LECTURE 1: DIVISIBILITY

1. Introduction

Number theory concerns itself with studying the multiplicative and additive structure of the natural numbers

$$\mathbb{N} = \{1, 2, 3, \dots\}.$$

Frequently, number theoretic questions are better asked in the set of all integers

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\},$$

and better answered by making use of the rational numbers

$$\mathbb{Q} = \left\{ \frac{p}{q} \,:\, p \in \mathbb{Z}, q \in \mathbb{N} \right\},$$

the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$, where more structure may become apparent.

Some form of number theory was developed by the ancient Babylonians, Egyptians and Greeks, and many modern problems are motivated by this work. When studying other long-standing areas of mathematics, such as Euclidean geometry, calculus or linear algebra, it is easy to come away with the conclusion that everything was worked out long ago. Number theory is not like that, and for many problems, some of them ancient, we have more conjectures than theorems. Nevertheless, new methods and results emerge in fits and starts, and the subject has seen many great advances in just the last few decades. Here are a few examples of number-theoretic problems that have been solved only recently or still remain open.

Problem 1.1 (Fermat's Last Theorem)**.** *For any integer $n \geqslant 3$, the equation $x^n + y^n = z^n$ has no solutions with $x, y, z \in \mathbb{N}$.*

This conjecture was stated by Fermat around 1637, and was motivated by much earlier work of Diophantus (c. 200-300AD). It was finally proven by Wiles in 1995.

Problem 1.2 (Congruent Number Problem)**.** *For which $n \in \mathbb{N}$ is there a right triangle with rational sides and area n ($2n = xy$ with $x^2+y^2 = z^2$, $x, y, z \in \mathbb{Q}$)?*

Such an n is called a *congruent number*. For instance, 6 is congruent because it is the area of the 3–4–5 right triangle, but it can be shown that 3 is not a congruent number. The question can be very subtle, as illustrated by the example 53, which is congruent, but for which the simplest suitable right triangle has legs $\frac{1472112483}{202332130}$ and $\frac{21447205780}{1472112483}$.

This question appeared in 10th century Arab manuscripts, but is possibly even older. In 1983, Tunnell gave a simple numerical criterion for determining whether a given n is congruent or not, but its correctness depends on the unproven Birch and

Swinnerton-Dyer conjecture (one of the $1 million Clay Millennium Prize problems). Nevertheless, there is an algorithm that is guaranteed to work if the BSD conjecture is true (even if we cannot prove it) and efficient in practice.

Problem 1.3 (Catalan's Conjecture)**.** *The only consecutive powers of natural numbers are* 8 *and* 9 *($x^n - y^m = 1$).*

This was conjectured by Catalan in 1844 (though again it is no doubt much older) and proven by Mihăilescu in 2002.

Problem 1.4 (Twin Prime Conjecture)**.** *There are infinitely many pairs of prime numbers that differ by* 2*.*

This conjecture was first stated in print by de Polignac in 1849, but its origins are probably much older, perhaps as far back as Euclid, who recorded a proof that there are infinitely many primes in the *Elements*, c. 300BC; we will study Euclid's proof early on in the course. In July 2014, the mathematics consortium D. H. J. Polymath, led by Terry Tao, has built on the pivotal work of Yitang Zhang and James Maynard to prove that there are infinitely many pairs of primes that differ by at most 246.

Problem 1.5 (Goldbach Conjecture)**.** *Every integer $n > 1$ can be expressed as the sum of at most three prime numbers.*

Goldbach stated this conjecture in a letter to Euler in June 1742. Euler replied that it is equivalent to the statement "every even integer $n > 2$ is the sum of two prime numbers", and this is often taken as the statement of the problem. In 2013, Helfgott claimed a proof of Goldbach's conjecture for *odd numbers* n. The problem for even n, including Euler's reformulation, remains open. However, it is known that "almost all" even natural numbers can indeed be written as the sum of two primes.

Problem 1.6 (ABC Conjecture)**.** *For each $\varepsilon > 0$, there exists $C_\varepsilon > 0$ (depending only on ε) such that whenever $abc \neq 0$ and $a + b + c = 0$, then*

$$\max\{|a|, |b|, |c|\} \leqslant C_\varepsilon \Big(\prod_{p|abc} p \Big)^{1+\varepsilon}$$

where the product is taken over distinct prime divisors of a, b and c.

