Lecture 5: Euclid's algorithm

Introduction

The fundamental arithmetic operations are addition, subtraction, multiplication and division. But there is a fifth operation which I would argue is just as fundamental — and that is the operation of taking greatest common divisors.

It might be thought that this operation is not fundamental because it depends on the others for its definition. But this argument also applies to multiplication, which is repeated addition, and division, which is repeated subtraction.

A better argument for the importance of this operation is that it is the key to unlocking many of the deeper properties of the natural numbers. These properties are interesting in themselves and pivotal in appreciating the applications of number theory to cryptography.

In the remainder of this lecture, I shall review the theory of the greatest common divisor of two natural numbers. I shall assume that you have met this before and so in the lecture itself I shall give a summary account whereas in the written notes I shall provide extra information for private study, if you need it.

I shall use the following notation. We denote by $\mathbb{N}$ the set of *natural numbers* — I include zero — and by $\mathbb{Z}$ the set of *integers*.

Gcd's

The following result is simple but at the same time very useful. It can be proved using the following idea. For simplicity let's assume that both a and b are positive. If $0 < a < b$ then $b \cdot 0 < a < b \cdot 1$. If $a \geq b$ then we can always find a q such that $bq \leq a < b(q+1)$. We therefore have the following.

Lemma 0.1 (Remainder Theorem)**.** *Let $a, b \in \mathbb{Z}$ where $b > 0$. Then there are unique integers q and r such that*

$$a = bq + r$$

where $0 \leq r < b$.

The number q is called the *quotient* and the number r is called the *remainder*. For example, if we consider the pair of natural numbers 14 and 3 then

$$14 = 3 \cdot 4 + 2$$

where 4 is the quotient and 2 is the remainder.

Let a and b be integers. We say that a *divides* b if there is a q such that $b = aq$. In other words, there is no remainder. We also say that a

is a *divisor* of b. We write $a \mid b$ to mean the same thing as 'a divides b'.[1]

Warning! $a \mid b$ does not mean the same thing as $\frac{a}{b}$. The latter is a number, the former is a statement about two numbers.

Let $a, b \in \mathbb{N}$. A number d which divides both a and b is called a *common divisor*. The largest number which divides both a and b is called the *greatest common divisor* of a and b and is denoted by $\gcd(a, b)$. A pair of natural numbers a and b is said to be *coprime* if $\gcd(a, b) = 1$.

Special case We define $\gcd(0, 0) = 0$ for completeness.

Example 0.2. Consider the numbers 12 and 16. The set of divisors of 12 is the set $\{1, 2, 3, 4, 6, 12\}$. The set of divisors of 16 is the set $\{1, 2, 4, 8, 16\}$. The set of common divisors is the intersection of these two sets: namely, $\{1, 2, 4\}$. The largest common divisor of 12 and 16 is therefore 4. Thus $\gcd(12, 16) = 4$.

A simple practical application of greatest common divisors is in simplifying fractions. For example, the fraction $\frac{12}{16}$ is equal to the fraction $\frac{3}{4}$ because we can divide out the common factor of numerator and denominator. The fraction which results cannot be simplified further and is in its *lowest terms*. We now justify this claim. The following result tells us that if we divide out the greatest common divisor of a pair of numbers, then the pair of numbers that results is coprime.

Lemma 0.3. *Let* $d = \gcd(a, b)$. *Then* $\gcd(\frac{a}{d}, \frac{b}{d}) = 1$.

Proof. Let $a = a'd$ and $b = b'd$. Suppose that $e \mid \frac{a}{d}$ and $e \mid \frac{b}{d}$. Then $\frac{a}{d} = ex$ and $\frac{b}{d} = ey$ for some natural numbers x and y. Thus $a = exd$ and $b = eyd$. Observe that $ed \mid a$ and $ed \mid b$. But d is the greatest common divisor and so $e = 1$, as required. □

If the numbers a and b are large, then calculating their gcd in the way suggested by the definition would be time-consuming and error-prone. The definition of the gcd of two numbers gives no clue that there might be a fast way of computing it. We want to find an *efficient* way of calculating the greatest common divisor. The following lemma is the basis of just such an efficient method.

[1] Observe that if a is nonzero, then $a \mid a$, if $a \mid b$ and $b \mid a$ then $a = \pm b$, and finally if $a \mid b$ and $b \mid c$ then $a \mid c$.

Lemma 0.4. *Let $a, b \in \mathbb{N}$, where $b \neq 0$, and let $a = bq + r$ where $0 \leq r < b$ by the Remainder Theorem. Then*

$$\gcd(a, b) = \gcd(b, r).$$

Proof. Let d be a common divisor of a and b. Since $a = bq + r$ we have that $a - bq = r$ so that d is also a divisor of r. It follows that any divisor of a and b is also a divisor of b and r.

