Number Theory

Ben Lynn

COLLABORATORS

	TITLE : Number Theory		
ACTION	*NAME*	*DATE*	*SIGNATURE*
WRITTEN BY	Ben Lynn	1980-01-01	

REVISION HISTORY

NUMBER	DATE	DESCRIPTION	NAME

Contents

1 Number Theory

I'm taking a loose informal approach, since that was how I learned. Once you have a good feel for this topic, it is easy to add rigour.

More formal approaches can be found all over the net, e.g: Victor Shoup, *A Computational Introduction to Number Theory and Algebra*.

One reader of these notes recommends I.N. Herstein, 'Abstract Algebra' for further reading.

I built a PDF version of these notes.

1.1 Overview

I have tried to order my pages so that the parts most relevant to cryptography are presented first.

Modular Arithmetic We begin by defining how to perform basic arithmetic modulo n, where n is a positive integer. Addition, subtraction, and multiplication follow naturally from their integer counterparts, but we have complications with division.

Euclid's Algorithm We will need this algorithm to fix our problems with division. It was originally designed to find the greatest common divisor of two numbers.

Division Once armed with Euclid's algorithm, we can easily compute divisions modulo n.

The Chinese Remainder Theorem We find we only need to study $\mathbb{Z}_{p^k}$ where p is a prime, because once we have a result about the prime powers, we can use the Chinese Remainder Theorem to generalize for all n.

Units While studying division, we encounter the problem of inversion. Units are numbers with inverses.

Exponentiation The behaviour of units when they are exponentiated is difficult to study. Modern cryptography exploits this.

Order of a Unit If we start with a unit and keep multiplying it by itself, we wind up with 1 eventually. The order of a unit is the number of steps this takes.

The Miller-Rabin Test We discuss a fast way of telling if a given number is prime that works with high probability.

Generators Sometimes powering up a unit will generate all the other units.

Cyclic Groups We focus only on multiplication and see if we can still say anything interesting.

Quadratic Residues Elements of $\mathbb{Z}_n$ that are perfect squares are called quadratic residues.

1.2 Bonus Material

The other topics are less relevant to cryptography, but nonetheless interesting.

2 Modular Arithmetic

Let n be a positive integer. We denote the set $[0..n-1]$ by $\mathbb{Z}_n$.

We consider two integers x, y to be the same if x and y differ by a multiple of n, and we write this as $x = y \pmod{n}$, and say that x and y are *congruent* modulo n. We may omit $\pmod{n}$ when it is clear from context. Every integer x is congruent to some y in $\mathbb{Z}_n$. When we add or subtract multiples of n from an integer x to reach some $y \in \mathbb{Z}_n$, we say are *reducing* x modulo n, and y is the *residue*.

We could have chosen different sets for $\mathbb{Z}_n$, e.g. we could add n to every member, but our default will be $[0..n-1]$. The elements in this particular representation of $\mathbb{Z}_n$ are called the *least residues*.

Example: $38 = 3 \pmod{5}$ since $38 = 7 \times 5 + 3$. $-3 = 11 \pmod{14}$ since $-3 = (-1) \times 14 + 11$.

What is the most natural way of doing arithmetic in $\mathbb{Z}_n$? Given two elements $x, y \in \mathbb{Z}_n$, we can add, subtract or multiply them as integers, and then the result will be congruent to one of the elements in $\mathbb{Z}_n$.

Example: $6+7=1 \pmod{12}$, $3\times 20=10 \pmod{50}$, $12-14=16 \pmod{18}$.

These operations behave similarly to their mundane counterparts. However, there is no notion of size. Saying $0<4 \pmod 8$ is nonsense for example, because if we add 4 to both sides we find $4<0 \pmod 8$. The regular integers are visualized as lying on a number line, where integers to the left are smaller than integers on the right. Integers modulo n however are visualized as lying on a circle (e.g. think of a clock when working modulo 12).

2.1 Division

Division is notably absent from the above discussion. If y divides x as integers, then one might guess we could use the usual definition. Let us see where this leads: we have $10=4 \pmod 6$. Dividing both sides by 2 gives the incorrect equation $5=2 \pmod 6$.

