MATHEMATICAL.
OLYMPIAD.CH
MATHEMATIK-OLYMPIADE
OLYMPIADES DE MATHÉMATIQUES
OLIMPIADI DELLA MATEMATICA

Number Theory I

Thomas Huber

Contents

1 Divisibility

In this script, you may assume that a and b denote integers. We say that a is *divisible* by b, or that b is a *divisor* of a, if there exists a $k \in \mathbb{Z}$ such that $a = kb$. Formally, we write this as: $b \,|\, a$. Note that every integer n is divisible by ± 1 and $\pm n$, and that any integer is a divisor of 0. Henceforth, unless specified otherwise, when considering the *divisors* of a positive number $a > 0$ we refer to only the set of its positive divisors.
A number $p \in N$ is called *prime* or a *prime number* if $p > 1$, and if its only divisors are p and 1.

Some simple but important properties:

- $a \,|\, b$ and $b \,|\, c \implies a \,|\, c$
- $a \,|\, b_1, \ldots, a \,|\, b_n$, then for arbitrary integers $c_1, \ldots, c_n$

$$a \,|\, \sum_{i=1}^{n} b_i c_i.$$

- $a \,|\, b$ and $c \,|\, \mathrm{d} \implies ac \,|\, bd$
- p prime and $p \,|\, ab \implies p \,|\, a$ or $p \,|\, b$
- $a \in \mathbb{N}$, $b \in \mathbb{Z}$ and $a \,|\, b \implies b = 0$ or $a \leq |b|$

Example 1 *Find all the natural numbers x, y such that*

$$x^2 - y! = 2001.$$

Solution. We notice that 2001 is divisible by 3 but not by 9. If $y \geq 3$, then $y!$ is divisible by 3, and so must x as well, in which case x^2 is divisible by 9. Now if $y \geq 6$, $y!$ would also be divisible by 9, and so would 2001, which isn't the case. Thus we're left with the following possibilities: $y = 1, 2, 3, 4, 5$. Testing all the cases, we find that the only solution is $(x, y) = (45, 4)$. □

If you have two integers, you can always do a division with remainder. That is:

Proposition 1.1 (Division with remainder) *Let a, b be integers with $b > 0$. Then there exist two unique integers q and r with $0 \leq r < b$, such that*

$$a = qb + r.$$

r is called the remainder *of the division, and we have that $r = 0$ if and only if $b \mid a$.*

One of the most important properties in number theory is the fact that any natural number can be uniquely written as a product of primes:

Theorem 1.2 (Unique prime factorisation) *Let a be a natural number. Then there exist distinct prime numbers $p_1, p_2, \ldots, p_r$ and natural numbers $n_1, n_2, \ldots, n_r$ such that*

$$a = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}.$$

The p_i and n_i depend only on a.

This theorem can be proven by induction, using the division with remainder, but we won't go over the details in this script. The case $a = 1$ corresponds to the empty product on the right-hand side of the equation, i.e. there are no prime factors and $r = 0$. This theorem has many important consequences, of which we will mention two.

Remark *Let $a = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$ be the prime factor decomposition of the natural number a. Then we have that:*

- *a has exactly $(n_1 + 1)(n_2 + 1) \cdots (n_r + 1)$ distinct positive divisors.*
- *a is the m-th power of a natural number if and only if all the exponents n_k are divisible by m.*

The following application of the theorem is one of Euclid's well known results:

Proposition 1.3 *There are infinitely many primes.*

Proof. Suppose that there is a finite number of primes $p_1, p_2, \ldots, p_n$. Let's consider the number $N = p_1 p_2 \cdots p_n + 1$. As $N > 1$, by the theorem 1.2 there exists a prime divisor q of N. But none of the primes p_k divide N, because otherwise we would get that $p_k \mid 1$, which is absurd. Thus q is a prime different from $p_1, p_2, \ldots, p_n$. Contradiction. □