The ABC Conjecture was stated by Oesterlé and Masser in 1985. It has many profound implications, but until very recently seemed far beyond reach. In 2012 Shinichi Mochizuki has recently claimed to have proved this conjecture, however, and there is considerable activity attempting to verify his proof.

Problem 1.7. *Is the number*

$$\zeta(k) = \sum_{n=1}^{\infty} \frac{1}{n^k}$$

irrational for every integer $k > 1$?

This question has its roots in the Basel problem from 1644, which asked for the value of $\zeta(2)$. Euler became famous when he solved the Basel problem in 1734,

proving that $\zeta(2) = \frac{\pi^2}{6}$. More generally, $\zeta(k)$ is a rational multiple of π^k whenever k is *even*. Lindemann proved that π is transcendental in 1844, and it follows that $\zeta(k)$ is irrational for every even k. Apéry stunned the number theory community in 1978 by proving that $\zeta(3)$ is irrational. Since then, Apéry's methods have been broadened to prove that $\zeta(k)$ is irrational for infinitely many odd values of k, but the full question remains open.

2. Divisibility

Definition 2.1. (i) Suppose that $a, b \in \mathbb{Z}$. We say that b **divides** a (written $b \mid a$) when there exists $c \in \mathbb{Z}$ such that $a = bc$. In such circumstances, we say that a is **divisible** by b, or that b is a **divisor** of a;
(ii) When a is not divisible by b, we write $b \nmid a$;
(iii) When $b \mid a$ and $1 \leqslant b < a$, we say that b is a **proper divisor** of a;
(iv) We write $a^k \parallel b$ when $a^k \mid b$ but $a^{k+1} \nmid b$.
It is understood that $b \mid a$ makes sense only when b is non-zero.

The next theorem records the basic properties of divisibility that are intuitively clear, but easily established from the definition.

Theorem 2.2. *(i) $a \mid a$ for every $a \in \mathbb{Z} \setminus \{0\}$;*
(ii) $a \mid 0$ for every $a \in \mathbb{Z} \setminus \{0\}$;
(iii) if $a \mid b$ and $b \mid c$, then $a \mid c$;
(iv) if $a \mid b$ and $a \mid c$, then for all $x, y \in \mathbb{Z}$, one has $a \mid (bx + cy)$;
(v) if $a \mid b$ and $b \mid a$, then $a = \pm b$;
(vi) if $a \mid b$ and $a > 0$ and $b > 0$, then $a \leqslant b$;
(vii) when $m \neq 0$, one has $a \mid b \Leftrightarrow ma \mid mb$.

Proof. We will leave these assertions as exercises, though in order to illustrate ideas, we will give a formal proof of part (vii). Suppose that $m \neq 0$ and $a \mid b$. Then there exists $c \in \mathbb{Z}$ with the property that $b = ac$, whence $mb = m(ac)$. So there exists $c \in \mathbb{Z}$ with the property that $(mb) = (ma)c$, whence by the definition of divisibility $(ma) \mid (mb)$. Conversely, if $m \neq 0$ and $ma \mid mb$, then there exists $c \in \mathbb{Z}$ with $mb = (ma)c$. But since $m \neq 0$, the latter implies that $b = ac$. So there exists $c \in \mathbb{Z}$ with the property that $b = ac$, so from the definition of divisibility, one has $a \mid b$. □

The next theorem lays the groundwork for the development of the theory of congruences.

Theorem 2.3 (The Division Algorithm)**.** *For any $a, b \in \mathbb{Z}$ with $a > 0$, there exist unique integers q and r with $b = qa + r$ and $0 \leqslant r < a$. If, further, $a \nmid b$, then the stronger inequality $0 < r < a$ holds.*

Proof. Let aq be the greatest multiple of a not exceeding b. Then if we put $r = b - aq$, one has $r \geqslant 0$. Moreover, by hypothesis one has $a(q + 1) > b$, and thus $r = b - aq < a$. This establishes the existence of the integers q and r as stated. In order to establish uniqueness, suppose that another pair

q', r' satisfy analogous conditions. If $r \neq r'$, there is no loss of generality in supposing that $r < r'$. Then since $aq'+r' = b = aq+r$, one has $a(q-q') = r'-r$, whence $a \mid (r'-r)$ and $0 < r'-r < a$. But the latter contradicts case (vi) of Theorem 2.2 (which would imply that $r'-r \geqslant a$). Thus we find that $r = r'$, and this now leads to the equation $qa = q'a$. But a is non-zero, so $q = q'$. Thus we find that $(q, r) = (q', r')$, and this establishes uniqueness.

