Chapter 1

Elementary number theory

1.1 Introduction to the subject

Number theory is really old, probably as old as mathematics.

Cryptography, viewed more widely as the art/science of secret messages, is really old, probably as old as written language. And yet, these topics are alive and kicking, with large research communities actively developing new ideas. There are many points of contact between the two subjects, both of historical and current interest. We will embrace both, using cryptographic applications as an excuse to learn some number theory, and using fundamental ideas in number theory as an excuse to explore cryptographic concepts.

Cryptography

The overarching goal of cryptography is simple to state: allow Alice to send a message to Bob in such a way that only Bob can read it, even if Eve manages to intercept the message on the way.

There have been numerous schemes that attempt to reach this goal, achieving some level of secrecy (provable or expected) at some cost (provable or expected).

There is in fact a scheme that achieves provable perfect secrecy, and dates back to 1882:

Example 1.1 (The One-Time Pad)**.** Alice and Bob share a random secret key $K \in \mathbf{F}_2^m$.

Alice wants to send a plaintext message $P \in \mathbf{F}_2^m$. The encoding is done by calculating $C = P \oplus K$, where $\oplus$ is the bitwise operation XOR (exclusive or):

$\oplus$	0	1
0	0	1
1	1	0

(Note that $A \oplus A = \mathbf{0}$ for all A.)

Bob decodes the ciphertext C by calculating $C \oplus K = (P \oplus K) \oplus K = P \oplus (K \oplus K) = P \oplus \mathbf{0} = P$.

For instance, if the key is $K = 01101110$, we could have

$$\begin{aligned} P &= 11010001 \\ C = P \oplus K &= 10111111 \\ P = C \oplus K &= 11010001. \end{aligned}$$

In 1949, Claude Shannon proved that the one-time pad is an unbreakable cryptosystem, and that any system that is unbreakable must feature a random secret key that is as large as the plaintext and is not reused even partially.

So we have a perfectly secure system, but it is rather impractical; moreover, any perfectly secure system is as impractical as the one-time pad.

Given this information-theoretic constraint, cryptography walks the fine line between secrecy and useability.

As a high-stakes applicable subject, the difference between theory and practice in cryptography is significant: implementation details matter tremendously, and many schemes that have conceptual elegance are riddled with implementation issues. This is definitely not a practical cryptography subject. We will be entirely focussed on theoretical aspects and the number theory behind them.

Number theory

Since the topic is thousands of years old, it now covers a whole lot of ground, making it difficult to distil into a concise yet accurate definition. One could say that it is the study of properties of the integers. This sounds fairly restrictive in scope but is actually a vast ongoing process that is unlikely to ever be "done".

We'll illustrate some features of number theoretic work with the famous

Example 1.2 (Pythagorean Triples)**.** A *Pythagorean triple* is a triple $(X, Y, Z) \in \mathbf{Z}^3_{\geqslant 1}$ such that $X^2 + Y^2 = Z^2$.

Given one such (X, Y, Z), we get infinitely many more by taking (aX, aY, aZ) with $a \in \mathbf{Z}_{\geqslant 1}$, so let's restrict to *primitive* triples where X, Y, Z are coprime.

Most people know $(3, 4, 5)$, and fewer of them know $(5, 12, 13)$. Even fewer know $(65, 72, 97)$. How do we find more? (Or even all of them?)

Divide through by Z to get a point on the unit circle

$$C : \ x^2 + y^2 = 1,$$

with $x = \frac{X}{Z}, y = \frac{Y}{Z} \in \mathbf{Q}$. Denote the set of rational points on C by

$$C(\mathbf{Q}) = \{(a, b) \in \mathbf{Q}^2 : \ a^2 + b^2 = 1\}.$$

Any primitive triple (X, Y, Z) gives rise to a point $(X/Z, Y/Z) \in C(\mathbf{Q})$; more precisely, this point lies strictly in the first quadrant, hence on the open arc

$$S = \{(a, b) \in \mathbf{Q}^2 : \ a^2 + b^2 = 1, a > 0, b > 0\}.$$

Conversely, any point $(a, b) \in S$ gives rise to a primitive triple (ad, bd, d), where d is the common denominator of the fractions a and b in lowest terms.

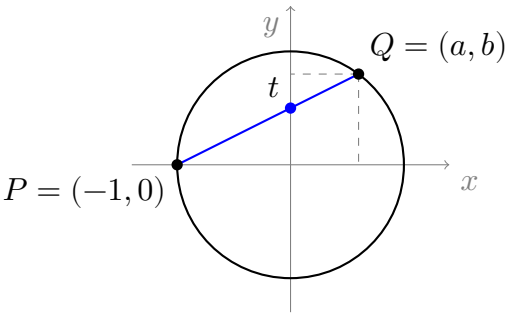
The problem is therefore reduced to: how do we find all the elements of S?

Fix the point $P = (-1, 0) \in C(\mathbf{Q})$.

Given a point $Q = (a, b) \in S$, the straight line joining P and Q has slope

$$t = \frac{b}{a + 1} \in (0, 1) \cap \mathbf{Q}.$$

Conversely, if $t \in (0, 1) \cap \mathbf{Q}$, the line of slope t passing through P intersects the circle in another point $Q = (a, b) \in S$.



We determine the corresponding (a, b) in terms of t: since $b = t(a+1)$ we have

$$1 = a^2 + b^2 = (1+t^2)a^2 + 2t^2 a + t^2,$$

in other words $x = a$ satisfies the quadratic equation

$$x^2 + \frac{2t^2}{1+t^2}\, x + \frac{t^2-1}{1+t^2} = 0.$$

However, so does $x = -1$, since $P = (-1, 0)$ is also on the line. Hence the degree zero term of the quadratic equation is

$$(-1)a = \frac{t^2-1}{1+t^2} \quad \Rightarrow \quad a = \frac{1-t^2}{1+t^2},\ b = t(a+1) = \frac{2t}{1+t^2}.$$

We get a nice parametrisation of primitive Pythagorean triples by writing $t = m/n$ with $0 < m < n$, $\gcd(m, n) = 1$:

$$a = \frac{n^2-m^2}{n^2+m^2}, \quad b = \frac{2mn}{n^2+m^2},$$

so that $(X, Y, Z) = (n^2 - m^2, 2mn, n^2 + m^2)$.

Note that to actually get a primitive triple, we must also require that exactly one of m, n is even (if they are both odd then all components of the triple are even).

Notation and conventions

We write $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ to denote respectively: the natural numbers, the integers, the rational numbers, the real numbers, and the complex numbers.

For us, $\mathbf{N}$ starts at zero:

$$\mathbf{N} = \{0, 1, 2, \dots\}.$$

This does cause some trouble, but, as the saying goes: "wer schön sein will, muss leiden". We will often alleviate the discomfort by using the terminology *positive integer* and the notation

$$\mathbf{Z}_{\geqslant 1} = \{1, 2, \dots\}.$$

Other notation will be introduced as we go.