2 GCD and LCM

For two natural numbers a, b, the $\gcd(a, b)$ is the *greatest common divisor* (also sometimes referred to as *greatest common factor*) of a and b; in other words, the largest positive integer that is a divisor both of a and of b. The $\operatorname{lcm}(a, b)$ is the *least/lowest common multiple*, i.e. the smallest positive number that has both a and b as one if its divisors. The GCD and LCM of more than two numbers are defined in a similar way; in which case, we use the abbreviations $(a_1, a_2, \ldots, a_n)$ for the GCD, and $[a_1, a_2, \ldots, a_n]$ for the LCM. The GCD can be characterised by the following equivalences:

(a) $c = \gcd(a, b)$

(b) $c > 0$ is a divisor of both a and of b, and for all positive numbers x, we have

$$x \mid a, x \mid b \quad \Longrightarrow \quad x \mid c.$$

Likewise for the LCM. If $\gcd(a, b) = 1$, then a and b are said to be *coprime*. We have the following properties:

- $\gcd(a, b) = \gcd(b, a)$
- $\gcd(a, b, c) = \gcd(\gcd(a, b), c)$
- $c \mid ab$ and $\gcd(a, c) = 1 \quad \Longrightarrow \quad c \mid b$
- $a \mid c$, $b \mid c$ and $\gcd(a, b) = 1 \quad \Longrightarrow \quad ab \mid c$
- If $d = \gcd(a, b)$, then there exist two coprime integers x and y such that $a = xd$ and $b = yd$. Furthermore, we get that $\operatorname{lcm}(a, b) = xyd$ (see theorem 2.1).
- If a, b are two coprime natural numbers such that ab is an m-th power, then a and b are also both m-th powers.

Using unique prime factorisation, we can calculate the GCD and the LCM explicitly:

Proposition 2.1 *Let $a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$ and $b = p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r}$ be the unique prime factorisations of a and b, with distinct primes p_k and exponents $\alpha_k, \text{and} \beta_k \geq 0$. Then we have*

$$\gcd(a, b) = p_1^{\min\{\alpha_1, \beta_1\}} p_2^{\min\{\alpha_2, \beta_2\}} \cdots p_r^{\min\{\alpha_r, \beta_r\}}$$
$$\operatorname{lcm}(a, b) = p_1^{\max\{\alpha_1, \beta_1\}} p_2^{\max\{\alpha_2, \beta_2\}} \cdots p_r^{\max\{\alpha_r, \beta_r\}}$$

Furthermore, knowing the formula $\min\{x, y\} + \max\{x, y\} = x + y$ we can immediately conclude that

$$\gcd(a, b) \cdot \operatorname{lcm}(a, b) = ab.$$

Example 2 *(Russia 95) Let m and n be two natural numbers with*

$$\gcd(m, n) + \operatorname{lcm}(m, n) = m + n.$$

Show that one number divides the other.

Solution. Let d be the greatest common divisor of m and n, and let's write $m = ad, n = bd$. Then we have that $\operatorname{lcm}(m, n) = abd$ by the theorem 2.1, and the equation becomes $d + abd = ad + bd$, or preferably $d(ab - a - b + 1) = 0$. We factor the left-hand side and find $d(a - 1)(b - 1) = 0$, which means that $a = 1$ or $b = 1$. In the first case, it follows that $m = d$, so $m \mid n$. Likewise, in the second case we get $n \mid m$. □

Theoretically, we can always calculate the gcd using the formulas of the theorem 2.1. Unfortunately it is not always easy to factor very large numbers. Luckily, there is a very simple and efficient algorithm for calculating the gcd, EUCLID*'s algorithm*. It is based on the following proposition:

Proposition 2.2 *For all integers a, b and n we have:*

$$(a, b) = (a, b + na). \tag{1}$$

Proof. We actually only have to show this for the case $n = \pm 1$, as the general case follows by iterating. If c is a common divisor of a and b, then c also divides $b \pm a$, so $(a, b) \mid (a, b \pm a)$. Conversely, let c be a common divisor of a and $b + a$, respectively $b - a$. Thus c also divides $(b + a) - a = b$, respectively $(b - a) + a = b$. Therefore we have $(a, b \pm a) \mid (a, b)$. □

To illustrate, we will calculate $(2541, 1092)$ by applying equation (1) until we get the result:

$$\begin{aligned}(2541, 1092) &= (2541 - 2 \cdot 1092, 1092) = (357, 1092)\\ &= (1092 - 3 \cdot 357, 357) = (21, 357)\\ &= (357 - 17 \cdot 21, 21) = (0, 21) = 21.\end{aligned}$$