Finally, if $r = 0$ then $b = qa$, whence $a \mid b$. The final assertion of the theorem is now immediate. □

Definition 2.4. (i) Suppose that $a \in \mathbb{Z} \setminus \{0\}$ and $b, c \in \mathbb{Z}$. We say that a is a **common divisor** of b and c when $a \mid b$ and $a \mid c$;
(ii) When b and c are not both zero, the number of common divisors of b and c is finite (see Theorem 2.2(vi)), and thus we may define the **greatest common divisor** (or **highest common factor**) of b and c to be the largest (positive) common divisor. The greatest common divisor of b and c is written (b, c) (or $\gcd(b, c)$ or $\mathrm{hcf}(b, c)$);
(iii) When $g_1, \dots, g_n$ are integers, not all zero, we similarly write $(g_1, \dots, g_n)$ for the greatest integer d satisfying the condition that $d \mid g_i$ $(1 \leqslant i \leqslant n)$.

Example 2.5. One has $(0, 2) = 2$, $(1, 3) = 1$ and $(1729, 182) = 91$ (at this point one can use trial and error, observing that (a, b) must be at most $\min\{|a|, |b|\}$).

The next theorem provides a useful tool to establish simple properties of greatest common divisors.

Theorem 2.6. *If $g = (b, c)$, then there exist integers x and y with $g = bx+cy$.*

Proof. Define the integer d by setting

$$d = \min\{bu + cv \,:\, u, v \in \mathbb{Z} \text{ and } bu + cv > 0\}.$$

Also, let x and y be the values of u and v corresponding to this minimum, so that $d = bx + cy$.

We first prove that $d \mid b$. If to the contrary $d \nmid b$, then by the Division Algorithm (Theorem 2.3), there exist integers r and q with $b = dq + r$ and $0 < r < d$. Then

$$r = b - dq = b - q(bx + cy) = b(1 - qx) + c(-qy),$$

whence

$$r \geqslant \min\{bu + cv \,:\, u, v \in \mathbb{Z} \text{ and } bu + cv > 0\} = d.$$

This gives a contradiction, since $r < d$, and thus we find that $d \mid b$.

A similar argument shows that $d \mid c$, and thus d is indeed a common divisor of b and c, which is to say that $d \leqslant (b, c)$. But $g = (b, c)$, and so there exist integers B and C with $b = gB$ and $c = gC$. Consequently, one has $d = g(Bx + Cy)$, and hence $g \mid d$. Thus $g > 0$, $d > 0$ and $g \mid d$, so by Theorem 2.2(vi) one has $g \leqslant d$. Then one has $d \geqslant (b, c)$ in addition to the relation $d \leqslant (b, c)$ which we derived above, so that necessarily $d = (b, c)$. But then $(b, c) = bx + cy$, and this completes the proof of the theorem. □

Theorem 2.7. *The greatest common divisor of b and c is:*
(i) the least positive value of $bx + cy$, as x and y range over $\mathbb{Z}$;
(ii) the positive common divisor of b and c that is divisible by all other such divisors.

Proof. The assertion (i) is plain from Theorem 2.6. For part (ii), observe that there exist integers x and y with $(b, c) = bx + cy$. Then if $d \mid b$ and $d \mid c$, say $b = dB$ and $c = dC$, one finds that $(b, c) = d(Bx + Cy)$, whence $d \mid (b, c)$. So (b, c) is divisible by all other positive common divisors of b and c. □

Remark 2.8. If $g_1, \dots, g_n$ are not all zero, then it follows as in the proof of Theorem 2.6 that there exist integers $x_1, \dots, x_n$ with $(g_1, \dots, g_n) = g_1x_1 + \cdots + g_nx_n$.

The criterion for determining the greatest common divisor recorded in Theorem 2.6, and (in modified form) in Theorem 2.7, provides a simple and direct approach to establishing simple properties of the greatest common divisor function.

Theorem 2.9. *Whenever $m \in \mathbb{N}$, one has $(ma, mb) = m(a, b)$.*

Proof. Making use of Theorem 2.7(i) (twice), one has

$$\begin{aligned}(ma, mb) &= \min\{max + mby : x, y \in \mathbb{Z} \text{ and } max + mby > 0\}\\ &= m \min\{ax + by : x, y \in \mathbb{Z} \text{ and } ax + by > 0\}\\ &= m(a, b).\end{aligned}$$

□

Remark 2.10. Similiarly, when $d \in \mathbb{N}$, and $d \mid a$ and $d \mid b$, one has $(a/d, b/d) = (a, b)/d$. In particular, if $g = (a, b)$, then $(a/g, b/g) = 1$.