Now let d be a common divisor of b and r. Since $a = bq + r$ we have that d divides a. Thus any divisor of b and r is a divisor of a and b.

It follows that the set of common divisors of a and b is the same as the set of common divisors of b and r. Thus $\gcd(a, b) = \gcd(b, r)$. □

The point is that $b < a$ and $r < b$. So calculating $\gcd(b, r)$ will be easier than calculating $\gcd(a, b)$ because the numbers involved are smaller. Compare

$$\overbrace{a = b}\,q + r$$

with

$$a = \underbrace{bq + r}.$$

The above result is the basis of an efficient algorithm for computing greatest common divisors. It was described by Euclid around 300 BC in his book the *Elements* in Propositions 1 and 2 of Book VII.

Algorithm 0.5 (Euclid's algorithm)**.**
Input: $a, b \in \mathbb{N}$ such that $a \geq b$ and $b \neq 0$.
Output: $\gcd(a, b)$.
Procedure: write $a = bq + r$ where $0 \leq r < b$. Then $\gcd(a, b) = \gcd(b, r)$. If $r \neq 0$ then repeat this procedure with b and r and so on. The last non-zero remainder is $\gcd(a, b)$

Example 0.6. Let's calculate $\gcd(19, 7)$ using Euclid's algorithm. I have highlighted the numbers that are involved at each stage.

$$\begin{array}{rcl}
\mathbf{19} & = & \mathbf{7} \cdot 2 + \mathbf{5} \\
\mathbf{7} & = & \mathbf{5} \cdot 1 + \mathbf{2} \\
\mathbf{5} & = & \mathbf{2} \cdot 2 + \mathbf{1} \ * \\
\mathbf{2} & = & \mathbf{1} \cdot 2 + 0
\end{array}$$

By our result above we have that

$$\gcd(19, 7) = \gcd(7, 5) = \gcd(5, 2) = \gcd(2, 1) = \gcd(1, 0).$$

The last non-zero remainder is 1 and so $\gcd(19, 7) = 1$ and, in this case, the numbers are coprime.

Theorem 0.7 (Bézout)**.** *There are* integers x *and* y *such that*

$$\gcd(a,b) = xa + yb.$$

I shall prove this theorem using the following.

Algorithm 0.8 (Extended Euclidean algorithm)**.**
Input: $a, b \in \mathbb{N}$ where $a \geq b$ and $b \neq 0$.
Output: numbers $x, y \in \mathbb{Z}$ such that $\gcd(a,b) = xa + yb$.
Procedure: apply Euclid's algorithm to a and b; working from bottom to top rewrite each remainder in turn.

Example 0.9. This is a little involved so I have split the process up into steps. I shall apply the extended Euclidean algorithm to the example I calculated above. I have highlighted the non-zero remainders wherever they occur, and I have discarded the last equality where the remainder was zero. I have also marked the last non-zero remainder.

$$\begin{aligned} 19 &= 7 \cdot 2 + \mathbf{5} \\ 7 &= \mathbf{5} \cdot 1 + \mathbf{2} \\ 5 &= \mathbf{2} \cdot 2 + \mathbf{1} \ * \end{aligned}$$

The first step is to rearrange each equation so that the non-zero remainder is alone on the lefthand side.

$$\begin{aligned} \mathbf{5} &= 19 - 7 \cdot 2 \\ \mathbf{2} &= 7 - \mathbf{5} \cdot 1 \\ \mathbf{1} &= \mathbf{5} - \mathbf{2} \cdot 2 \end{aligned}$$

Next we reverse the order of the list

$$\begin{aligned} \mathbf{1} &= \mathbf{5} - \mathbf{2} \cdot 2 \\ \mathbf{2} &= 7 - \mathbf{5} \cdot 1 \\ \mathbf{5} &= 19 - 7 \cdot 2 \end{aligned}$$

We now start with the first equation. The lefthand side is the gcd we are interested in. We treat all other remainders as algebraic quantities and systematically substitute them in order. Thus we begin with the first equation

$$\mathbf{1} = \mathbf{5} - \mathbf{2} \cdot 2.$$

The next equation in our list is

$$\mathbf{2} = 7 - \mathbf{5} \cdot 1$$

so we replace **2** in our first equation by the expression on the right to get

$$\mathbf{1} = \mathbf{5} - (7 - \mathbf{5} \cdot 1) \cdot 2.$$

We now rearrange this equation by collecting up like terms treating the highlighted remainders as algebraic objects to get

$$\mathbf{1} = 3 \cdot \mathbf{5} - 2 \cdot 7.$$

We can of course make a check at this point to ensure that our arithmetic is correct. The next equation in our list is

$$\mathbf{5} = 19 - 7 \cdot 2$$

so we replace **5** in our new equation by the expression on the right to get

$$\mathbf{1} = 3 \cdot (19 - 7 \cdot 2) - 2 \cdot 7.$$

Again we rearrange to get

$$\mathbf{1} = 3 \cdot \mathit{19} - 8 \cdot \mathit{7}.$$

The algorithm now terminates and we can write

$$\gcd(19, 7) = 3 \cdot \mathit{19} + (-8) \cdot \mathit{7},$$

as required. We can also, of course, easily check the answer!