Visibly, the idea is to continue the calculations with the smallest number, and the remainder of the largest number divided by the smallest. This is formalised in Euclid's algorithm:

Algorithm 2.3 (EUCLID) *To calculate (a, b) for $a, b \geq 0$.*

1. *Let $a_1 = \max\{a, b\}$, $a_2 = \min\{a, b\}$ and $n = 2$.*
2. *Let $a_{n-1} = q_n a_n + a_{n+1}$ with $0 \leq a_{n+1} < a_n$ (division with remainder).*
3. *If $a_{n+1} = 0$, then we have $(a, b) = a_n$, otherwise we increase n by 1 and return to step 2.*

The validity of this algorithm follows directly from formula (1). For our example, we had essentially done these calculations:

$$\begin{aligned}2541 &= 2 \cdot 1092 + 357\\ 1092 &= 3 \cdot 357 + 21\\ 357 &= 17 \cdot 21 + 0.\end{aligned}$$

Since the remainder in the last line is 0, we get $(2541, 1092) = 21$.

Proposition 2.4 (BEZOUT) *If a, b are coprime, then there exist integers x, y such that*

$$xa + yb = 1.$$

More generally: if $d = \gcd(a, b)$, then there exist integers x, y such that

$$xa + yb = d.$$

Proof. This follows directly from Euclid's algorithm, as in the second to last line, we get an equation for $\gcd(a, b) = a_n$. By substituting the expression for a_n into the equation on the $(n-1)$-th line and iterating this procedure all the way up, we eventually get an equation of the form: $\gcd(a, b) = xa + yb$. □

Reusing our example above:

$$\begin{aligned} 21 &= 1 \cdot 1092 - 3 \cdot 357 \\ &= 1 \cdot 1092 - 3(2541 - 2 \cdot 1092) \\ &= (-3) \cdot 2541 + 7 \cdot 1092. \end{aligned}$$

Now let's consider the linear Diophantine equation in two variables.

Proposition 2.5 *Let a, b, c be integers. The equation*

$$ax + by = c$$

has a solution (x, y) with $x, y \in \mathbb{Z}$ if and only if $d = \gcd(a, b) \,|\, c$. In that case, if (x_0, y_0) is a solution, then the set of all solutions is given by

$$(x, y) = \left(x_0 + k \cdot \frac{b}{d}\,,\; y_0 - k \cdot \frac{a}{d} \right), \qquad k \in \mathbb{Z}.$$

Proof. Assume that (x_0, y_0) is a solution. Then d divides the left-hand term, and so it also divides c. On the other hand, if $d \,|\, c$, the existence of a solution (x_0, y_0) follows directly from Bezout's theorem. Let (x, y) be another solution. Then $a(x - x_0) + b(y - y_0) = c - c = 0$, which gives

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

Since $\frac{a}{d}$ and $\frac{b}{d}$ are coprime, $(x - x_0)$ is divisible by $\frac{b}{d}$ and $(y - y_0)$ by $\frac{a}{d}$. It follows that all solutions are of the given form. By introducing these values in the equation, we show that these are indeed solutions. □

3 Estimations

Estimations are a very important method for solving problems in number theory. They can often help us reduce a problem to testing only for a few special cases, or make a problem more accessible. When estimating, we want to compare the growth of certain quantities in an equation. Here we will present some situations where this method applies.

This first example uses divisibility arguments:

Example 3 *Find all natural numbers n with $n^2 + 11 \mid n^3 + 13$.*

Solution. $n^2 + 11$ divides $n^3 + 13$, so $n^2 + 11$ also must divide $n(n^2 + 11) - (n^3 + 13) = 11n - 13$. Clearly $n = 1$ isn't a solution. When $n \geq 2$ we have that $11n - 13 > 0$, and since this number must also be divisible by $n^2 + 11$, we have

$$n^2 + 11 \leq 11n - 13.$$

This is our estimation. Since the left-hand side is quadratic and the right-hand side is linear, this inequation can only be satisfied for small values of n. Rearranging, we get $n^2 - 11n + 24 = n(n - 11) + 24 \leq 0$. But when $n \geq 12$, we get that $n(n - 11) + 24 \geq 12 \cdot 1 + 24 > 0$, which means that $n \leq 11$. Testing all the cases, we find that the solutions are $n = 3$ and $n = 8$. □

The core observation here was that $a \mid b$ and $b > 0$ implies that $|a| \leq |b|$. Keep this principle in mind as it can come in handy any time, even with IMO problems!

