Divisibility and prime numbers

Definition

We say that an integer x *divides* an integer y and write $x \mid y$ if there is an integer k such that $y = kx$. Otherwise we say that x *does not divide* y and write $x \nmid y$.
If $x \mid y$ we also say that y is *divisible* by x; if $x \nmid y$ we say that y is not divisible by x.

Example

$3 \mid 3$, $3 \mid 6$, $3 \mid 9$, and $3 \mid 12$, but $3 \nmid 4$, $3 \nmid 5$, $3 \nmid 7$.
$4 \mid 4$, $4 \mid 8$, $4 \nmid 9$, $4 \nmid 10$.

Definition

Let m be an integer. A positive integer d such that $d \mid m$ is called a (positive) *divisor* of m.

Example

- 1 is a divisor of every integer;
- If $m > 1$ is an integer then m has at least two different divisors: 1 and m.
- Divisors of 6 are 1, 2, 3 and 6;
- Divisors of 25 are 1, 5, 25
- Divisors of 28 are 1, 2, 4, 7, 14, 28.
- An integer m such that m is equal to the sum of all its divisors (except m, of course) is called *perfect*. Only 49 even perfect numbers are known presently (the first 4, namely 6, 28, 496 and $8,128$, were already known about 2,000 years ago; the next one, $33,550,336$, was only discovered in the 15th century). No odd perfect numbers are known.

Some properties of divisibility

- $x \mid 0$ for all $x \in \mathbb{Z}$, but $0 \nmid x$ for all $x \in \mathbb{Z}$ (except $x = 0$).
- $1 \mid x$ for all $x \in \mathbb{Z}$ but $x \mid 1$ implies that $x = 1$ or $x = -1$
- If $x \mid y$ and $y \mid x$ then $y = x$ or $y = -x$
- If $x \mid y$ and $y \mid z$ then $x \mid z$.
- If $2 \mid n$ then we say that n is *even*. If $2 \nmid n$ then we say that n is *odd*.
- If $n \mid x$ and $n \mid y$ then $n \mid ax + by$ for any integers a and b.
- If $m, n > 0$ and $n \mid m$ then $n \leq m$.
  Indeed, if $n \mid m$ then $m = kn$ for some $k \in \mathbb{Z}$. Since $m, n > 0$ we also have $k > 0$. Then $kn = n + (k-1)n \geq n$.

Prime numbers

Definition

An integer $p > 1$ is called a *prime* (or a *prime number*) if it has *precisely two different divisors*, namely 1 and p.
An integer $m > 1$ which is not a prime is called a *composite number*

Important observation

A positive integer m is a *composite number* if and only if m can be written as $m = ab$ where $a, b \in \mathbb{N}$ and $1 < a, b < m$.

Example

2, 3, 5 and 7 are primes; $4 = 2 \cdot 2$, $6 = 3 \cdot 2$, $9 = 3 \cdot 3$, $72 = 9 \cdot 8$ and $3,341 = 257 \cdot 13$ are composite numbers.
2 is the only *even* prime number.

"There are two facts about the distribution of prime numbers. The first is that, [they are] the most arbitrary and ornery objects studied by mathematicians: they grow like weeds among the natural numbers, seeming to obey no other law than that of chance, and nobody can predict where the next one will sprout. The second fact is even more astonishing, for it states just the opposite: that the prime numbers exhibit stunning regularity, that there are laws governing their behavior, and that they obey these laws with almost military precision."

— Don Zagier

Theorem

Every integer which is greater than 1 is a product of finitely many primes and, in particular, is divisible by a prime.

We regard one prime as a product (with just one factor); we also use the convention that the empty product equals 1.

Proof.

We use strong induction. We want to prove that if $m \in \mathbb{Z}$ and $m > 1$ then m is a product of finitely many primes.
If $m = 2$ then m is a prime and there is nothing to do.
Suppose that $m > 2$ and every integer k such that $2 \leq k < m$ is a product of primes. We want to show that then m is a product of primes.
If m is a prime then we are done.
If m is not a prime then $m = kn$ where $1 < k, n < m$. Then k and n are products of finitely many primes by the induction hypothesis and therefore m is also a product of finitely many primes. $\square$

The fundamental theorem of arithmetic

Theorem

Let m be an integer and assume that $m > 1$. Then m can be written, as a product of finitely many prime numbers in a unique way, up to a permutation.
That is, $m = p_1 \cdots p_r$ where $r \geq 1$ and $p_1 \leq p_2 \leq \cdots \leq p_r$ are primes, and if $m = q_1 \cdots q_s$ where $s \geq 1$ and $q_1 \leq q_2 \leq \cdots \leq q_s$ are primes then $s = r$ and $q_i = p_i$, $1 \leq i \leq r$.

Example

$6 = 2 \cdot 3$, $8 = 2^3$, 7 is a prime, $9 = 3^2$, $10 = 2 \cdot 5$, 11 is a prime,
$12 = 2^2 \cdot 3$, 13 is a prime, $14 = 2 \cdot 7$, $15 = 3 \cdot 5$, $16 = 2^4$, 17 is a prime,
$18 = 2 \cdot 3^2$, 19 is a prime, $20 = 2^2 \cdot 5$, ...