Thus we have to change what division means. Intuitively, division should "undo multiplication", that is to divide x by y means to find a number z such that y times z is x. The problem above is that there are different candidates for z: in $\mathbb{Z}_6$ both 5 and 2 give 4 when multiplied by 2.

Which answer should we choose for "$4/2$", 5 or 2? We could introduce some arbitrary convention, such as choosing the smallest answer when considering the least residue as an integer, but then division will behave strangely.

Instead, we require uniqueness, that is x divided by y modulo n is only defined when there is a unique $z \in \mathbb{Z}_n$ such that $x=yz$.

We can obtain a condition on y as follows. Suppose $z_1 y = z_2 y \pmod n$. Then by definition, this means for some k we have $y(z_1-z_2)=kn$. Let d be the greatest common divisor of n and y. Then n/d divides z_1-z_2 since it cannot divide y, thus we have

$$z_1 y = z_2 y \pmod n$$

if and only if

$$z_1 = z_2 \pmod{n/d}.$$

Thus a unique z exists modulo n only if the greatest common divisor of y and n is 1.

2.2 Inverses

We shall see that a unique z exists if and only if it is possible to find a $w \in \mathbb{Z}_n$ such that $yw=1 \pmod n$. If such a w exists, it must be unique: suppose yw' is also 1. Then multiplying both sides of $yw=yw'$ by w gives $wyw=wyw'$, which implies $w=w'$ since $wy=1$. When it exists, we call this unique w the *inverse* of y and denote it by y^{-1}.

How do we know if y^{-1} exists, and if it does, how do we find it? Since there are only n elements in $\mathbb{Z}_n$, we can multiply each element in turn by y and see if we get 1. If none of them work then we know y does not have an inverse. In some sense, modular arithmetic is easier than integer arithmetic because there are only finitely many elements, so to find a solution to a problem you can always try every possbility.

We now have a good definition for division: x divided by y is x multiplied by y^{-1} if the inverse of y exists, otherwise the answer is undefined.

To avoid confusion with integer division, many authors avoid the $/$ symbol completely in modulo arithmetic and if they need to divide x by y, they write xy^{-1}. Also some approaches to number theory start with inversion, and define division using inversion without discussing how it relates to integer division, which is another reason $/$ is often avoided. We will follow convention, and reserve the $/$ symbol for integer division.

Example: $2\times 3+4(5^{-1})=2 \pmod 6$.

3 Euclid's Algorithm

Given three integers a, b, c, can you write c in the form

$$c = ax + by$$

for integers x and y? If so, is there more than one solution? Can you find them all? Before answering this, let us answer a seemingly unrelated question:

How do you find the greatest common divisor (gcd) of two integers a, b?

We denote the greatest common divisor of a and b by $\gcd(a,b)$, or sometimes even just (a,b). If $(a,b) = 1$ we say a and b are *coprime*.

The obvious answer is to list all the divisors a and b, and look for the greatest one they have in common. However, this requires a and b to be factorized, and no one knows how to do this efficiently.

A few simple observations lead to a far superior method: Euclid's algorithm, or the Euclidean algorithm. First, if d divides a and d divides b, then d divides their difference, a - b, where a is the larger of the two. But this means we've shrunk the original problem: now we just need to find $\gcd(a, a-b)$. We repeat until we reach a trivial case.

Hence we can find $\gcd(a,b)$ by doing something that most people learn in primary school: division and remainder. We give an example and leave the proof of the general case to the reader.

Suppose we wish to compute $\gcd(27,33)$. First, we divide the bigger one by the smaller one:

$$33 = 1 \times 27 + 6$$

Thus $\gcd(33,27) = \gcd(27,6)$. Repeating this trick:

$$27 = 4 \times 6 + 3$$

and we see $\gcd(27,6) = \gcd(6,3)$. Lastly,

$$6 = 2 \times 3 + 0$$

Since 6 is a perfect multiple of 3, $\gcd(6,3) = 3$, and we have found that $\gcd(33,27) = 3$.

This algorithm does not require factorizing numbers, and is fast. We obtain a crude bound for the number of steps required by observing that if we divide a by b to get $a = bq + r$, and $r > b/2$, then in the next step we get a remainder $r' \leq b/2$. Thus every two steps, the numbers shrink by at least one bit.