In terms of cross-referencing among the various materials: Exercise refers to the self-study exercises, Tutorial Question refers to tutorial sheets, Assignment Question refers to assignments, any other type of item refers to the lecture notes.

1.2 Divisibility

A goal that has driven millenia of development in number theory is the solution of *Diophantine equations*; these are equations in one or several variables for which we want to determine the **integer** solutions. They are named after Diophantus of Alexandria, whose work Arithmetica consists of thirteen volumes of arithmetic problems that are solved using algebraic methods.

Despite the long history of such questions, the situation in 1900 was unsatisfactory enough that David Hilbert proposed, as one of his challenge problems:

Hilbert's Tenth Problem. *Find an algorithm to determine whether a given polynomial Diophantine equation with integer coefficients has an integer solution.*

In 1970, Matiyasevich proved that this question is undecidable, that is no such algorithm exists.
The solvability of special types of Diophantine equations can, however, be determined.

Example 1.3. The equation $15x + 9y = 17$ has no integer solutions, since 3 always divides the LHS of this equation, but $3 \nmid 17$.

This is pretty clear on an intuitive level, but let's back up for a second and introduce (or most likely: recall) the relevant terminology.

The starting point is the humble division with remainder:

Theorem 1.4 (Division Algorithm)**.** *For any positive integers $a \geqslant m$, there exist unique integers q (the quotient) and r (the remainder) such that*

$$a = qm + r \qquad \text{and } 0 \leqslant r < m.$$

For the proof of this result, see Exercise 1.1. It uses the

Well-Ordering Property of N. *If $S \subseteq \mathbf{N}$ is a non-empty subset, then S has a minimal element m (that is, $m \in S$ and for all $s \in S$ we have $m \leqslant s$).*

In the special case where $r = 0$, so that $a = qm$, we say that m *divides* a and write $m \mid a$ (with lots of synonyms: m is a *divisor* of a, a is a *multiple* of m, a is *divisible* by m, etc.) We extend the notion of divisibility to all integers a by imposing that $m \mid -a$ if and only if $m \mid a$. If $r > 0$ we write $m \nmid a$.

It is customary to extend these definition to all integers: given $a, m \in \mathbf{Z}$, we say that m *divides* a if there exists $q \in \mathbf{Z}$ such that $a = qm$.

Note that any divisor d of a satisfies $|d| \leqslant |a|$.

Congruences

Suppose now that we have a positive integer m and integers a, b. We say that $a \equiv b \pmod{m}$ (in words: *a is congruent to b modulo m*) if $m \mid a - b$.

Note that, by definition: if r is the remainder of the division of a by m, then $a \equiv r \pmod{m}$.

The notion of congruence plays particularly well with algebraic operations: if $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then

$$a + c \equiv b + d \pmod{m} \qquad \text{and} \qquad ac \equiv bd \pmod{m}.$$

For a proof, see Exercise 1.6.

We can revisit Example 1.3: suppose $15x + 9y = 17$ has an integer solution (x, y), then modulo 3 we get:

$$15x + 9y \equiv 17 \pmod{3}, \qquad \text{so} \qquad 0 \equiv 2 \pmod{3},$$

contradiction.

Greatest common divisor

Given positive integers a and b, the set of common divisors

$$S = \{d : \ d \mid a \text{ and } d \mid b\}$$

is non-empty (since $1 \in S$) and finite, so it has a maximal element called the *greatest common divisor* of a and b, denoted $\gcd(a, b)$.

This is extended to arbitrary integers a and b by setting $\gcd(a, b) = \gcd(|a|, |b|)$.

Lemma 1.5. *Let $a, b \in \mathbf{Z}_{\geqslant 1}$ and let $a = qb + r$ be the result of the Division Algorithm. Then*

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

Proof. It suffices to prove that the set of common divisors of a and b is equal to the set of common divisors of b and r.

Suppose d divides both a and b, then

$$r = a - qb \equiv 0 - 0 \pmod{d},$$

so d divides r.

Conversely, if d divides both b and r, then

$$a = qb + r \equiv 0 + 0 \pmod{d},$$

so d divides a. □

This is key to the *Euclidean Algorithm* for computing $\gcd(a, b)$: apply the Division Algorithm to get $a = qb + r$. If $r > 0$, replace a by b and b by r and repeat. If $r = 0$, then $\gcd(a, b) = b$. Lemma 1.5 ensures that the gcd is the same at each step. Since r starts out as $r < b$ and decreases at each step, after finitely many (at most b) steps it reaches 0 and the algorithm terminates.

We illustrate the process with an example, which will also reveal a useful side effect:

Example 1.6. Compute $\gcd(2178, 1560)$.

$$\begin{aligned}
2178 &= 1 \cdot 1560 + 618 \\
1560 &= 2 \cdot 618 + 324 \\
618 &= 1 \cdot 324 + 294 \\
324 &= 1 \cdot 294 + 30 \\
294 &= 9 \cdot 30 + 24 \\
30 &= 1 \cdot 24 + 6 \\
24 &= 4 \cdot 6 + \mathbf{0}
\end{aligned}$$

We conclude that the gcd is 6.

Observe that we can isolate 6 in the second to last step, then iteratively work our way back up to get an integer linear combination of the numbers we started with:

$$\begin{aligned}
6 &= 30 - 1 \cdot 24 = 30 - (294 - 9 \cdot 30) = 10 \cdot 30 - 294 = 10(324 - 294) - 294 \\
&= 10 \cdot 324 - 11 \cdot 294 = 10 \cdot 324 - 11(618 - 324) = 21 \cdot 324 - 11 \cdot 618 \\
&= 21(1560 - 2 \cdot 618) - 11 \cdot 618 = 21 \cdot 1560 - 53 \cdot 618 = 21 \cdot 1560 - 53(2178 - 1560) \\
&= 74 \cdot 1560 - 53 \cdot 2178.
\end{aligned}$$

Proposition 1.7 (Bézout's Lemma)**.** *Let $a, b \in \mathbf{Z}_{\geqslant 1}$ and let $g = \gcd(a, b)$.*

The equation $ua + vb = d$ is solvable with $u, v \in \mathbf{Z}$ if and only if d is an integer multiple of g.

Proof. In one direction: suppose $d = kg$ with $k \in \mathbf{Z}$. We have seen that the Euclidean Algorithm produces $x, y \in \mathbf{Z}$ such that $xa + yb = g$, so we get $(kx)a + (ky)b = kg = d$.

In the other direction: suppose there exist $u, v \in \mathbf{Z}$ such that $d = ua + vb$. Since $g \mid a$ and $g \mid b$, we conclude that $g \mid ua + vb = d$. □

Primes and factorisation

An integer $p > 1$ is said to be *prime* if it has precisely two divisors: 1 and p. An integer $n > 1$ is said to be *composite* if it is not prime. The integer 1 is neither prime nor composite (it is a *unit*).

We sometimes extend these definitions to negative integers: a negative integer q is said to be prime if the positive integer $-q$ is prime, etc.