Example (This factorization was found in December, 2009)

$$
\begin{aligned}
&1230186684530117755130494958384962720772853569595334792197322\\
&4521517264005072636575187452021997864693899564749427740638459\\
&2519255732630345373154826850791702612214291346167042921431160\\
&222124047927473779408066535141959745985690214341 3\\
&= 33478071698956898786044169848212690817704794983713768568912\\
&\quad 43138898288379387800228761471165253174308773781446799948 9\\
&\times 36746043666799590428244633799627952632279158164343087642676\\
&\quad 03228381573966651127923337341714339681027009279873630891 7
\end{aligned}
$$

(see http://www.isilonsystems.org/emc-plus/rsa-labs/historical/rsa-768-factored.htm for more information)

How many primes are there?

Theorem (Euclid)

There are infinitely many prime numbers.

Proof.

Suppose that there are only finitely many primes, say $p_1, \ldots, p_n$ for some n. Let $N = p_1 \cdots p_n + 1$. Then $N \in \mathbb{N}$, $N > 1$ and so N is divisible by a prime by our previous theorem. Since all primes are listed, N is divisible by p_i for some $1 \leq i \leq n$. Since p_i also divides the product $p_1 \cdots p_n$, we conclude that $p_i \mid N - p_1 \cdots p_n = 1$. So, $p_i \leq 1$, which contradicts the definition of a prime. $\square$

A naive sieve algorithm for finding all primes $\leq N$

input : N

initialization: $L := \{2, 3, \ldots, N\}$;

for $2 \leq i \leq N$ **do**

| $L := L \setminus \{ki : 2 \leq k \leq N/i\}$;

end

output: L

Sieve of Eratosthenes (κόσκινον Ερατοσθένους)

An efficient algorithm for finding all primes $\leq N$.

input : N
initialization: $P := \emptyset$; $p := 2$; $L := \{2, 3, \ldots, N\}$;
while $p^2 \leq N$ **do**
  $L := L \setminus \{kp : 1 \leq k \leq N/p\}$; $P := P \cup \{p\}$; $p := \min L$;
end
output: $P \cup L$

Example: Find all primes less than 140.

Initialization: $P = \emptyset$, $p = 2$

$$
\begin{array}{llllllllll}
L = \{ & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10, & \\
& 11, & 12, & 13, & 16, & 15, & 16, & 17, & 18, & 19, & 20, \\
& 21, & 22, & 23, & 26, & 25, & 26, & 27, & 28, & 29, & 30, \\
& 31, & 32, & 33, & 36, & 35, & 36, & 37, & 38, & 39, & 40, \\
& 41, & 42, & 43, & 46, & 45, & 46, & 47, & 48, & 49, & 50, \\
& 51, & 52, & 53, & 56, & 55, & 56, & 57, & 58, & 59, & 60, \\
& 61, & 62, & 63, & 66, & 65, & 66, & 67, & 68, & 69, & 70, \\
& 71, & 72, & 73, & 76, & 75, & 76, & 77, & 78, & 79, & 80, \\
& 81, & 82, & 83, & 86, & 85, & 86, & 87, & 88, & 89, & 90, \\
& 91, & 92, & 93, & 96, & 95, & 96, & 97, & 98, & 99, & 100, \\
& 101, & 102, & 103, & 104, & 105, & 106, & 107, & 108, & 109, & 110, \\
& 111, & 112, & 113, & 114, & 115, & 116, & 117, & 118, & 119, & 120, \\
& 121, & 122, & 123, & 124, & 125, & 126, & 127, & 128, & 129, & 130 \\
& 131, & 132, & 133, & 134, & 135, & 136, & 137, & 138, & 139, & 140\}
\end{array}
$$

Example: Find all primes less than 140.

First iteration: remove all multiples of 2

$$
\begin{array}{lcccccccccc}
L = \{ & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10, & \\
 & 11, & 12, & 13, & 16, & 15, & 16, & 17, & 18, & 19, & 20, \\
 & 21, & 22, & 23, & 26, & 25, & 26, & 27, & 28, & 29, & 30, \\
 & 31, & 32, & 33, & 36, & 35, & 36, & 37, & 38, & 39, & 40, \\
 & 41, & 42, & 43, & 46, & 45, & 46, & 47, & 48, & 49, & 50, \\
 & 51, & 52, & 53, & 56, & 55, & 56, & 57, & 58, & 59, & 60, \\
 & 61, & 62, & 63, & 66, & 65, & 66, & 67, & 68, & 69, & 70, \\
 & 71, & 72, & 73, & 76, & 75, & 76, & 77, & 78, & 79, & 80, \\
 & 81, & 82, & 83, & 86, & 85, & 86, & 87, & 88, & 89, & 90, \\
 & 91, & 92, & 93, & 96, & 95, & 96, & 97, & 98, & 99, & 100, \\
 & 101, & 102, & 103, & 104, & 105, & 106, & 107, & 108, & 109, & 110, \\
 & 111, & 112, & 113, & 114, & 115, & 116, & 117, & 118, & 119, & 120, \\
 & 121, & 122, & 123, & 124, & 125, & 126, & 127, & 128, & 129, & 130 \\
 & 131, & 132, & 133, & 134, & 135, & 136, & 137, & 138, & 139, & 140\}
\end{array}
$$