3.1 Extended Euclidean Algorithm

The above equations actually reveal more than the gcd of two numbers. We can use them to find integers m, n such that

$$3 = 33m + 27n$$

First rearrange all the equations so that the remainders are the subjects:

$$6 = 33 - 1 \times 27$$

$$3 = 27 - 4 \times 6$$

Then we start from the last equation, and substitute the next equation into it:

$$3 = 27 - 4 \times (33 - 1 \times 27) = (-4) \times 33 + 5 \times 27)$$

And we are done: $m = -4, n = 5$.

If there were more equations, we would repeat until we have used them all to find m and n.

Thus in general, given integers a and b, let $d = \gcd(a,b)$. Then we can find integer m and n such that

$$d = ma + nb$$

using the extended Euclidean algorithm.

3.2 The General Solution

We can now answer the question posed at the start of this page, that is, given integers a, b, c find all integers x, y such that

$$c = xa + yb.$$

Let $d = \gcd(a,b)$, and let $b = b'd, a = a'd$. Since $xa + yb$ is a multiple of d for any integers x, y, solutions exist only when d divides c.

So say $c = kd$. Using the extended Euclidean algorithm we can find m, n such that $d = ma + nb$, thus we have a solution $x = km, y = kn$.

Suppose x', y' is another solution. Then

$$c = xa + yb = x'a + y'b$$

Rearranging,

$$(x' - x)a = (y - y')b$$

Dividing by d gives:

$$(x' - x)a' = (y - y')b'$$

The numbers a' and b' are coprime since d is the greatest common divisor, hence $(x' - x)$ is some multiple of b', that is:

$$x' - x = tb/d$$

for some integer t. Then solving for $(y - y')$ gives

$$y' - y = ta/d$$

Thus $x' = x + tb/d$ and $y' = y - ta/d$ for some integer t.

But if we replace t with any integer, x' and y' still satisfy $c = x'a + y'b$. Thus there are infinitely many solutions, and they are given by

$$x = km + tb/d, y = kn + ta/d.$$

for all integers t.

Later, we shall often wish to solve $1 = xp + yq$ for coprime integers p and q. In this case, the above becomes

$$x = m + tq, y = n + tp.$$

4 Division

Intuitively, to divide x by y means to find a number z such that y times z is x, but we had trouble adopting this defintion of division because sometimes there is more than one possibility for z modulo n.

We solved this by only defining division when the answer is unique. We stated without proof that when division defined in this way, one can divide by y if and only if y^{-1}, the inverse of y exists.

We shall now show why this is the case. We wish to find all z such that $yz = x \pmod{n}$, which by definition means

$$x = zy + kn$$

for some integer k. But this is precisely the problem we encountered when discussing Euclid's algorithm! Let $d = \gcd(y, n)$.

Suppose $d > 1$. Then no solutions exist if x is not a multiple of d. Otherwise the solutions for z, k are

$$z = r + tn/d, k = s - ty/d$$

for some integers r, s (that we get from Euclid's algorithm) and for all integers t. But this means z does not have a unique solution modulo n since $n/d < n$. (Instead z has a unique solution modulo n/d.)

On the other hand, if $d = 1$, that is if y and n are coprime, then x is always a multiple of d so solutions exist. Recall we find them by using Euclid's algorithm to find r, s such that

$$1 = ry + sn$$

Then the solutions for z, k are given by

$$z = xr + tn, k = zs - ty$$

for all integers t. Thus z has a unique solution modulo n, and division makes sense for this case.

Also, r satisfies $ry = 1 \pmod{n}$ so in fact $y^{-1} = r$. Thus our claims are correct: we can divide x by y if and only if y has an inverse.

We can also see that y has an inverse if and only if $\gcd(y, n) = 1$. (Actually, we have only proved some of these statements in one direction, but the other direction is easy.)

Note: even though we cannot always divide x by y modulo n, sometimes we still need to find all z such that $yz = x$.

4.1 Computing Inverses

We previously asked: given $y \in \mathbb{Z}_n$, does y^{-1} exist, and if so, what is it?

Our answer before was that since $\mathbb{Z}_n$ is finite, we can try every possibility. But if n is large, say a 256-bit number, this cannot be done even if we use the fastest computers available today.