Proposition 1.8. *An integer $p > 1$ is prime if and only if: for all integers a and b, if $p \mid ab$ then $p \mid a$ or $p \mid b$.*

Proof. Let $p > 1$ and suppose p is not prime. Then p has some divisor a such that $1 < a < p$. Write $p = ab$, then b is also a divisor such that $1 < b < p$. In particular, $p \nmid a$ and $p \nmid b$, but certainly $p \mid ab$.

Conversely, suppose $p > 1$ is prime and let $a, b \in \mathbf{Z}$ be such that $p \mid ab$. If $p \mid a$, we are done. Otherwise, $p \nmid a$, so $\gcd(a, p) = 1$, so by Bézout's Lemma there exist $u, v \in \mathbf{Z}$ such that $ua + vp = 1$. Multiplying by b we get $uab + vpb = b$, which shows that $p \mid b$. □

Corollary 1.9. *Suppose p is prime and $p \mid q_1^{f_1} \dots q_s^{f_s}$ with $q_1, \dots, q_s$ distinct primes and $f_1, \dots, f_s \in \mathbf{Z}_{\geqslant 1}$. Then there exists $j \in \{1, \dots, s\}$ such that $p = q_j$.*

An important property of multiplication in the integers is the unique factorisation into primes:

Theorem 1.10 (Fundamental Theorem of Arithmetic)**.** *Every integer $n > 1$ can be written uniquely in the form*

$$n = p_1^{e_1} p_2^{e_2} \dots p_r^{e_r} \qquad \text{with } p_1 < p_2 < \dots < p_r \text{ primes, } e_1, e_2, \dots, e_r \in \mathbf{Z}_{\geqslant 1}.$$

Proof. For the existence of factorisations, see Tutorial Question 2.6.

We move on to the uniqueness claim. Suppose it is false, in other words there exist integers $n > 1$ with at least two distinct factorisations. Then the set of such integers is a non-empty subset of $\mathbf{N}$, hence by the Well-Ordering Principle there is a minimal element $n_0 > 1$ with two factorisations

$$n_0 = p_1^{e_1} p_2^{e_2} \dots p_r^{e_r} = q_1^{f_1} q_2^{f_2} \dots q_s^{f_s}$$

where $p_1 < \dots < p_r$ and $q_1 < \dots < q_s$ are primes, $e_1, \dots, e_r, f_1, \dots, f_s \in \mathbf{Z}_{\geqslant 1}$.

Without loss of generality $p_1 \leqslant q_1$ (otherwise interchange the two factorisations to achieve this).

Since $p_1 \mid n_0$, we have that $p_1 \mid q_1^{f_1} \dots q_s^{f_s}$, so by Corollary 1.9 we have $p_1 = q_j$ for some j. But $p_1 \leqslant q_1$, so $p_1 = q_1$. Cancelling one p_1 in the first factorisation and one q_1 in the second, we get

$$\frac{n_0}{p_1} = p_1^{e_1 - 1} p_2^{e_2} \dots p_r^{e_r} = q_1^{f_1 - 1} q_2^{f_2} \dots q_s^{f_s}.$$

Since $(n_0/p_1) < n_0$, either $(n_0/p_1) = 1$ (so $n_0 = p_1 = q_1$, contradiction) or (n_0/p_1) is uniquely factorisable, so that $r = s$, $p_j = q_j$ and $e_j = f_j$ for all j, contradiction. □

1.3 Abelian groups

We introduce the main theoretical structure that will be used repeatedly in the course of the subject: the notion of an abelian group. This bit of formalism will allow us to conceptualise results and techniques.

Recall the definition of a real vector space: it is a set V with an operation of addition and an operation of multiplication by scalars, satisfying a number of axioms. We are interested in a structure that is more general than this, where we are only given an operation of addition.

An *abelian group* is a set G together with an operation $+$ such that:

(a) $g + (h + k) = (g + h) + k$ for all $g, h, k \in G$;

(b) $g + h = h + g$ for all $g, h \in G$;

(c) there exists $0 \in G$ such that $0 + g = g$ for all $g \in G$;

(d) for all $g \in G$, there exists $-g \in G$ such that $g + (-g) = 0$.

You will see in Exercise 1.13 that the neutral element and the inverse are unique.

Example 1.11. Some abelian groups you know:

(a) $\mathbf{Z}$ with usual addition;

(b) $\{0\}$;

(c) any vector space V over any field F, with vector addition;

(d) special case of (c): any field F (e.g. $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$) with its addition;

(e) special case of (c): for any field F, the subset of nonzero elements $F^\times$ with multiplication;

(f) special case of (c): $M_n(F)$ with matrix addition;

(g) but not the subset $\mathsf{GL}_n(F)$ of invertible matrices with matrix multiplication (since matrix multiplication is not commutative); this is a group, but not an abelian group.

Caveat: "addition" is just a name we give the operation so we can refer to it. In concrete settings, this could actually be given by multiplication, composition, etc.

Example 1.12 (Integers modulo m)**.** Fix an integer $m \geqslant 1$ and consider the set $G = \{0, 1, 2, \dots, m-1\}$. Define the addition on G as follows: given $g, h \in G$, let r be the remainder of the division of the integer $g + h$ by m, and define the element $g + h \in G$ to be r. Then G is an abelian group. It is denoted $\mathbf{Z}/m\mathbf{Z}$ (read as: "zed mod m zed").

The set $\mathbf{Z}/m\mathbf{Z}$ also has an operation of multiplication, defined similarly. Is it an abelian group under this operation as well? Certainly not when $m \geqslant 2$, since $0 \in \mathbf{Z}/m\mathbf{Z}$ does not have a multiplicative inverse.

It is not enough to remove 0: in $\mathbf{Z}/6\mathbf{Z}$, 2 does not have a multiplicative inverse.

Proposition 1.13. *Given an integer $m \geqslant 2$, the elements of $\mathbf{Z}/m\mathbf{Z}$ that have multiplicative inverses are precisely those $a \in \mathbf{Z}/m\mathbf{Z}$ such that $\gcd(a, m) = 1$.*

Proof. Let $a \in \mathbf{Z}/m\mathbf{Z}$ have a multiplicative inverse $x \in \mathbf{Z}/m\mathbf{Z}$, in other words $ax \equiv 1 \pmod{m}$. Then there exists $y \in \mathbf{Z}$ such that $ax - 1 = my$, that is $ax - my = 1$. By Proposition 1.7, this implies that $\gcd(a, m) = 1$.

Conversely, if $\gcd(a, m) = 1$ then by Proposition 1.7 there exist $x, y \in \mathbf{Z}$ such that $ax + my = 1$, so that $ax \equiv 1 \pmod{m}$, hence x is the multiplicative inverse of a. □

We denote this subset as

$$(\mathbf{Z}/m\mathbf{Z})^\times = \{a \in \mathbf{Z}/m\mathbf{Z} : \ \gcd(a, m) = 1\}.$$

It is an abelian group under multiplication. Its cardinality is called *Euler's phi function* $\phi(m)$.