Example: Find all prime numbers less than $N = 140$

After first iteration: $P = \{2\}$, $p = \min L = 3$

$$
\begin{aligned}
L = \{ \quad & 3, && 5, && 7, && 9, && 11, \\
& 13, && 15, && 17, && 19, && 21, \\
& 23, && 25, && 27, && 29, && 31, \\
& 33, && 35, && 37, && 39, && 41, \\
& 43, && 45, && 47, && 49, && 51, \\
& 53, && 55, && 57, && 59, && 61, \\
& 63, && 65, && 67, && 69, && 71, \\
& 73, && 75, && 77, && 79, && 81, \\
& 83, && 85, && 87, && 89, && 91, \\
& 93, && 95, && 97, && 99, && 101, \\
& 103, && 105, && 107, && 109, && 111, \\
& 113, && 115, && 117, && 119, && 121, \\
& 123, && 125, && 127, && 129, && 131, \\
& 133, && 135, && 137, && 139 && \}
\end{aligned}
$$

Second iteration: remove multiples of $p = 3$

$$
\begin{aligned}
L = \{ \quad & 3, && 5, && 7, && 9, && 11, \\
& 13, && 15, && 17, && 19, && 21, \\
& 23, && 25, && 27, && 29, && 31, \\
& 33, && 35, && 37, && 39, && 41, \\
& 43, && 45, && 47, && 49, && 51, \\
& 53, && 55, && 57, && 59, && 61, \\
& 63, && 65, && 67, && 69, && 71, \\
& 73, && 75, && 77, && 79, && 81, \\
& 83, && 85, && 87, && 89, && 91, \\
& 93, && 95, && 97, && 99, && 101, \\
& 103, && 105, && 107, && 109, && 111, \\
& 113, && 115, && 117, && 119, && 121, \\
& 123, && 125, && 127, && 129, && 131, \\
& 133, && 135, && 137, && 139 && \}
\end{aligned}
$$

After second iteration: $P = \{2, 3\}$, $p = \min L = 5$

$$
\begin{aligned}
L = \{ & & 5, & 7, & & 11, \\
& 13, & & 17, & 19, & \\
& 23, & 25, & & 29, & 31, \\
& & 35, & 37, & & 41, \\
& 43, & & 47, & 49, & \\
& 53, & 55, & & 59, & 61, \\
& & 65, & 67, & & 71, \\
& 73, & & 77, & 79, & \\
& 83, & 85, & & 89, & 91, \\
& & 95, & 97, & & 101, \\
& 103, & & 107, & 109, & \\
& 113, & 115, & & 119, & 121, \\
& & 125, & 127, & & 131, \\
& 133, & & 137, & 139 & \}
\end{aligned}
$$

Third iteration: remove multiples of $p = 5$

$$
\begin{aligned}
L = \{ & & 5, & 7, & & 11, \\
& 13, & & 17, & 19, & \\
& 23, & 25, & & 29, & 31, \\
& & 35, & 37, & & 41, \\
& 43, & & 47, & 49, & \\
& 53, & 55, & & 59, & 61, \\
& & 65, & 67, & & 71, \\
& 73, & & 77, & 79, & \\
& 83, & 85, & & 89, & 91, \\
& & 95, & 97, & & 101, \\
& 103, & & 107, & 109, & \\
& 113, & 115, & & 119, & 121, \\
& & 125, & 127, & & 131, \\
& 133, & & 137, & 139 & \}
\end{aligned}
$$

After third iteration: $P = \{2, 3, 5\}$, $p = \min L = 7$

$$
\begin{aligned}
L = \{ & & 7, & & 11, \\
& 13, & 17, & 19, & \\
& 23, & & 29, & 31, \\
& & 37, & & 41, \\
& 43, & 47, & 49, & \\
& 53, & & 59, & 61, \\
& & 67, & & 71, \\
& 73, & 77, & 79, & \\
& 83, & & 89, & 91, \\
& & 97, & & 101, \\
& 103, & 107, & 109, & \\
& 113, & & 119, & 121, \\
& & 127, & & 131, \\
& 133, & 137, & 139 & \}
\end{aligned}
$$

Fourth iteration: remove multiples of $p = 7$

$$
\begin{aligned}
L = \{ & & 7, & & 11, \\
& 13, & 17, & 19, & \\
& 23, & & 29, & 31, \\
& & 37, & & 41, \\
& 43, & 47, & 49, & \\
& 53, & & 59, & 61, \\
& & 67, & & 71, \\
& 73, & 77, & 79, & \\
& 83, & & 89, & 91, \\
& & 97, & & 101, \\
& 103, & 107, & 109, & \\
& 113, & & 119, & 121, \\
& & 127, & & 131, \\
& 133, & 137, & 139 & \}
\end{aligned}
$$