Here is an application of Bézout's theorem.

Lemma 0.10. *$c \mid a$ and $c \mid b$ iff $c \mid \gcd(a, b)$*

Proof. Let $d = \gcd(a, b)$. Suppose that $c \mid a$ and $c \mid b$. We can find integers x and y such that $d = xa + yb$. It follows immediately that $c \mid d$. Conversely, suppose that $c \mid d$. By definition $d \mid a$ and $d \mid b$ so it is immediate that $c \mid a$ and $c \mid b$. □

Euclid's *Elements*

Euclid's algorithm turns out to be of fundamental importance in modern cryptography. It is therefore perhaps surprising that it is 2,300 years old.

We know virtually nothing about Euclid himself although by the close scrutiny of ancient texts scholars have deduced that he lived around 300BC, that he was probably educated in Athens, and that his working life was spent in Alexandria.

Despite the obscurity of his life, he is famous because of the book he wrote known in English as the *Elements* from the Greek *Stoicheia*. This book is the single most influential maths book ever written, arguably the most influential science book ever written, and one of the most influential books — period, as the Americans would say — ever written.

I have some difficulty in whether to call it a 'book' or 'books'. It is usually described as consisting of thirteen books, numbered I-XIII, but we would nowadays regard these as individual chapters each of which being originally written on a single roll of papyrus.

Euclid's *magnum opus* is commonly regarded as a geometry book and it is certainly true that it contains the foundations of both plane and solid geometry: Pythagoras' theorem is the highlight of Book I, Book IV constructs some regular polygons and Book XIII is all about the Platonic solids. The geometric aspects of the Elements were the foundations of building and surveying — the great European cathedrals contain embodiments of some of Euclid's theorems — but they also stirred the imagination of subsequent mathematicians leading to the development of non-Euclidean geometry in the nineteenth century. Our modern understanding of the large-scale structure of the universe is based on this mathematics.

But this book also contains some basic algebra, although it is disguised to our eyes as geometry, in Book II, and, particularly relevant to this course, it contains the basics of number theory in Books VII and IX.

We shall meet some more of Euclid's results over the course of the next few lectures.

Blankinship's algorithm

This is an alternative procedure to the extended Euclidean algorithm that delivers exactly the same information but in a much easier form and is the one I recommend. It uses matrix theory and was described by W. A. Blankinship in 'A new version of the Euclidean algorithm' *American Mathematical Monthly* **70** (1963), 742–745. To explain how it works, let's go back to the basic step of Euclid's algorithm. If $a \geq b$ then we divide b into a and write

$$a = bq + r$$

where $0 \leq r \leq b$. The key point is that $\gcd(a, b) = \gcd(b, r)$. We shall now think of (a, b) and (b, r) as column matrices

$$\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} r \\ b \end{pmatrix}.$$

We want the 2×2 matrix that maps

$$\begin{pmatrix} a \\ b \end{pmatrix} \text{ to } \begin{pmatrix} r \\ b \end{pmatrix}.$$

This is the matrix

$$\begin{pmatrix} 1 & -q \\ 0 & 1 \end{pmatrix}.$$

Thus

$$\begin{pmatrix} 1 & -q \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} r \\ b \end{pmatrix}.$$

Finally, we can describe the process by the following matrix operation

$$\left(\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b \end{array}\right) \to \left(\begin{array}{rr|c} 1 & -q & r \\ 0 & 1 & b \end{array}\right)$$

by carrying out an elementary row operation. This procedure can be iterated. It will terminate when one of the entries in the righthand column is 0. The non-zero entry will then be the greatest common divisor of a and b and the matrix on the lefthand side will tell you how to get to $(0, \gcd(a, b))$ from (a, b) and so will provide the information that the Euclidean algorithm provides. All of this is best illustrated by means of an example.