Corollary 1.14. *If p is prime, then $(\mathbf{Z}/p\mathbf{Z})^\times = \{1, 2, \ldots, p-1\}$. Thus $\mathbf{Z}/p\mathbf{Z}$ is a field, denoted*[1] $\mathbf{F}_p$.

We should note another consequence of (the proof of) Proposition 1.13: given $a \pmod m$, one can find a multiplicative inverse for a by applying the extended Euclidean Algorithm to a and m to get $x, y \in \mathbf{Z}$ such that $ax + my = g$, where $g = \mathsf{gcd}\,(a, m)$; if $g = 1$ then $x \pmod m$ is the desired inverse, whereas if $g > 1$ then the inverse does not exist.

Subgroups

A *subgroup* of an abelian group G is a subset H of G such that:

(a) $0 \in H$;

(b) for all $x, y \in H$, we have $x - y \in H$.

Example 1.15. Let $G = \mathbf{Z}$ with the usual addition and let $H = m\mathbf{Z}$ be the subset of all multiples of a fixed $m \geqslant 0$.

Then $m\mathbf{Z}$ is a subgroup of $\mathbf{Z}$.

Moreover, all subgroups of $\mathbf{Z}$ are of this form (see Tutorial Question 2.8).

Example 1.16. Let $G = \mathbf{C}^\times$ with the usual complex multiplication and let $H = \mathbf{S}^1$, where

$$\mathbf{S}^1 = \{z \in \mathbf{C} : \ |z| = 1\}.$$

Then $\mathbf{S}^1$ is a subgroup of $\mathbf{C}^\times$.

(Note that we have put an abelian group structure on the unit circle.)

Given a subset S of an abelian group G, the *subgroup generated by S*, denoted $\langle S \rangle$, is the smallest subgroup of G that contains S, that is: $S \subseteq \langle S \rangle \subseteq G$, $\langle S \rangle$ is a subgroup of G, and if H is any subgroup of G such that $S \subseteq H \subseteq G$, then $\langle S \rangle \subseteq H$.

A subgroup H of an abelian group G is said to be *cyclic* if there exists an element $h \in H$ such that $H = \langle h \rangle$.

If G_1 and G_2 are abelian groups, their Cartesian product $G_1 \times G_2$ is also an abelian group, with operation defined by

$$(g_1, g_2) + (g_1', g_2') = (g_1 + g_1', g_2 + g_2').$$

Example 1.17. Taking $G_1 = G_2 = \mathbf{Z}/2\mathbf{Z}$, we get

$$(\mathbf{Z}/2\mathbf{Z}) \times (\mathbf{Z}/2\mathbf{Z}) = \{(0,0), (1,0), (0,1), (1,1)\}.$$

Example 1.18. The group $G = (\mathbf{Z}/2\mathbf{Z}) \times (\mathbf{Z}/2\mathbf{Z})$ is not cyclic. Since its cardinality is small, we can check this by computing the subgroups generated by each of its elements:

$$\begin{aligned}
\langle (0,0) \rangle &= \{(0,0)\} \\
\langle (1,0) \rangle &= \{(1,0), (0,0)\} \\
\langle (0,1) \rangle &= \{(0,1), (0,0)\} \\
\langle (1,1) \rangle &= \{(1,1), (0,0)\}.
\end{aligned}$$

However:

[1] Computer scientists sometimes write $\mathsf{GF}\,(p)$ instead of $\mathbf{F}_p$.

Example 1.19. The group $G = (\mathbf{Z}/2\mathbf{Z}) \times (\mathbf{Z}/3\mathbf{Z})$ is cyclic, since:

$$\langle (1,1) \rangle = \langle (1,1), (0,2), (1,0), (0,1), (1,2), (0,0) \rangle = G.$$

We will elucidate the patterns behind the last two examples in a little while.

If g is an element in an abelian group G, the *order of g* is the cardinality of the subgroup $\langle g \rangle$ generated by g. If this cardinality is not finite, we say that *g has infinite order*.

Example 1.20. The order of the neutral element in any abelian group G is 1.

The order of $(1,1) \in (\mathbf{Z}/2\mathbf{Z}) \times (\mathbf{Z}/3\mathbf{Z})$ is 6. The order of $(1,0) \in (\mathbf{Z}/2\mathbf{Z}) \times (\mathbf{Z}/3\mathbf{Z})$ is 2.

The order of $1 \in \mathbf{Z}$ is infinite.

Proposition 1.21. *If H is a subgroup of an abelian group G and both are finite, then the cardinality of H divides the cardinality of G.*

For a proof, see Exercise 1.14.

Directly from the definitions we get

Corollary 1.22. *The order of an element in a finite abelian group G divides the cardinality of G.*

Corollary 1.23 (Euler's Theorem)**.** *Let $m \geqslant 2$ and a be integers. If $\gcd(a, m) = 1$, then*

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

Proof. Since $a \in (\mathbf{Z}/m\mathbf{Z})^\times$, its order divides the cardinality $\phi(m)$ of the group, hence $a^{\phi(m)} = 1$ in the group. □

Corollary 1.24 (Fermat's little Theorem)**.** *If p is a prime and a is an integer not divisible by p, then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Proof. Special case $m = p$ prime of Euler's Theorem. □

Homomorphisms

(Think: "linear transformations for abelian groups".)

Given abelian groups G and H, a *group homomorphism* is a function $f : G \longrightarrow H$ such that

$$f(x + y) = f(x) + f(y) \qquad \text{for all } x, y \in G.$$

Example 1.25. Consider the map $f : \mathbf{R} \longrightarrow \mathbf{S}^1$ given by $f(\theta) = e^{i\theta}$, where $\mathbf{S}^1$ is defined in Example 1.16.

Then f is a group homomorphism.

Example 1.26. Fix $m \geqslant 1$ and consider the map $f : \mathbf{Z} \longrightarrow \mathbf{Z}/m\mathbf{Z}$ given by: for $a \in \mathbf{Z}$, let $f(a)$ be the remainder of the division of a by m.

Then f is a group homomorphism.

Example 1.27. Let G be an abelian group.

Some group homomorphisms $G \longrightarrow G$:

- $[2] : G \longrightarrow G$ given by $[2](x) = x + x$;
- $[1] = \mathsf{id}_G$;

- $[0]: G \longrightarrow G$ given by $[0](x) = 0_G$;
- $[-1]: G \longrightarrow G$ given by $[-1](x) = -x$;
- $[-2]: G \longrightarrow G$ given by $[-2](x) = [2](-x) = (-x) + (-x)$.

Most generally, for any $m \in \mathbf{Z}$ there is a group homomorphism $[m]: G \longrightarrow G$.

Similar to the case of linear transformations between vector spaces, given a group homomorphism $f: G \longrightarrow H$, we define its *kernel*

$$\ker f = \{g \in G : f(g) = 0\}$$

and its *image*

$$\mathsf{im}\, f = \{h \in H : \text{ there exists } g \in G \text{ such that } f(g) = h\}.$$

You will see in Exercise 1.16 that $\ker f$ is a subgroup of G and $\mathsf{im}\, f$ is a subgroup of H.

