gcd(a, b) = ma + nb$  is an easy consequence of the Euclidean algorithm, though Euclid apparently missed it. We also cannot be sure that Âryabhaṭa noticed it, since his book contains

only a few lines on the problem of solving  $ax + by = c$ , and these were only made intelligible by the efforts of later commentators. The first of these was Bhâskara I, who observed in 522 CE that, by dividing  $a$  and  $b$  by their gcd, the problem reduces to solving

$$a'x + b'y = 1,$$

where  $\gcd(a', b') = 1$ , and that the latter problem can always be solved. Thus Bhâskara I assumed that  $1 = \gcd(a', b') = m'a' + n'b'$  for some integers  $m'$  and  $n'$ , and it follows that  $\gcd(a, b) = ma + nb$ , after multiplying both sides by  $\gcd(a, b)$ .

Bhâskara I also introduced the vivid term *kuttaka*, meaning *pulverizer*, for the Euclidean algorithm. The numbers  $a$  and  $b$  are “pulverized” by the algorithm into smaller and smaller parts, with the smallest part being their gcd. The Indian pulverizer was the division-with-remainder form of the algorithm, though of course the word applies equally well to the subtractive form. To solve the equation  $ax + by = c$ , where  $\gcd(a, b)$  divides  $c$ , the pulverizer was combined with substitution to find coefficients  $m$  and  $n$  such that  $ma + nb = \gcd(a, b)$ , and multiplication by a suitable factor to obtain  $x$  and  $y$  such that  $ax + by = c$ . Examples may be seen in Srinivasiengar (1967).

### EXERCISES

Finding  $m$  and  $n$  such that  $\gcd(a, b) = ma + nb$  can be done by running the Euclidean algorithm on the numbers  $a$  and  $b$ , in parallel with the algorithm on the literal *symbols*  $a$  and  $b$ . The symbols  $a$  and  $b$  hold the numbers  $m$  and  $n$  as their coefficients. For example, to find  $m$  and  $n$  such that  $1 = 21m + 17n$ , one runs the Euclidean algorithm starting with the pair  $(21, 17)$ , and also with  $(a, b)$ , doing to the symbols exactly what is done to the numbers.

The first few steps look like this:

$(21, 17)$	$(a, b)$
$(17, 21 - 17)$	$(b, a - b)$
$(17, 4)$	$(b, a - b)$
$(17 - 4, 4)$	$(b - (a - b), a - b)$
$(13, 4)$	$(-a + 2b, a - b)$

So far, this gives  $13 = -21 + 2 \times 17$  and  $4 = 21 - 17$  in the form  $21m + 17n$ .

**5.3.1** Complete the running of the Euclidean algorithm on  $(21, 17)$ , and hence find integers  $m$  and  $n$  such that  $1 = 21m + 17n$ .

**5.3.2** Hence find integers  $x, y$  such that  $21x + 17y = 3$ .

## 5.4 Pell's Equation in Brahmagupta

Where linear Diophantine equations are concerned, Indian mathematics and Chinese mathematics are very similar. In fact, the resemblance is even greater than has been suggested so far, because Chinese remainder problems were also studied in India. This suggests possible contact and sharing of ideas. On the other hand, the two mathematical cultures diverge in other respects. The Chinese developed algebra and approximation methods for high-degree equations, but *not* integer solutions for nonlinear equations (except for the Pythagorean equation). The Indians made less progress in algebra, but had striking success finding integer solutions of the Pell equation—the first major advance in number theory since Diophantus.

The author of this advance was Brahmagupta, whose *Brâhma-sphuṭa-siddhânta* of 628 CE can be read in the English translation of Colebrooke (1817). Brahmagupta's treatment of the Pell equation

$$x^2 - Ny^2 = 1, \quad \text{where } N \text{ is a nonsquare integer,}$$

is based on his discovery [see Colebrooke (1817), p. 363] that

$$(x_1^2 - Ny_1^2)(x_2^2 - Ny_2^2) = (x_1x_2 + Ny_1y_2)^2 - N(x_1y_2 + x_2y_1)^2,$$

which is a generalization of an identity discovered by Diophantus

$$(x_1^2 + y_1^2)(x_2^2 + y_2^2) = (x_1x_2 - y_1y_2)^2 + (x_1y_2 + x_2y_1)^2,$$

to which we shall return later in connection with complex numbers. Like Diophantus' identity, Brahmagupta's is easily checked by multiplying out both sides, though not easily discovered in the first place.

Brahmagupta used his identity to find solutions of

$$x^2 - Ny^2 = 1$$

via a sequence of equations of the form

$$x^2 - Ny^2 = k_i.$$

His identity shows that if

$$x = x_1, \quad y = y_1 \quad \text{is a solution of} \quad x^2 - Ny^2 = k_1,$$

and

$$x = x_2, \quad y = y_2 \quad \text{is a solution of} \quad x^2 - Ny^2 = k_2,$$

then

$$x = x_1x_2 + Ny_1y_2, \quad y = x_1y_2 + x_2y_1 \quad \text{is a solution of} \quad x^2 - Ny^2 = k_1k_2.$$

This is called *composition* of the triples  $(x_1, y_1, k_1)$  and  $(x_2, y_2, k_2)$  to form the triple  $(x_1x_2 + Ny_1y_2, x_1y_2 + x_2y_1, k_1k_2)$ .