A better way is to use what we just proved: y^{-1} exists if and only if $\gcd(y, n) = 1$ (which we can check using Euclid's algorithm), and y^{-1} can be computed efficiently using the extended Euclidean algorithm.

Example: does $7^{-1} \pmod{19}$ exist, and if so, what is it? Euclid's algorithm gives

$$19 = 2 \times 7 + 5$$

$$7 = 1 \times 5 + 2$$

$$5 = 2 \times 2 + 1$$

Thus an inverse exists since $\gcd(7, 19) = 1$. To find the inverse we rearrange these equations so that the remainders are the subjects. Then starting from the third equation, and substituting in the second one gives

$$\begin{aligned} 1 &= 5 - 2 \times 2 \\ &= 5 - 2 \times (7 - 1 \times 5) \\ &= (-2) \times 7 + 3 \times 5 \end{aligned}$$

Now substituting in the first equation gives

$$\begin{aligned} 1 &= (-2) \times 7 + 3 \times (19 - 2 \times 7) \\ &= (-8) \times 7 + 3 \times 19 \end{aligned}$$

from which we see that $7^{-1} = -8 = 11 \pmod{19}$.

5 The Chinese Remainder Theorem

Suppose we wish to solve

$$x = 2 \pmod{5}$$

$$x = 3 \pmod{7}$$

for x. If we have a solution y, then $y + 35$ is also a solution. So we only need to look for solutions modulo 35. By brute force, we find the only solution is $x = 17 \pmod{35}$.

For any system of equations like this, the Chinese Remainder Theorem tells us there is always a unique solution up to a certain modulus, and describes how to find the solution efficiently.

Theorem: Let p, q be coprime. Then the system of equations

$$x = a \pmod{p}$$

$$x = b \pmod{q}$$

has a unique solution for x modulo pq.

The reverse direction is trivial: given $x \in \mathbb{Z}_{pq}$, we can reduce x modulo p and x modulo q to obtain two equations of the above form.

Proof: Let $p_1 = p^{-1} \pmod{q}$ and $q_1 = q^{-1} \pmod{p}$. These must exist since p, q are coprime. Then we claim that if y is an integer such that

$$y = aqq_1 + bpp_1 \pmod{pq}$$

then y satisfies both equations:

Modulo p, we have $y = aqq_1 = a \pmod{p}$ since $qq_1 = 1 \pmod{p}$. Similarly $y = b \pmod{q}$. Thus y is a solution for x.

It remains to show no other solutions exist modulo pq. If $z = a \pmod{p}$ then $z - y$ is a multiple of p. If $z = b \pmod{q}$ as well, then $z - y$ is also a multiple of q. Since p and q are coprime, this implies $z - y$ is a multiple of pq, hence $z = y \pmod{pq}$.

This theorem implies we can represent an element of $\mathbb{Z}_{pq}$ by one element of $\mathbb{Z}_p$ and one element of $\mathbb{Z}_q$, and vice versa. In other words, we have a bijection between $\mathbb{Z}_{pq}$ and $\mathbb{Z}_p \times \mathbb{Z}_q$.

Examples: We can write $17 \in \mathbb{Z}_{35}$ as $(2,3) \in \mathbb{Z}_5 \times \mathbb{Z}_7$. We can write $1 \in \mathbb{Z}_{pq}$ as $(1,1) \in \mathbb{Z}_p \times \mathbb{Z}_q$.

In fact, this correspondence goes further than a simple relabelling. Suppose $x, y \in \mathbb{Z}_{pq}$ correspond to $(a,b), (c,d) \in \mathbb{Z}_p \times \mathbb{Z}_q$ respectively. Then a little thought shows $x+y$ corresponds to $(a+c, b+d)$, and similarly xy corresponds to (ac, bd).

A practical application: if we have many computations to perform on $x \in \mathbb{Z}_{pq}$ (e.g. RSA signing and decryption), we can convert x to $(a,b) \in \mathbb{Z}_p \times \mathbb{Z}_q$ and do all the computations on a and b instead before converting back. This is often cheaper because for many algorithms, doubling the size of the input more than doubles the running time.

Example: To compute $17 \times 17 \pmod{35}$, we can compute $(2 \times 2, 3 \times 3) = (4,2)$ in $\mathbb{Z}_5 \times \mathbb{Z}_7$, and then apply the Chinese Remainder Theorem to find that $(4,2)$ is $9 \pmod{35}$.