A *group isomorphism* is by definition a bijective group homomorphism. We say that abelian groups G and H are *isomorphic* if there exists a group isomorphism $f: G \longrightarrow H$.

1.4 Structure of integers modulo m

Chinese Remainder Theorem

Theorem 1.28 (Chinese Remainder Theorem, Congruence Version)**.** *Let $m_1, \ldots, m_r$ be pairwise coprime positive integers, and let $a_1, \ldots, a_r$ be integers. The system of congruences*

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ &\vdots \\ x &\equiv a_r \pmod{m_r} \end{aligned}$$

has a unique integer solution x with $0 \leqslant x < m_1 \ldots m_r$.

Corollary 1.29 (Chinese Remainder Theorem, Groups Version)**.** *Let $m_1, \ldots, m_r$ be pairwise coprime positive integers. There are group isomorphisms*

$$\begin{aligned} \mathbf{Z}/(m_1 \ldots m_r)\mathbf{Z} &\cong (\mathbf{Z}/m_1\mathbf{Z}) \times \cdots \times (\mathbf{Z}/m_r\mathbf{Z}), \\ (\mathbf{Z}/(m_1 \ldots m_r)\mathbf{Z})^\times &\cong (\mathbf{Z}/m_1\mathbf{Z})^\times \times \cdots \times (\mathbf{Z}/m_r\mathbf{Z})^\times. \end{aligned}$$

Corollary 1.30 (Group Structure of $\mathbf{Z}/m\mathbf{Z}$)**.** *Let $m \geqslant 2$ be an integer and consider its prime factorisation*

$$m = p_1^{e_1} \ldots p_r^{e_r}.$$

Then there is a group isomorphism

$$\mathbf{Z}/m\mathbf{Z} \cong (\mathbf{Z}/p_1^{e_1}\mathbf{Z}) \times \cdots \times (\mathbf{Z}/p_r^{e_r}\mathbf{Z}).$$

Corollary 1.31 (Group Structure of $(\mathbf{Z}/m\mathbf{Z})^\times$)**.** *Let $m \geqslant 2$ be an integer and consider its prime factorisation*

$$m = p_1^{e_1} \ldots p_r^{e_r}.$$

Then there is a group isomorphism

$$(\mathbf{Z}/m\mathbf{Z})^\times \cong (\mathbf{Z}/p_1^{e_1}\mathbf{Z})^\times \times \cdots \times (\mathbf{Z}/p_r^{e_r}\mathbf{Z})^\times.$$

Corollary 1.32. *Let $m \geqslant 2$ be an integer and consider its prime factorisation*

$$m = p_1^{e_1} \dots p_r^{e_r}.$$

Then

$$\phi(m) = \phi(p_1^{e_1}) \dots \phi(p_r^{e_r}).$$

It remains to describe more precisely the group structure of $(\mathbf{Z}/p^e\mathbf{Z})^\times$ for prime powers p^e.

Primitive roots

Theorem 1.33 (Primitive Root Theorem)**.** *For any prime p, the group $\mathbf{F}_p^\times$ is cyclic.*

Any element g such that $\mathbf{F}_p^\times = \langle g \rangle$ is called a *primitive root* modulo p.

Proof. This is slightly more involved than one might think. We will use two ingredients:

- The fact that a nonzero polynomial of degree d over any field K has at most d distinct roots in K.
- The fact that in an abelian group G, given an element of order n and an element of order m, there exists an element of order $\mathsf{lcm}\,(n, m)$ (see Tutorial Question 2.9 for a proof of this).

Now for the proof: let $m = \mathsf{max}\,\{\mathsf{ord}\,(a) : \ a \in \mathbf{F}_p^\times\}$, then $m \leqslant p-1$, the cardinality of $\mathbf{F}_p^\times$.

We need to prove that $m \geqslant p-1$. For this, let $b \in \mathbf{F}_p^\times$ be arbitrary and let $n = \mathsf{ord}\,(b)$. I claim that $n \mid m$. Otherwise by the second fact quoted above, there exists an element of $\mathbf{F}_p^\times$ of order $\mathsf{lcm}\,(n, m) > m$, contradicting the maximality of m. So $b^n \equiv 1 \pmod p$ for some $n \mid m$, hence $b^m \equiv 1 \pmod p$, so that b is a root of the polynomial $x^m - 1$ over the field $\mathbf{F}_p$. We conclude that this polynomial of degree m has at least $p-1$ distinct roots, hence $m \geqslant p-1$. □

Example 1.34 (Primitive roots for primes up to 50)**.**

p	primitive roots g with $1 \leqslant g < p$	count
2	1	1
3	2	1
5	2, 3	2
7	3, 5	2
11	2, 6, 7, 8	4
13	2, 6, 7, 11	4
17	3, 5, 6, 7, 10, 11, 12, 14	8
19	2, 3, 10, 13, 14, 15	6
23	5, 7, 10, 11, 14, 15, 17, 19, 20, 21	10
29	2, 3, 8, 10, 11, 14, 15, 18, 19, 21, 26, 27	12
31	3, 11, 12, 13, 17, 21, 22, 24	8
37	2, 5, 13, 15, 17, 18, 19, 20, 22, 24, 32, 35	12
41	6, 7, 11, 12, 13, 15, 17, 19, 22, 24, 26, 28, 29, 30, 34, 35	16
43	3, 5, 12, 18, 19, 20, 26, 28, 29, 30, 33, 34	12
47	5, 10, 11, 13, 15, 19, 20, 22, 23, 26, 29, 30, 31, 33, 35, 38, 39, 40, 41, 43, 44, 45	22

Artin's Conjecture. *Let $g \in \mathbf{Z}$ be such that $g \neq -1$ and g is not a perfect square. Then there exist infinitely many primes p such that $g \pmod p$ is a primitive root.*

There are more precise statements, for instance if g is squarefree and $g \not\equiv 1 \pmod 4$, then the density of primes such that $g \pmod p$ is a primitive root is

$$\lim_{X\to\infty} \frac{\#\{p \text{ prime} : \ p \leqslant X \text{ and } g \pmod p \text{ is a primitive root}\}}{\#\{p \text{ prime} : \ p \leqslant X\}} = \prod_p \left(1 - \frac{1}{p(p-1)}\right) \approx 0.374,$$

where the infinite product is over all prime numbers p.

Artin stated his conjecture in 1927. Despite a lot of work going into it, the conjecture is still open, even for a single value of g such as $g = 2$. In 1986, Heath-Brown proved that at least one of the numbers $2, 3, 5$ is a primitive root for infinitely many primes.