After fourth iteration: $P = \{2, 3, 5, 7\}$, $p = \min L = 11$

$$
\begin{aligned}
L = \{ & & & & 11, \\
& 13, & 17, & 19, & \\
& 23, & & 29, & 31, \\
& & 37, & & 41, \\
& 43, & 47, & & \\
& 53, & & 59, & 61, \\
& & 67, & & 71, \\
& 73, & & 79, & \\
& 83, & & 89, & \\
& & 97, & & 101, \\
& 103, & 107, & 109, & \\
& 113, & & & 121, \\
& & 127, & & 131, \\
& & 137, & 139 & \}
\end{aligned}
$$

5th iteration: remove multiples of $p = 11$

$$
\begin{aligned}
L = \{ & & & & 11, \\
& 13, & 17, & 19, & \\
& 23, & & 29, & 31, \\
& & 37, & & 41, \\
& 43, & 47, & & \\
& 53, & & 59, & 61, \\
& & 67, & & 71, \\
& 73, & & 79, & \\
& 83, & & 89, & \\
& & 97, & & 101, \\
& 103, & 107, & 109, & \\
& 113, & & & 121, \\
& & 127, & & 131, \\
& & 137, & 139 & \}
\end{aligned}
$$

After 5th iteration: $P = \{2, 3, 5, 7, 11\}$, $p = \min L = 13$, the algorithm stops since $13^2 > 140$

$$
\begin{aligned}
P \cup L = \{2,\ & 3,\ && 5,\ && 7,\ && && 11,\\
& 13,\ && && 17,\ && 19,\ &&\\
& 23,\ && && 29,\ && 31,\\
& && && 37,\ && 41,\\
& 43,\ && && 47,\ && &&\\
& 53,\ && && && 59,\ && 61,\\
& && && 67,\ && 71,\\
& 73,\ && && && 79,\ &&\\
& 83,\ && && && 89,\ &&\\
& && && 97,\ && 101,\\
& 103,\ && && 107,\ && 109,\ &&\\
& 113,\ && && && && \\
& && && 127,\ && 131,\\
& && && 137,\ && 139 && \}
\end{aligned}
$$

Answer: prime numbers < 140

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53
59, 61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109,
113, 127, 131, 137, 139

Fermat primes ($p = 2^{2^n} + 1$, $n \geq 0$): 3, 5, 17

Mersenne primes ($p = 2^q - 1$, q prime): 3, 7, 31, 127

Twin primes (p, $p + 2$ are both prime): (3,5), (5,7), (11,13), (17,19), (29,31), (41,43), (59,61), (71,73), (101,103), (107,109), (137,139)

Sophie Germain primes (a prime p is a Sophie Germain prime if $2p + 1$ is also a prime): 2, 3, 5, 11, 23, 29, 41, 53, 83, 89, 113, 131

Some open questions

- The only *Fermat primes* are the five known ones (3, 5, 17, 257 and $65,537$)
- There are infinitely many *Mersenne primes* (and hence *even perfect numbers*, since a positive even number m is perfect if and only if $m = 2^{p-1}(2^p - 1)$ where p and $2^p - 1$ are primes)
- There are infinitely many pairs of *twin primes*
- There are infinitely many *Sophie Germain primes*
- (Goldbach's conjecture, 1742) Every even integer > 2 is a sum of two primes (verified for numbers up to $4 \cdot 10^{18}$). For example, $4 = 2 + 2$, $6 = 3 + 3$, $8 = 3 + 5$, $10 = 3 + 7 = 5 + 5$, $12 = 5 + 7$, $14 = 7 + 7 = 11 + 3$, $16 = 11 + 5 = 13 + 3$, $18 = 11 + 7 = 13 + 5$, $20 = 13 + 7 = 17 + 3$, $22 = 11 + 11 = 17 + 5 = 19 + 3$, $\ldots$

Why does this algorithm work? That is

- why does it find all primes $\leq N$?
- why every number it finds is a prime?

Lemma

Let n be a positive integer, $n > 1$. If n is not a prime then there exists a prime p such that $p \mid n$ and $p^2 \leq n$.

Proof.

We prove a weaker statement: if n is not a prime number then there exists a positive integer k such that $k \mid n$ and $k^2 \leq n$. Indeed, if n is not a prime then $n = ab$ where a, b are integers and $1 < a, b < n$. Suppose that $a^2 > n$ and $b^2 > n$. Then $a^2 b^2 > n^2 = a^2 b^2$, which is absurd. So, either $a^2 \leq n$ or $b^2 \leq n$. □

After each iteration of our algorithm the set L consists of integers not divisible by all primes already found. The smallest of them is a prime.

After we found all primes p such that $p^2 \leq N$ and removed all their multiples, all remaining integers $\leq N$ (our set L) are also primes. Indeed, these are integers $k \leq N$ which are not divisible by any prime p such that $p^2 \leq N$ and, in particular, by any prime p such that $p^2 \leq k$. By our Lemma they cannot be composite numbers.

Finally, suppose that $n \in \{2, \ldots, N\} \setminus (P \cup L)$ that is, n is one of the numbers that are not in the output of our algorithm. Then $p \mid n$ for some $p \in P$ and $n \neq p$. Therefore, n is not a prime.

Largest known primes

$2^{74\,207\,281} - 1$ (22 338 618 digits 2016) Mersenne 49??