Proof. The first assertion follows from Theorem 2.9 by means of the relation $(d(a/d), d(b/d)) = d(a/d, b/d)$, and the second is immediate from the first. □

Theorem 2.11. *Whenever a, b, m are integers with $(a, m) = (b, m) = 1$, one has $(ab, m) = 1$.*

Proof. By Theorem 2.6, there exist integers x, y, u, v with $1 = ax + my = bu + mv$. Thus we obtain

$$(ax)(bu) = (1 - my)(1 - mv) = 1 - mw,$$

say, with $w = y + v - mvy$. Consequently, one has $(ab)(xu) + mw = 1$. But then by Theorem 2.2(iv), any common divisor of ab and m divides 1. We therefore conclude that $(ab, m) = 1$. □

Theorem 2.12. *For any integer x, and for any integers a and b, not both zero, one has*

$$(a, b) = (b, a) = (a, -b) = (a, b + ax).$$

Proof. The first assertions of the theorem are plain from Theorem 2.7(i). In order to prove that $(a, b) = (a, b+ax)$, observe that by Theorem 2.6, there exist integers u and v with $(a, b) = au+bv$, whence $(a, b) = a(u-xv)+(b+ax)v$. We therefore have $(a, b+ax) \mid (a, b)$. But $(a, b) \mid a$ and $(a, b) \mid b$, so $(a, b) \mid (b+ax)$. But now we have $(a, b + ax) \mid (a, b) \mid (a, b + ax)$, and so by virtue of positivity, Theorem 2.2(v) establishes the desired conclusion. □

Theorem 2.13. *Suppose that $c \mid ab$ and $(b, c) = 1$. Then $c \mid a$.*

Proof. By Theorem 2.9, the hypotheses of the theorem imply that $(ab, ac) = |a|(b, c) = |a|$. But by hypothesis, one has $c \mid ab$, which implies that $c \mid (ab, ac)$. We thus conclude that $c \mid a$. □

At last we are positioned to describe an algorithm for calculating greatest common divisors. Of course, by exhaustive checking one could determine the greatest common divisor of two integers b and c with $O(\min\{|b|, |c|\})$ applications of the Division Algorithm, but the Euclidean Algorithm requires only $O(\log \min\{|b|, |c|\})$ such divisions.

Theorem 2.14 (Euclidean Algorithm)**.** *Suppose that $b \in \mathbb{Z}$ and $c \in \mathbb{N}$. Define the integers r_i and q_i for $i \geqslant 1$ by repeated application of the Division Algorithm thus:*

$$\begin{aligned} b &= cq_1 + r_1, \quad \textit{with } 0 < r_1 < c, \\ c &= r_1q_2 + r_2, \quad \textit{with } 0 < r_2 < r_1, \\ r_1 &= r_2q_3 + r_3, \quad \textit{with } 0 < r_3 < r_2, \\ &\vdots \\ r_{j-2} &= r_{j-1}q_j + r_j, \quad \textit{with } 0 < r_j < r_{j-1}, \\ r_{j-1} &= r_jq_{j+1}. \end{aligned}$$

(Here we adopt obvious conventions if the process terminates prematurely.) Then $(b, c) = r_j$, the last non-zero remainder in the division process.

Proof. Repeated application of Theorem 2.12 yields

$$\begin{aligned} (b, c) &= (b - cq_1, c) = (r_1, c) \\ &= (c - r_1q_2, r_1) = (r_2, r_1) \\ &= (r_1 - r_2q_3, r_2) = (r_3, r_2) \\ &= \cdots = (r_j, r_{j-1}) = (r_j, 0) = r_j. \end{aligned}$$

This conclusion of the theorem follows at once. □

Observation 2.15. One can apply the Euclidean Algorithm to obtain integral solutions (x, y) to linear equations of the shape $bx + cy = (b, c)$ by "reversing" the application of the algorithm. In general, one can apply this method to solve the equation $bx + cy = k$ whenever $(b, c) \mid k$. (Why? Convince yourself that this is the case.)