Going back to the aim of this section, Theorem 1.33 explicates the group structure of $(\mathbf{Z}/p^e\mathbf{Z})^\times$ when $e = 1$. We can express the result as the existence of a group isomorphism

$$(\mathbf{Z}/p\mathbf{Z})^\times \cong \mathbf{Z}/(p-1)\mathbf{Z}.$$

The case of general $e \geqslant 1$ is similar:

Theorem 1.35. *If p is an odd prime and $e \in \mathbf{Z}_{\geqslant 1}$, then the group $(\mathbf{Z}/p^e\mathbf{Z})^\times$ is cyclic, so that there is a group isomorphism*

$$(\mathbf{Z}/p^e\mathbf{Z})^\times \cong \mathbf{Z}/p^{e-1}(p-1)\mathbf{Z}.$$

*If $p = 2$ and $e \geqslant 3$, then the group $(\mathbf{Z}/2^e\mathbf{Z})^\times$ is **not** cyclic; there is a group isomorphism*

$$(\mathbf{Z}/2^e\mathbf{Z})^\times \cong \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2^{e-2}\mathbf{Z}.$$

The group $(\mathbf{Z}/2^2\mathbf{Z})^\times$ is cyclic of order 2.

Corollary 1.36. *For any prime p and any integer $e \geqslant 1$ we have*

$$\phi(p^e) = p^{e-1}(p-1).$$

1.5 The RSA cryptosystem

We discuss the Rivest–Shamir–Adleman public key cryptosystem, introduced in the 1970s.

The setup

Alice wants to send a message to Bob.

In preparation for this, Bob chooses distinct primes p and q of similar size and computes $N = pq$. He then chooses an element $e \in (\mathbf{Z}/\phi(N)\mathbf{Z})^\times$ and computes its multiplicative inverse d, in other words solves for d in the congruence

$$ed \equiv 1 \pmod{\phi(N)}.$$

Note that since $N = pq$ we have $\phi(N) = (p-1)(q-1)$.

The pair (N, e) forms Bob's *public key*, which can be made available publicly (in particular to Alice).

The integer d is Bob's *private key*, which must remain known only to Bob. The primes p and q are not needed anymore (but must be disposed of).

Encryption and decryption

Alice obtains Bob's public key (N, e). She encodes her *plaintext* message as an integer $1 \leqslant m < N$.

She then computes the *ciphertext*

$$c \equiv m^e \pmod{N}$$

and sends this to Bob.

Bob receives c and computes $c^d \pmod{N}$.

I claim that $c^d \equiv m \pmod{N}$, so that Bob now has the plaintext m.

Here is why the claim is true, in the special case where $m \in (\mathbf{Z}/N\mathbf{Z})^\times$: since $ed \equiv 1 \pmod{\phi(N)}$, there exists ℓ such that $ed - 1 = \ell\phi(N)$. Therefore

$$c^d \equiv (m^e)^d = m^{ed} = m^{\ell\phi(N)} \cdot m = \left(m^{\phi(N)}\right)^\ell \cdot m \equiv 1^\ell \cdot m = m \pmod{N},$$

where we used Euler's Theorem (Corollary 1.23).

How restrictive is the assumption that $m \in (\mathbf{Z}/N\mathbf{Z})^\times$? You get to see in Assignment Question 1.3.

Practicalities

Generally, the size of the public key N is at least 2048 bits (so p and q should be primes in the interval $[2^{1023}, 2^{1024}]$). (The largest number N of this type that is known[2] to have been factored has 829 bits; this was in 2020.)

How does one produce such primes? We'll find out in the next section.

The other half e of the pair (N, e) forming the public key is typically chosen so that m^e is not too onerous to compute, but without sacrificing safety. It is common to pick $e = 65537 = 2^{16} + 1$.

The strategy used for computing m^e is called *repeated squaring* or (more descriptively) *square-and-multiply*. The basic observation is that computing $m^{2^r} \pmod{N}$ requires only r multiplications mod N: let $m_0 = m$ and for $j \geqslant 1$ let $m_j = m_{j-1}^2 \pmod{N}$, then $m_r = m^{2^r} \pmod{N}$. Of course, there is no reason why e should be expected to be a power of 2; in general we write e in base 2:

$$e = (1e_{r-1} \dots e_1 e_0)_2 = 2^r + 2^{r-1}e_{r-1} + \cdots + 2e_1 + e_0, \quad \text{where } r \geqslant 0 \text{ and } e_{r-1}, \dots, e_0 \in \{0, 1\}.$$

Set $m_0 = m$ and for $j \geqslant 1$ set

$$m_j = \begin{cases} m_{j-1}^2 \pmod{N} & \text{if } e_{r-j} = 0, \\ m_{j-1}^2 \cdot m \pmod{N} & \text{if } e_{r-j} = 1. \end{cases}$$

Then $m_r = m^e \pmod{N}$. Since this takes $r = \lfloor \log_2(e) \rfloor - 1$ steps and each step requires at most 2 multiplications, we conclude that the encryption step costs roughly $2\log_2(e)$ multiplications mod m.

The typical choice $e = 65537 = 2^{16} + 1 = (1000000000000001)_2$ has $r = 16$ and uses only 17 multiplications. It has the advantage that it is prime, so that checking the condition $e \in (\mathbf{Z}/\phi(N)\mathbf{Z})^\times$ reduces to verifying that e does not divide either of $p - 1$ and $q - 1$.

1.6 Primality

In order to understand how RSA and related schemes work, we need to know a bit more about prime numbers.

[2] Conspiracy theories abound regarding shadowy government entities having factored many of the RSA challenge numbers without telling anyone about it.

Distribution

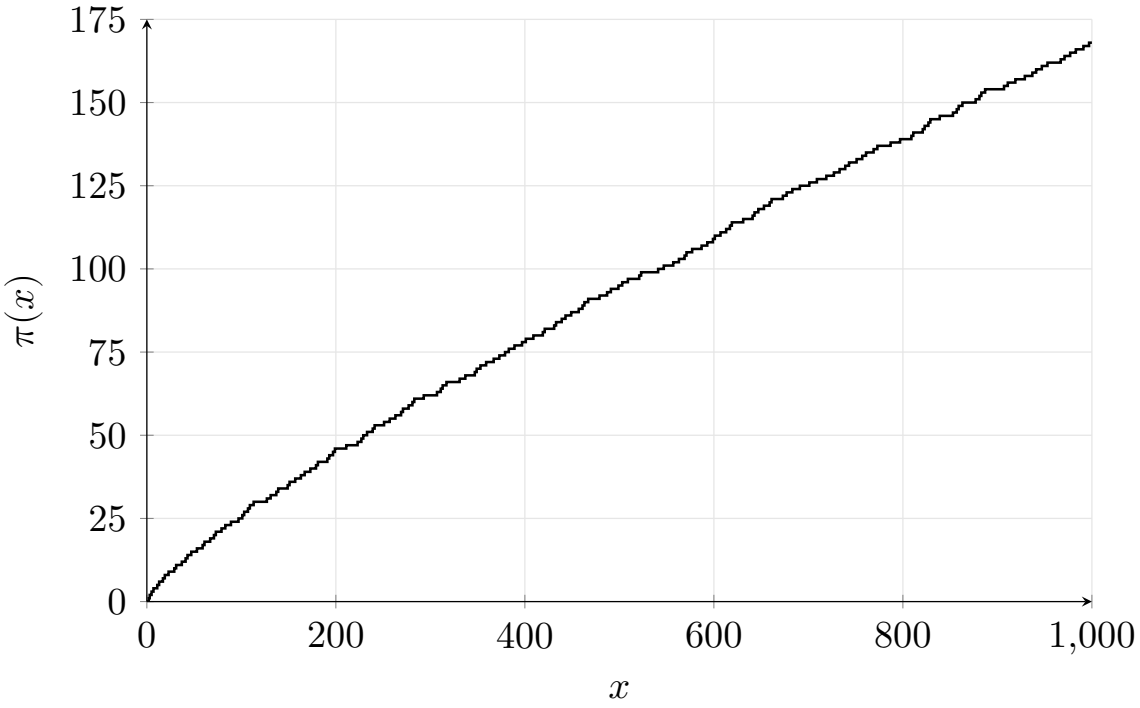
We know already from Euclid that there are infinitely many prime numbers.