$2^{57\,885\,161} - 1$ (17 425 170 digits 2013) Mersenne 48?

$2^{43\,112\,609} - 1$ (12 978 189 digits 2008) Mersenne 47?

$2^{42\,643\,801} - 1$ (12 837 064 digits 2009) Mersenne 46?

$2^{37\,156\,667} - 1$ (11 185 272 digits 2008) Mersenne 45

$2^{32\,582\,657} - 1$ (9 808 358 digits 2006) Mersenne 44

$2^{30\,402\,457} - 1$ (9 152 052 digits 2005) Mersenne 43

49 Mersenne primes are presently known

Largest known non-Mersenne prime: $10\,223 \cdot 2^{31\,172\,165} + 1$ (9 383 761 digits 2016), between Mersenne 43 and Mersenne 44.

Largest known pair of twin primes: $299\,686\,3034\,895 \cdot 2^{1\,290\,000} \pm 1$ (388 342 digits 2016)

For more information see http://primes.utm.edu/largest.html

Euclidean division

Theorem

Let x be an integer and n be a positive integer. Then there exist unique integers q and r such that $x = qn + r$ and $0 \leq r < n$.

Existence.

We first prove the theorem for $x \in \mathbb{N}$. The set $S = \{y \in \mathbb{N} : yn \leq x\}$ contains finitely many elements and so has the maximal element which we denote by q. Let $r = x - qn$. By the choice of q we have $r \geq 0$. If $r \geq n$ then $x = r + qn \geq (q+1)n$ and so $q + 1 \in S$. Since q is the maximal element of S, this is a contradiction. Thus, $0 \leq r < n$.
If $x < 0$ then we can write $-x = q'n + r'$ for some $q' \in \mathbb{N}$ and $0 \leq r' < n$. Then $x = -q'n - r'$. If $r' = 0$ then $x = qn + r$ with $q = -q'$ and $r = 0$ and we are done. If $0 < r < n$ then $0 < n - r < n$ and we can write $x = -q'n - r' = -q'n - r' + n - n = -(q'+1)n + (n - r') = qn + r$, where $q = -(q'+1)$, $r = n - r'$. □

Uniqueness.

Assume that $qn + r = q'n + r'$ for some $q, q' \in \mathbb{Z}$ and $0 \leq r, r' < n$. We may assume, without loss of generality, that $r' \geq r$. Then $(q - q')n = r' - r$ and so $n \mid r' - r$. If $r' > r$ then this implies that $n \leq r' - r$ which is a contradiction since $0 \leq r, r' < n$. Thus, $r' = r$. Then $(q - q')n = 0$ and so $q = q'$. $\square$

Definition

If $x \in \mathbb{Z}$ and n is a positive integer, we call the unique r such that $x = qn + r$, $q \in \mathbb{Z}$, $0 \leq r < n$ the *remainder* of x when divided by n.

Example

- Using Euclidean division, we can write any integer n as $2k + r$ where $k \in \mathbb{Z}$ and $0 \leq r < 2$. Thus, either $n = 2k$ and is even or $n = 2k + 1$ and is odd.
- Using Euclidean division, we can write any integer n as $3k + r$ where $k \in \mathbb{Z}$ and $0 \leq r < 3$. Thus, every integer is of one of 3 forms: $3k$, $3k + 1$ or $3k + 2$, $k \in \mathbb{Z}$.
- Let $n \in \mathbb{N}^+$ be a perfect square. Then $n = m^2$ for some $m \in \mathbb{N}^+$. If m is even then $m = 2k$ and so $n = (2k)^2 = 4k^2$. Thus, $n = 4l$ for some $l \in \mathbb{Z}$ (actually, $l \in \mathbb{N}^+$). If m is odd then $m = 2k + 1$ and so $n = (2k + 1)^2 = 4k^2 + 4k + 1 = 4(k^2 + k) + 1 = 4l + 1$, $l = k^2 + k \in \mathbb{N}^+$. Thus, a perfect square can only have reminders 0 or 1 when divided by 4.

By this example, an integer $n = 4k + 3$ cannot be a sum of two perfect squares. In particular, it is not true that all positive integers are sums of two perfect squares. Likewise, a difference of two perfect squares cannot be of the form $4k + 2$.

Questions

- Let n be an integer. What possible reminders can n^2 have when divided by 5?

Greatest common divisor

Definition

Let x, y be non-zero integers. Their greatest common divisor, denoted $\gcd(x, y)$, is the *maximal positive* integer d such that $d \mid x$ and $d \mid y$.
If $\gcd(x, y) = 1$ we say that x and y are relatively prime or coprime.

Example

$\gcd(15, 18) = 3$, $\gcd(10, 12) = 2$, $\gcd(15, 46) = 1$.

Theorem

Let x, y be non-zero integers and let $d = \gcd(x, y)$. Then d is the minimal positive integer in the set $\langle x, y\rangle := \{ax + by : a, b \in \mathbb{Z}\}$. In particular, $d = ax + by$ for some $a, b \in \mathbb{Z}$ and if $n \mid x, y$ then $n \mid d$.