If  $k_1 = 1$  or  $k_2 = 1$ , composition is a way to generate infinitely many solutions of  $x^2 - Ny^2 = 1$  when one is known (if only one of  $k_1, k_2$  is 1, compose the corresponding triple with itself). More surprisingly, it is often possible to find a solution of  $x^2 - Ny^2 = 1$  from solutions of

$$x^2 - Ny^2 = k_1 \quad \text{and} \quad x^2 - Ny^2 = k_2 \quad \text{for integers } k_1, k_2 > 1.$$

The reason is that composing  $(x_1, y_1, k_1)$  with itself gives a solution of  $x^2 - Ny^2 = k_1^2$ , say  $x = X, y = Y$ , and hence a *rational* solution  $x = X/k_1, y = Y/k_1$  of  $x^2 - Ny^2 = 1$ . With a bit of luck, this solution will be integral, or else it will yield an integral solution when composed further.

**Example:**  $x^2 - 92y^2 = 1$ . [This is Brahmagupta's first example; he says that "a person solving this problem within a year is a mathematician." See Colebrooke (1817), p. 364.]

**Solution.** Since  $10^2 - 92 \times 1^2 = 8$ , we have the triple  $(10, 1, 8)$ . Composing this with itself gives the triple

$$(10 \times 10 + 92 \times 1 \times 1, 10 \times 1 + 1 \times 10, 8 \times 8) = (192, 20, 64),$$

which means

$$192^2 - 92 \times 20^2 = 8^2.$$

Dividing both sides by  $8^2$  gives

$$24^2 - 92 \times (5/2)^2 = 1,$$

and hence the new "nearly integer" triple  $(24, 5/2, 1)$ . Composing  $(24, 5/2, 1)$  with itself finally gives the integer triple

$$\begin{aligned} (24^2 + 92 \times (5/2)^2, 24 \times (5/2) + (5/2) \times 24, 1) &= (576 + 575, 120, 1) \\ &= (1151, 120, 1). \end{aligned}$$

Thus  $x = 1151, y = 120$  is an integer solution of  $x^2 - 92y^2 = 1$ .  $\square$

## Exercises

**5.4.1** Explain the solutions  $x_{n+1} = x_n + 2y_n$ ,  $y_{n+1} = x_n + y_n$  of  $x^2 - 2y^2 = (-1)^n$  (the “side and diagonal numbers” of Section 3.4) in terms of Brahmagupta composition.

**5.4.2** Derive Brahmagupta’s identity using the factorization

$$(x_1^2 - Ny_1^2)(x_2^2 - Ny_2^2) = (x_1 - \sqrt{N}y_1)(x_1 + \sqrt{N}y_1)(x_2 - \sqrt{N}y_2)(x_2 + \sqrt{N}y_2),$$

and combining the first factor with the third, and the second with the fourth.

**5.4.3** Show that  $\sqrt{N}$  is irrational when  $N$  is a nonsquare integer. Deduce that if  $a_1 - \sqrt{N}b_1 = a_2 - \sqrt{N}b_2$  for integers  $a_1, b_1, a_2, b_2$ , then  $a_1 = a_2$  and  $b_1 = b_2$ .

**5.4.4** If  $(x_3, y_3, 1)$  is the composite of  $(x_1, y_1, 1)$  and  $(x_2, y_2, 1)$ , use Exercise 5.4.3 to show that  $x_3, y_3$  may also be defined as the integers such that

$$(x_1 - \sqrt{N}y_1)(x_2 - \sqrt{N}y_2) = x_3 - \sqrt{N}y_3.$$

Now we free  $x$  and  $y$  from the restriction to integer or rational values, and define the Brahmagupta composite of *any* triples  $(x_1, y_1, 1)$  and  $(x_2, y_2, 1)$  to be  $(x_1x_2 + Ny_1y_2, x_1y_2 + x_2y_1, 1)$

**5.4.5** (For readers familiar with hyperbolic functions.) Show that the functions  $x = \cosh u$ ,  $y = \frac{1}{\sqrt{N}} \sinh u$  define a one-to-one correspondence between the real numbers  $u$  and the points  $(x, y)$  on the branch of the hyperbola  $x^2 - Ny^2 = 1$  where  $x > 1$ . Show also that the Brahmagupta composite of  $(\cosh u_1, \frac{1}{\sqrt{N}} \sinh u_1, 1)$  and  $(\cosh u_2, \frac{1}{\sqrt{N}} \sinh u_2, 1)$  is  $(\cosh(u_1 + u_2), \frac{1}{\sqrt{N}} \sinh(u_1 + u_2), 1)$ , hence Brahmagupta’s composition corresponds to addition of real numbers  $u$ .

**5.4.6** Use the functions  $x = \cos \theta$  and  $y = \sin \theta$  parameterizing the unit circle to show similarly that “Diophantus’ composition” of  $(x_1, y_1)$  and  $(x_2, y_2)$  to form  $(x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$  corresponds to addition of angles  $\theta$ .

## 5.5 Pell’s Equation in Bhâskara II

Brahmagupta found integer solutions of many Pell equations  $x^2 - Ny^2 = 1$  by his composition method, but he was not able to apply it uniformly for all values of  $N$ . The best he could do was show that if  $x^2 - Ny^2 = k$  has an integer solution for  $k = \pm 1, \pm 2$ , or  $\pm 4$  then  $x^2 - Ny^2 = 1$  also has an integer solution. His proofs that composition succeeds in these cases may be found in Srinivasiengar (1967), p. 111.

The first general method for solving the Pell equation was given by Bhâskara II in his *Bījaganita* of 1150 CE. He completed Brahmagupta’s