It took longer to make this more precise. Consider the *prime-counting function* $\pi : \mathbf{R}_{>0} \longrightarrow \mathbf{N}$ given by

$$\pi(x) = \#\{p \leqslant x : \ p \text{ prime}\}.$$

For instance, $\pi(1.3) = 0$, $\pi(10) = 4$, and $\pi(4\pi) = 5$.

Apart from being (obviously) non-decreasing and (obviously) going up by 1 whenever it does go up, the behaviour of π is somewhat erratic:



It was conjectured to be asymptotic to $x/\log(x)$, which was eventually proved by Hadamard and de la Vallée Poussin:

Theorem 1.37 (Prime Number Theorem)**.**

$$\pi(x) \sim \frac{x}{\log(x)} \quad \textit{as } x \longrightarrow \infty,$$

where the *asymptotic notation* $f(x) \sim g(x)$ means that

$$\lim_{x \longrightarrow \infty} \frac{f(x)}{g(x)} = 1.$$

In practice, one can do better using the *logarithmic integral* Li:

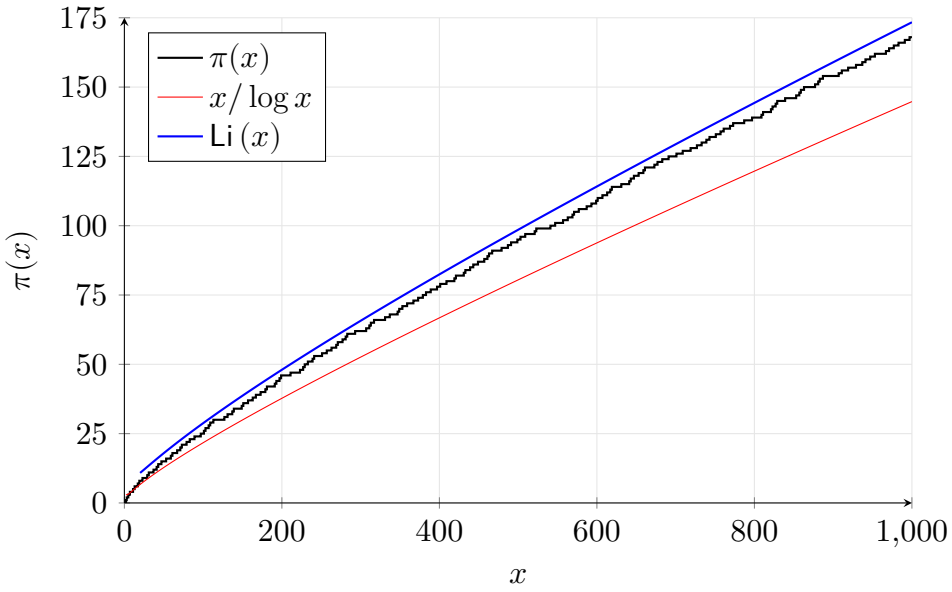
$$\pi(x) \sim \mathsf{Li}\,(x), \qquad \text{where} \qquad \mathsf{Li}\,(x) = \int_2^x \frac{dt}{\log(t)}.$$

Various types of bounds are known for the errors $|\pi(x) - x/\log(x)|$ and $|\pi(x) - \mathsf{Li}\,(x)|$, and they are pretty ugly. If one is willing to believe the Riemann Hypothesis, then the bounds become better and nicer:

$$\pi(x) - \mathsf{Li}\,(x) = O(\sqrt{x}\,\log(x)).$$

The *big-oh notation* $f(x) = O(g(x))$ means that: there exists a constant $C \in \mathbf{R}_{>0}$ such that

$$|f(x)| \leqslant C\, g(x) \qquad \text{for all } x.$$



What does this mean in terms of primes suitable for use in the RSA system? If we want a key N of size roughly 2048 bits, then p and q are primes in the interval $[2^{1023}, 2^{1024}]$. The number of such primes is then

$$\pi(2^{1024}) - \pi(2^{1023}) \approx \mathsf{Li}\,(2^{1024}) - \mathsf{Li}\,(2^{1023}) \approx 1.267 \cdot 10^{305},$$

so there is definitely a good supply.

Primality testing

Continuing the back-of-the-envelope calculation, if we pick a random integer in the interval $[2^{1023}, 2^{1024}]$, the probability of it being a prime number should be about

$$\frac{1.267 \cdot 10^{305}}{2^{1023}} \approx 0.14\%.$$

If we're moderately awake and pick a random integer with last digit in $\{1, 3, 7, 9\}$, the probability goes up to 0.35%. This means that we might expect success if we repeat this about 300 times, as long as we have some quick way of checking whether a number of that size is prime.

Recall that Fermat's little Theorem says: if p is prime, then for every integer a coprime to p we have $a^{p-1} \equiv 1 \pmod{p}$. We can state this in a more uniform way as: if p is prime, then for every integer a we have $a^p \equiv a \pmod{p}$. The contrapositive of this is:

Corollary 1.38 (Fermat Compositeness Test)**.** *Let $n > 1$ be an integer. If there exists an integer a such that $a^n \not\equiv a \pmod{n}$, then n is composite.*

We call such an integer a a *Fermat witness* (of the compositeness of n).

Combined with the fast modular exponentiation we discussed in the RSA section, this gives us a very efficient way of ruling out candidates. But what if we try a few a's and they stubbornly refuse to be witnesses for n? When can we conclude that n is actually prime?

Well, in some sense, never. We say that an integer $n > 1$ is a *Carmichael number* if n is composite and $a^n \equiv a \pmod{n}$ for all a. Such things do in fact exist; the smallest is 561, and in 1984 Alford–Granville–Pomerance proved that there are infinitely many Carmichael numbers (although they are somewhat rare).