Proof

First of all, the set $\langle x, y\rangle$ contains $x = 1 \cdot x + 0 \cdot y$ and $-x = (-1) \cdot x + 0 \cdot y$. Since $x \neq 0$, either $x > 0$ or $-x > 0$. Thus, the set $\{z \in \langle x, y\rangle : z > 0\}$ is not empty.
Next, every non-empty set of positive integers has a minimal element. Indeed, let S be such a set and let $n \in S$. Then there are finitely many positive integers m such that $m < n$. Let m_0 be minimal positive integer such that $m_0 \leq m$ and $m \in S$ (it exists because there are finitely many integers with that property). Then for every $x \in S$, $m_0 \leq x$. Thus, $m_0 = \min S$.
From all the above we conclude that the set $\langle x, y\rangle$ indeed contains a minimal positive element, say d'. We want to show that $d' = \gcd(x, y) = d$.

Proof (cont.)

Since $d = \gcd(x, y)$, $d \mid x, y$ and so $d \mid z$ for every $z \in \langle x, y \rangle$.
Thus, $d \mid d'$ In particular, $d \leq d'$.
We will now show that $d' \mid x$ and $d' \mid y$. We only prove that $d' \mid x$, because the argument for $d' \mid y$ is similar.
Write $x = qd' + r$ where $0 \leq r < d'$ and $q \in \mathbb{Z}$. By assumption, $d' \in \langle x, y \rangle$ and so $d' = ax + by$ for some $a, b \in \mathbb{Z}$. Then
$r = x - qd' = x - q(ax + by) = (1 - qa)x + (-b)y$ and so $r \in \langle x, y \rangle$. But $0 \leq r < d'$ and d' is *minimal positive* element of $\langle x, y \rangle$. So, $r = 0$, that is, $d' \mid x$.
So, d' is a common divisor of x and y and thus $d' \leq d$. We already established that $d \leq d'$. This is only possible if $d = d'$.
Since $\gcd(x, y) \in \langle x, y \rangle$, $\gcd(x, y) = ax + by$ for some $a, b \in \mathbb{Z}$. Now, if $n \mid x$ and $n \mid y$ then n divides any element of $\langle x, y \rangle$ and in particular $n \mid \gcd(x, y)$.

Our definition of gcd does not immediately imply that every common divisor of x and y divides $\gcd(x, y)$. Now we proved it.

Algorithm (Euclid) for finding $\gcd(x, y)$

input : x, y, $y \geq x > 0$
initialization: $r_0 := y$, $r_1 := x$, $n := 1$;
while $r_n > 0$ **do**
    $r_{n+1} :=$ the remainder of r_{n-1} when divided by r_n;
    $n := n + 1$;
end
output: r_{n-1}

Example

Example

Find $\gcd(2130, 3024)$

$$
\begin{aligned}
& & r_0 &= 3024 \\
& & r_1 &= 2130 \\
3024 &= 1 \cdot 2130 + 894, & r_2 &= 894 \\
2130 &= 2 \cdot 894 + 342, & r_3 &= 342 \\
894 &= 2 \cdot 342 + 210, & r_4 &= 210 \\
342 &= 1 \cdot 210 + 132, & r_5 &= 132 \\
210 &= 1 \cdot 132 + 78, & r_6 &= 78 \\
132 &= 1 \cdot 78 + 54, & r_7 &= 54 \\
78 &= 1 \cdot 54 + 24, & r_8 &= 24 \\
54 &= 2 \cdot 24 + 6, & r_9 &= 6 \\
24 &= 4 \cdot 6 + 0, & r_{10} &= 0.
\end{aligned}
$$

Why does the algorithm work?

Lemma

Let $k \geq 1$. If $r_k > 0$ then $\gcd(r_{k-1}, r_k) = \gcd(r_k, r_{k+1})$.

Proof.

Let CD_k be the set of common divisors of r_k and r_{k+1}. We claim that $CD_k = CD_{k-1}$ as sets.

By construction of our sequence $r_0, r_1, \ldots$ we have $r_{k+1} = qr_k + r_{k-1}$ for some $q \in \mathbb{Z}$.

Suppose first that $d \in CD_{k-1}$. Then $d \mid r_k$, $d \mid r_{k-1}$ and hence $d \mid qr_k + r_{k-1} = r_{k+1}$. Thus, $d \in CD_k$ and so $CD_{k-1} \subset CD_k$.

On the other hand, if $d' \in CD_k$ then $d' \mid r_k$ and $d' \mid r_{k+1}$ hence $d' \mid r_{k+1} - qr_k = r_{k-1}$ and so $d' \in CD_{k-1}$. Thus, $CD_k \subset CD_{k-1}$. Therefore, $CD_k = CD_{k-1}$.

Now, since CD_k and CD_{k-1} coincide, their maximal elements also coincide. But the maximal element of CD_k is $\gcd(r_k, r_{k+1})$ and the maximal element of CD_{k-1} is $\gcd(r_{k-1}, r_k)$. □

We have $\gcd(x, y) = \gcd(r_0, r_1)$. By our lemma, at each iteration we have $\gcd(x, y) = \gcd(r_{k+1}, r_k)$. Thus, when we exit, $\gcd(r_n, r_{n-1}) = \gcd(x, y)$. But $r_n = 0$ so $\gcd(r_n, r_{n-1}) = r_{n-1}$. Thus, $r_{n-1} = \gcd(x, y)$. The algorithm is correct.