Proof. Using the notation employed in the statement of the Euclidean Algorithm, one finds that r_1 is a linear combination of b and c, and then that r_2 is a linear combination of c and r_1, and hence of b and c, and that r_3 is a linear combination of r_1 and r_2, and hence of b and c, and so on. In this way, we see that every remainder r_i that occurs in the algorithm is itself a linear combination of b and c, and the desired conclusion follows. □

Example 2.16. Determine the greatest common divisor of 2016 and 323, and find integers x and y with $2016x + 323y = (2016, 323)$.

Proof. Applying the Euclidean Algorithm, we obtain

$$\begin{aligned} 2016 &= 323 \cdot 6 + 78 \\ 323 &= 78 \cdot 4 + 11 \\ 78 &= 11 \cdot 7 + 1 \\ 11 &= 1 \cdot 11, \end{aligned}$$

and so $(2016, 323) = 1$. Reversing this application of the Euclidean Algorithm, we find that

$$\begin{aligned} 1 &= 78 - 11 \cdot 7 \\ &= 78 - (323 - 78 \cdot 4) \cdot 7 = 78 \cdot 29 - 323 \cdot 7 \\ &= (2016 - 323 \cdot 6) \cdot 29 - 323 \cdot 7 = 2016 \cdot 29 - 323 \cdot 181. \end{aligned}$$

Thus we see that the equation $2016x + 323y = 1$ has the solution $(x, y) = (29, -181)$. □

Note 2.17. One can obtain integral solutions to linear equations in more variables by breaking the equation down into subequations of two variables each. In order to illustrate the strategy, consider the equation $18x+39y+77z = 1$. One can verify easily that $(18, 39) = 3$, and so the equation $18x + 39y = 3$ possesses an integral solution, say $18x_0+39y_0 = 3$, which may be found via the Euclidean Algorithm. Now substitute this solution into the original equation with an additional parameter, and solve the resulting equation. We obtain the equation $3l + 77z = 1$. Since $(3, 77) = 1$, the latter equation has an integral solution $(l, z) = (l_0, z_0)$, say, which may be found via the Euclidean Algorithm. A solution of the original equation is then given by $(x, y, z) = (l_0x_0, l_0y_0, z_0)$.

We finish this section by introducing the concept of least common multiples.

Definition 2.18. (i) Integers $a_1, \dots, a_n$ are said to have a *common multiple* b when $a_i \mid b$ for $1 \leqslant i \leqslant n$.
(ii) The *least common multiple* of the integers $a_1, \dots, a_n$ is the smallest positive common multiple of these integers, which we denote by $[a_1, \dots, a_n]$ or $\operatorname{lcm}(a_1, \dots, a_n)$.

Theorem 2.19. *(i) If m is a positive integer, then $[ma, mb] = m[a, b]$.*
(ii) One has $[a, b](a, b) = |ab|$.

Proof. First consider the assertion of part (i) of the theorem. Let $D = [ma, mb]$ and $d = [a, b]$. Then md is a multiple of both ma and mb, so that $md \geqslant D$. Also, D is a multiple of both ma and mb, so that D/m is a multiple of both a and b. Then $D/m \geqslant d$. We have therefore shown that $md \leqslant D \leqslant md$, whence $D = md$. This establishes part (i) of the theorem.

Now consider part (ii). Suppose first that $(a, b) = 1$. There is no loss of generality in supposing that $a > 0$ and $b > 0$. Write $[a, b] = ma$, with $b \mid ma$. Since $(a, b) = 1$, it follows from Theorem 2.13 that $b \mid m$, whence $b \leqslant m$. Then $ba \leqslant ma$. But $ba \geqslant [a, b] = ma$. We therefore conclude that $ab = [a, b]$, and since $(a, b) = 1$, this yields the desired conclusion $(a, b)[a, b] = |ab|$.

Turning to the general case, put $g = (a, b)$ and $a' = a/g$, $b' = b/g$. Then $(a', b') = (a, b)/g = 1$ and $[a', b'] = [a, b]/g$, so by the above,

$$\frac{(a, b)[a, b]}{g^2} = (a', b')[a', b'] = |a'b'| = \frac{|ab|}{g^2}.$$

The desired identity follows on multiplying by g^2. □

Theorem 2.20. *Suppose that $b_1, \ldots, b_n$ are integers and that $k = [b_1, \ldots, b_n]$. Then the set of all common multiples of $b_1, \ldots, b_n$ is given by $\{km : m \in \mathbb{Z}\}$.*

Proof. Exercise. □