This points to the need for a better test. Given $n > 1$ odd, write $n - 1 = 2^k m$ with m odd. We call an integer a with $\gcd(a, n) = 1$ a *Miller–Rabin witness* for n if

$$a^m \not\equiv 1 \pmod{n} \qquad \text{and} \qquad a^{2^j m} \not\equiv -1 \pmod{n} \text{ for all } j \in \{0, 1, \ldots, k-1\}.$$

Proposition 1.39. *Let p be an odd prime and write $p-1=2^k m$ with m odd, and let a be an integer with $\gcd(a,p)=1$. Then one of the following properties holds:*

(a) $a^m \equiv 1 \pmod{p}$;

(b) there exists $j \in \{0,1,\dots,k-1\}$ *such that* $a^{2^j m} \equiv -1 \pmod{p}$.

Proof. Since $\mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$ is a field, the polynomial x^2-1 has at most two roots in $\mathbf{F}_p$. We know that 1 and -1 are roots, and they are distinct for odd p, so in our setting $x^2-1=0$ has exactly the solutions $1,-1 \in \mathbf{Z}/p\mathbf{Z}$.

Let's consider a now. By Fermat's little Theorem we have $(a^m)^{2^k} = a^{p-1} \equiv 1 \pmod{p}$. In particular, $(a^m)^{2^{k-1}}$ is a root of x^2-1, so it is either 1 or -1. If it is -1, we are done. Otherwise $(a^m)^{2^{k-1}} \equiv 1 \pmod{p}$, so $(a^m)^{2^{k-2}}$ is a root of x^2-1, so it is either 1 or -1. We continue in this manner; as soon as we hit a -1, we are done and we can stop. If this does not happen we get down all the way to a^m being a root of x^2-1, and if it is not -1 then it must be 1 and we are done. □

Corollary 1.40 (Miller–Rabin Compositeness Test)**.** *If a is a Miller–Rabin witness for n, then n is composite.*

Example 1.41. Consider $n=561$. Write $n-1=560=2^4\cdot 35$. Take $a=2$ and compute

$$2^{35} \equiv 263 \pmod{561},\quad 2^{2\cdot 35} \equiv 166 \pmod{561},\quad 2^{2^2\cdot 35} \equiv 67 \pmod{561},\quad 2^{2^3\cdot 35} \equiv 1 \pmod{561}.$$

Hence 2 is a Miller–Rabin witness for 561, so 561 is composite.

Two features make this significantly better than the Fermat approach: (i) there are no analogues of the Carmichael numbers for Miller–Rabin; (ii) any composite number is guaranteed to have lots of Miller–Rabin witnesses.

Proposition 1.42. *If $n>9$ is an odd composite number, then at least 75% of all $a \in (\mathbf{Z}/n\mathbf{Z})^\times$ are Miller–Rabin witnesses for n.*

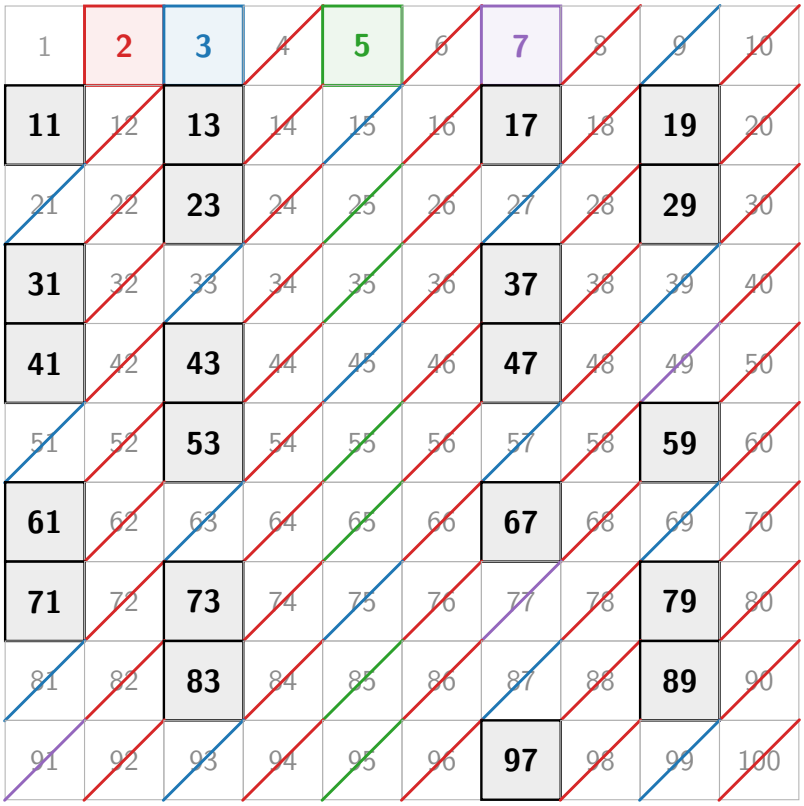
This gives a probabilistic primality test: given n odd, choose a random $a \in (\mathbf{Z}/n\mathbf{Z})^\times$ and see if it is a Miller–Rabin witness for n. If yes, n is proved composite. If no, n is prime with $1/4$ probability of failure. If we repeat this k times and do not find a Miller–Rabin witness, the probability of failure is reduced to $1/4^k$. Pick k as desired.

The test is so fast that running it 300 times for random numbers with last digit in $\{1,3,7,9\}$ in the interval $[2^{1023}, 2^{1024}]$ is very quick.

Enumerating all primes up to a bound

This is worth a quick detour: what can we do if, instead of needing to find a couple of primes of a given size, we need to find all the primes up to some bound M?

It is a very old problem with a very old solution: the *sieve of Eratosthenes*.



The efficiency of this approach is pretty astounding, especially for an algorithm this old.

It is possible to do better though, and you can hear about it in David Harvey's pure mathematics seminar talk on Friday 14 August at 2pm in Peter Hall 162.

Integer factorisation

So far we have focussed on the generation and testing of primes, as efficient approaches to these tasks are crucial to the practicality of the RSA cryptosystem.

Another aspect is even more crucial though: that the factorisation of very large integers appears to be a computationally difficult problem. The safety of RSA hinges on this aspect; otherwise Eve could simply factor the N of the public key and use the resulting p and q to easily compute the decryption exponent d.

We will say a bit more about factorisation methods once we have seen more of the relevant concepts. For the moment it will suffice to know that the best known factorisation algorithms have a running time that is subexponential[3] in the size of N, whereas everything involved in the production of N and the encryption/decryption process runs in time polynomial in the size of N.

A minor caveat is that, just because no polynomial-time factorisation algorithms are known, it does not mean that they do not exist (in the mathematical sense of existence).

A major caveat is that all of these considerations are based on the models of computation of conventional computers. In 1994, Shor described a quantum computation algorithm that factors integers in polynomial time. Despite the fact that quantum computers that can reliably run a computation of this size do not exist at the moment, it is likely that they will become a reality in the not-too-distant future. Therefore, there is a move (rush?) towards transitioning to the use of quantum-resistant cryptosystems, based on problems that are known (or suspected) to be hard to solve even with quantum computers.

[3] This has a precise meaning, but for now just think "barely faster than exponential".