Let's calculate x, y such that $\gcd(2520, 154) = xa + yb$. We start with the matrix

$$\left(\begin{array}{cc|r} 1 & 0 & 2520 \\ 0 & 1 & 154 \end{array}\right)$$

If we divide 154 into 2520 it goes 16 times plus a remainder. Thus we subtract 16 times the second row from the first to get

$$\left(\begin{array}{rr|r} 1 & -16 & 56 \\ 0 & 1 & 154 \end{array}\right)$$

We now repeat the process but, since the larger number, 154, is on the bottom, we have to subtract some multiple of the first row from the second. This time we subtract twice the first row from the second to get

$$\left(\begin{array}{rr|r} 1 & -16 & 56 \\ -2 & 33 & 42 \end{array}\right)$$

Now repeat this procedure to get

$$\left(\begin{array}{rr|r} 3 & -49 & 14 \\ -2 & 33 & 42 \end{array}\right)$$

And again

$$\left(\begin{array}{rr|r} 3 & -49 & 14 \\ -11 & 180 & 0 \end{array}\right)$$

The process now terminates because we have a zero in the rightmost column. The non-zero entry in the rightmost column is $\gcd(2520, 154)$. We also know that

$$\begin{pmatrix} 3 & -49 \\ -11 & 180 \end{pmatrix} \begin{pmatrix} 2520 \\ 154 \end{pmatrix} = \begin{pmatrix} 14 \\ 0 \end{pmatrix}.$$

Now this matrix equation corresponds to two equations. The bottom one can be verified. The top one says that

$$14 = 3 \times 2520 - 49 \times 154.$$

Gauss's Lemma

We can use Bézout's theorem to prove a result which is one of the most useful in number theory. Suppose that $c \mid ab$. In general, we cannot make any deductions about whether c divides a or b. For example, $15 \mid 6 \times 35$ but neither 6 nor 35 are divisible by 15. However, if we know *in addition* that $\gcd(c, a) = 1$, that is that c and a are coprime, then we can deduce that $c \mid b$. This result is called *Gauss's Lemma.* Here is the proof. We are told that $\gcd(c, a) = 1$. By Bézout's theorem there exist integers x and y such that

$$1 = cx + ay.$$

Multiply both sides of this equation by b to get

$$b = bcx + aby.$$

Now $c \mid bcx$ and $c \mid aby$ and so $c \mid b$, as required.

Linear Diophantine equations

This is an application of Bézout's theorem which again I assume you have met before. The details here are therefore for private study if you haven't. Named after the third century Greek mathematician Diophantus, a *Diophantine equation* is an equation where we are interested only in the integer solutions. In this section, we are interested in equations of the form

$$ax + by = c$$

where a, b, c are integers and where we require solutions (x, y) to be integers as well. Think geometrically: $ax + by = c$ is a line in the plane and we want to know which lattice points are on this line where a *lattice point* is a point (x, y) where both x and y are integers.

Theorem 0.11. *A necessary and sufficient condition for the equation*

$$ax + by = c$$

to have an integer solution is that $\gcd(a,b) \mid c$*. If this condition is satisfied, and* (x_0, y_0) *is any one solution and* $d = \gcd(a,b)$ *then all solutions are obtained as follows*

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + n \begin{pmatrix} \frac{b}{d} \\ -\frac{a}{d} \end{pmatrix}$$

where $n \in \mathbb{Z}$ *is arbitrary.*

Proof. I shall sketch out the proof. Suppose first that $ax_0 + by_0 = c$ is a solutions in integers. Then clearly $\gcd(a,b) \mid c$.

We now prove tha converse. Suppose that $\gcd(a,b) \mid c$. Put $d = \gcd(a,b)$. Then $\frac{a}{d}$ and $\frac{b}{d}$ are coprime. Thus by Bézout's theorem there are integers x' and y' such that

$$1 = \frac{a}{d}x' + \frac{b}{d}y'.$$

If we multiply both sides by c we get that

$$c = \frac{a}{d}cx' + \frac{b}{d}cy'.$$

We may write this as

$$c = a(\frac{c}{d}x') + b(\frac{c}{d}y').$$

Put $x_0 = \frac{c}{d}x'$ and $y_0 = \frac{c}{d}y'$ both integers by our assumption. We have proved that the equation has a solution and we have shown how to find one.

It is an easy exercise to check that for any $n \in \mathbb{Z}$, the following are all solutions

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + n \begin{pmatrix} \frac{b}{d} \\ -\frac{a}{d} \end{pmatrix}$$

It remains to show now that every solution has the above form. Let $ax + by = c$ be any solution. Subtract from this the solution $ax_0 + by_0 = c$ to get $a(x - x_0) + b(y - y_0) = 0$. Thus

$$\frac{a}{d}(x - x_0) = -\frac{b}{d}(y - y_0).$$

But $\frac{a}{d}$ and $\frac{b}{d}$ are coprime. We now apply Gauss's Lemma. We deduce that $\frac{b}{d}$ divides $x - x_0$. We may therefore write

$$x = x_0 + n\frac{b}{d}$$

for some $n \in \mathbb{Z}$. But this implies that $y - y_0 = -n\frac{a}{d}$. □