Our second example here uses the fact that between two *consecutive* squares (or any n-th power), there is no other square. This can be useful if you're dealing with an expression that doesn't look that far off from being a square, but that you know is a square. This idea might sound rather trivial, but it can be surprisingly effective.

Example 4 *(Germany 95) Find all pairs of non-negative integers (x, y) that satisfy the following equation:*

$$x^3 + 8x^2 - 6x + 8 = y^3.$$

Solution. We can see that the left-hand side is a 3-rd power (in this case y^3), but at the same time it also is quite close to x^3. Let's explore the 3-rd powers close to x:

$$(x + 2)^3 = x^3 + 6x^2 + 12x + 8,$$
$$(x + 3)^3 = x^3 + 9x^2 + 27x + 27.$$

Considering the coefficients of x^2 in both equations, we can see that the first term appears to be smaller and the second term larger than the left-hand side of our original equation. Let's see when this is the case:

$$(x + 2)^3 < x^3 + 8x^2 - 6x + 8 \Leftrightarrow 2x^2 - 18x > 0 \Leftrightarrow x > 9,$$
$$(x + 3)^3 > x^3 + 8x^2 - 6x + 8 \Leftrightarrow x^2 + 33x + 15 > 0 \quad \text{true for all } x \geq 0.$$

When $x > 9$, we have that our expression lies between two 3-rd powers, contradiction. Therefore $x \leq 9$. Testing all the cases, we find that the solutions are $(0, 2)$ and $(9, 11)$. $\square$

Now we'll present a third estimation technique that can help when dealing with the gcd:

Proposition 3.1 *Let a, b, c be three integers. Then $(a, bc) \leq (a, b) \cdot (a, c)$, with equality when b, c are coprime. (There are other equality cases but it is easier to deal with them on a case-by-case basis rather than try find a general expression.)*

Proof. Let $d = (a, bc)$. There exist two (not necessarily unique) integers d_1, d_2 such that $d = d_1 d_2$ and $d_1 \mid b, d_2 \mid c$. Furthermore, we also have that $d_1 \mid a$ and $d_2 \mid a$ since $d_1 d_2 = d \mid a$. This gives us

$$(a, bc) = d = d_1 d_2 \leq (a, b) \cdot (a, c).$$

This proves the first part of the proposition. For the second part, suppose that b, c are coprime. In which case it follows that $d_1 = (a, b)$ and $d_2 = (a, c)$ are also coprime, since if k is a common divisor of d_1 and d_2, then k is also a common divisor of b and c, hence $k = 1$. Thus since d_1 and d_2 are coprime and are both divisors of a, their product $d_1 d_2$ also divides a. Clearly $d_1 d_2$ divides bc, so $d_1 d_2 \leq (a, bc) \leq (a, b) \cdot (a, c) = d_1 d_2$ and we get the equality. $\square$

In the case of an equality, we notice that for example $(16, 2 \cdot 4) = (16, 2) \cdot (16, 4)$ even though 2 and 4 aren't coprime, and yet $(4, 2 \cdot 4) = 4 < 8 = (4, 2) \cdot (4, 4)$.

This inequality is particularly useful when we want to prove that two terms are coprime. If one of the two terms is a product, then we can use this equality to break down our calculation into several simpler calculations, as we'll see in the following example.

Example 5 *Let m, n be two coprime integers. Then mn and $m + n$ are also coprime.*

Solution. Thanks to the inequality we've just shown, we get $(mn, m + n) \leq (m, m + n) \cdot (n, m + n)$. Thus, if we're able to prove that $(m, m + n) = (n, m + n) = 1$, then we're done. Now we can use Euclid's algorithm to find that $(m, m + n) = (m, n) = 1$. We do the same thing for the other term, and this finishes off the proof.

$\square$