The fundamental property of primes

Theorem (Euclid's Lemma)

Let n, a, b be non-zero integers and suppose that $\gcd(n, a) = 1$ and $n \mid ab$. Then $n \mid b$. In particular, if p is a prime and a, b are non-zero integers then $p \mid ab$ implies that $p \mid a$ or $p \mid b$.

Proof.

Since $\gcd(n, a) = 1$, $1 = xa + yn$ for some $x, y \in \mathbb{Z}$. Then we have $b = b \cdot 1 = b \cdot (xa + yn) = xab + ynb$. Since $n \mid ab$, this implies that $n \mid b$. The first assertion is proven.

If p is a prime then either $\gcd(p, a) = p$ (in which case $p \mid a$) or $\gcd(p, a) = 1$. In the first case we are done. In the second case we use the first assertion to conclude that $p \mid b$. □

Example

We can use this to prove that if p is a prime then $\sqrt{p}$ is not a rational number. More precisely, we can prove that there is no rational number x such that $x^2 = p$.

Indeed, suppose that there is a rational number x such that $x^2 = p$. Write $x = a/b$ with $a, b \in \mathbb{Z}$ and $\gcd(a, b) = 1$. Then $b^2 p = a^2$ and so $p \mid a^2$. By Euclid's Lemma this implies that $p \mid a$. Thus, $a = kp$ for some $k \in \mathbb{Z}$ and so $b^2 p = k^2 p^2$ which implies that $b^2 = k^2 p$. Thus, $p \mid b^2$ and so $p \mid b$ by Euclid's Lemma. Therefore, $p \mid a, b$ and so $\gcd(a, b) \geq p > 1$ which is a contradiction.

A similar argument can be used to prove that if $n > 1$ and p is a prime then $\sqrt[n]{p}$ is not a rational number. That is, there is no rational number y such that $y^n = p$.

Corollary

Suppose that a, b are positive integers and $\gcd(a, b) = 1$. Let n be an integer. Then $ab \mid n$ if and only if $a \mid n$ and $b \mid n$.

Proof.

If $ab \mid n$ then $n = kab = (kb)a = (ka)b$ and so $a \mid n$ and $b \mid n$.
Conversely, suppose that $a \mid n$ and $b \mid n$. Then $n = ak$ for some $k \in \mathbb{Z}$.
Since $b \mid n = ak$ and $\gcd(a, b) = 1$, Euclid's Lemma implies that $b \mid k$.
Thus, $k = bm$ for some $m \in \mathbb{Z}$ and so $n = ak = abm$, whence $ab \mid n$. $\square$

The fundamental theorem of arithmetic and its consequences

Theorem (Unique factorization of integers)

Every integer $n > 1$ can be written uniquely as $n = p_1 \cdots p_r$ where $r \geq 1$ and $p_1 \leq p_2 \leq \cdots \leq p_r$ are primes.

Proof.

We already proved that every integer > 1 is a product of finitely many primes.
To prove uniqueness, we can use strong induction on the *minimal number of factors in our decomposition*.
Given $N > 1$, we can consider the set of all positive integers r such that N can be written as a product of r prime factors. This set is non-empty and so has a minimal element which we will denote by $\ell(N)$.
Suppose that $\ell(N) = 1$. Then N is a prime and so cannot be written as a product of primes in any other way.

Proof (contd.)

Suppose that the claim is proven for all $n > 1$ with $\ell(n) < r$. Let N be such that $\ell(N) = r$ and write $N = p_1 \cdots p_r$ where $p_1 \leq \cdots \leq p_r$ are primes. Suppose that we also have $N = q_1 \cdots q_s$ where $q_1 \leq \cdots \leq q_s$ are primes. Then $s \geq r$ (because r is the *minimal number of primes* needed to write N as a product of primes). In particular, $p_1 \cdot (p_2 \cdots p_r) = q_1 \cdots q_s$ and so $p_1 \mid q_1 \cdots q_s$. Then by Euclid's Lemma, we can find $1 \leq i \leq s$ such that $p_1 \nmid q_j$, $j < i$ and $p_1 \mid q_i$. Since both p_1 and q_i are primes it follows that $p_1 = q_i$. Then

$$M = p_2 \cdots p_r = q_1 \cdots \hat{q}_i \cdots q_s$$

where $\hat{\cdot}$ means omission. But now the induction hypothesis applies to M since $\ell(M) \leq r - 1 < r$. This means that $r - 1 = s - 1$ (and so $r = s$) and that

$$p_j = \begin{cases} q_{j-1} & 2 \leq j \leq i \\ q_j, & i+1 \leq j \leq r. \end{cases}$$

We have $p_1 \leq p_2 = q_1 \leq q_i$. If $q_1 \neq q_i$ (and so $i > 1$ by the choice of i) we conclude that $p_1 < q_i = p_1$ which is absurd. So, $i = 1$ and we have $p_j = q_j$ for all $1 \leq j \leq r$. The inductive step is proven. □

Given a prime p, let $e_p(n)$ be the maximal non-negative integer k such that $p^k \mid n$.

Clearly, $e_p(1) = 0$ for all primes p, $e_p(p^k) = k$ and $e_p(n) = 0$ if $p \nmid n$.

Also if we fix n then $e_p(n) = 0$ for all but finitely many primes.

Example

- $e_p(6) = 0$ if $p \notin \{2, 3\}$ while $e_2(6) = e_3(6) = 1$
- $e_2(50) = 1$, $e_5(50) = 2$, $e_p(50) = 0$ if $p \notin \{2, 5\}$.
- Let $n = 199\,797\,500 = 2^2 \cdot 5^4 \cdot 7^3 \cdot 233$. Then $e_2(n) = 2$, $e_5(n) = 4$, $e_7(n) = 3$, $e_{233}(n) = 1$ and $e_p(n) = 0$ if $p \notin \{2, 5, 7, 233\}$.

The fundamental theorem of arithmetic states that an integer n is uniquely determined by the values $e_p(n)$ for all primes p. We can also extend it to all non-zero integers; then the values of $e_p(n)$ determine n uniquely up to its sign.

Lemma

Let m be a positive integer. Then $d \mid m$ if and only if $e_p(d) \leq e_p(m)$ for all primes p. In particular, the number of divisors of m is

$$\prod_{p \text{ prime}} (e_p(m) + 1).$$

The notation $\prod$ means "product"; we multiply all values $e_p(m) + 1$ for all primes p. Note that only finitely many of them are different from 1 so this makes sense.

Proof.

Let p be a prime and d be a divisor of m. Recall that $a \mid b$, $b \mid c$ implies that $a \mid c$. Thus, if $p^k \mid d$ then $p^k \mid m$ and so $e_p(d) \leq e_p(m)$.
Conversely, suppose that $e_p(d) \leq e_p(m)$ for all primes p. Let n be the unique positive integer such that $e_p(n) = e_p(m) - e_p(d)$ for all prime p. Then $dn = m$ and so $d \mid m$. □

Example

- We have $3025 = 5 \cdot 605 = 5^2 \cdot 121 = 5^2 \cdot 11^2$. Then $e_5(3025) = e_{11}(3025) = 2$ and $e_p(3025) = 0$ if $p \notin \{5, 11\}$. Divisors of 3025 are integers d with $e_p(d) = 0$, $p \notin \{5, 11\}$, $0 \leq e_5(d) \leq 2$ and $0 \leq e_{11}(d) \leq 2$. Thus, 3025 has $(2+1) \cdot (2+1) = 9$ divisors: 1, 5, 11, $25 = 5^2$, $55 = 5 \cdot 11$, $121 = 11^2$, $275 = 5^2 \cdot 11$, $605 = 5 \cdot 11^2$ and 3025.
- $17,199 = 3^3 \cdot 7^2 \cdot 13$. Its divisors are all integers d with $e_p(d) = 0$, $p \notin \{3, 7, 13\}$, $0 \leq e_3(d) \leq 3$, $0 \leq e_7(d) \leq 2$ and $0 \leq e_{13}(d) \leq 1$. it has $(3+1) \cdot (2+1) \cdot (1+1) = 24$ divisors, namely 1, 3, 7, 9, 13, 21, 27, 39, 49, 63, 91, 117, 147, 189, 273, 351, 441, 637, 819, 1323, 1911, 2457, 5733, 17199

Theorem

Let x, y be positive integers. Then $\gcd(x, y)$ is the unique positive integer d with $e_p(d) = \min(e_p(x), e_p(y))$ for all primes p.

Example

As we saw previously, $3024 = 2^4 \cdot 3^3 \cdot 7$, $2130 = 2 \cdot 3 \cdot 5 \cdot 71$.
Let $d = \gcd(3024, 2130)$. Then $e_p(d) = 0$, $p \notin \{2, 3, 5, 7, 71\}$ and $e_2(d) = \min(4, 1) = 1$, $e_3(d) = \min(3, 1) = 1$, $e_5(d) = \min(0, 1) = 0$, $e_7(d) = \min(1, 0) = 0$ and $e_{71}(d) = \min(0, 1) = 0$. Thus, $\gcd(3024, 2130) = 2^1 \cdot 3^1 = 6$.
One may ask why use Euclid's algorithm to find the GCD instead of finding prime factors and taking their minimal *exponents* or just finding all divisors of both numbers and choosing the maximal one. The answer is that finding prime factors of an integer (and hence all its divisors) is *an exponentially hard problem*, while Euclid's algorithm is very fast and efficient.

Proof.

Since $e_p(d) \leq e_p(x)$ for all primes p we have $d \mid x$. Similarly, since $e_p(d) \leq e_p(y)$ for all primes p we have $d \mid y$. Thus, d is a common divisor of x and y.

Now, if $n \mid x$ and $n \mid y$ then $e_p(n) \leq e_p(x)$ and $e_p(n) \leq e_p(y)$ for all prime p. Therefore, $e_p(n) \leq \min(e_p(x), e_p(y)) = e_p(d)$. This means that $n \mid d$ and so $n \leq d$. Thus, d is the greatest common divisor of x and y. □