Let us restate the Chinese Remainder Theorem in the form it is usually presented.

5.1 For Several Equations

Theorem: Let $m_1, ..., m_n$ be pairwise coprime (that is $\gcd(m_i, m_j) = 1$ whenever $i \neq j$). Then the system of n equations

$$x = a_1 \pmod{m_1}$$

$$...$$

$$x = a_n \pmod{m_n}$$

has a unique solution for x modulo M where $M = m_1 ... m_n$.

Proof: This is an easy induction from the previous form of the theorem, or we can write down the solution directly.

Define $b_i = M/m_i$ (the product of all the moduli except for m_i) and $b_i' = b_i^{-1} \pmod{m_i}$. Then by a similar argument to before,

$$x = \sum_{i=1}^{n} a_i b_i b_i' \pmod{M}$$

is the unique solution.

5.2 Prime Powers First

An important consequence of the theorem is that when studying modular arithmetic in general, we can first study modular arithmetic a prime power and then appeal to the Chinese Remainder Theorem to generalize any results. For any integer n, we factorize n into primes $n = p_1^{k_1} ... p_m^{k_m}$ and then use the Chinese Remainder Theorem to get

$$\mathbb{Z}_n = \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_m^{k_m}}$$

To prove statements in $\mathbb{Z}_{p^k}$, one starts from $\mathbb{Z}_p$, and inductively works up to $\mathbb{Z}_{p^k}$. Thus the most important case to study is $\mathbb{Z}_p$.

6 Roots of Polynomials

Example: What are the roots of x^2-1 modulo some prime p?

Clearly, $x=\pm 1$ works, but are there any other solutions?

Suppose $(x+1)(x-1)=0 \pmod p$. As p is prime, we see p divides $(x+1)$ or p divides $(x-1)$. These cases correspond to the solutions we already have, so there are no more solutions.

More generally, we have the following:

Theorem: Let $f(x)$ be a polynomial over $\mathbb{Z}_p$ of degree n. Then $f(x)$ has at most n roots.

Proof: We induct. For degree 1 polynomials $ax+b$, we have the unique root $x=-ba^{-1}$.

Suppose $f(x)$ is a degree n with at least one root a. Then write $f(x)=(x-a)g(x)$ where $g(x)$ has degree $n-1$. Now $f(x)=0 \pmod p$ means $(x-a)g(x)$ is a multiple of p.

Since p is prime, p divides $(x-a)$, or p divides $g(x)$. In the former case, we have the root a, and the latter case is equivalent to saying $g(x)=0 \pmod p$, which by induction has at most $n-1$ solutions.

6.1 Composite Moduli

Let n be a product of distinct primes: $n=p_1...p_k$. The Chinese Remainder Theorem implies we can solve a polynomial $f(x)$ over each $\mathbb{Z}_{p_i}$ and then combine the roots together to find the solutions modulo n. This is because a root a of $f(x)$ in $\mathbb{Z}_n$ corresponds to

$$(a_1,...,a_k) \in \mathbb{Z}_{p_1} \times ... \times \mathbb{Z}_{p_k}$$

where each a_i is a root of $f(x)$ in $\mathbb{Z}_{p_i}$.

Example: Solve $x^2-1 \pmod{77}$.

x^2-1 has the solutions $\pm 1 \pmod 7$ and $\pm 1 \pmod{11}$ (since they are both prime), thus the solutions modulo 77 are the ones corresponding to:

- $x=1 \pmod 7, x=1 \pmod{11}$: this is $x=1 \pmod{77}$
- $x=1 \pmod 7, x=-1 \pmod{11}$: this is $x=43 \pmod{77}$
- $x=-1 \pmod 7, x=1 \pmod{11}$: this is $x=34 \pmod{77}$
- $x=-1 \pmod 7, x=-1 \pmod{11}$: this is $x=-1(=76) \pmod{77}$

Generalizing the last example, whenever N is the product of two distinct odd primes we always have four square roots of unity. (When one of the primes is 2 we have a degenerate case because $1=-1 \pmod 2$.) An interesting fact is that if we are told one of the non-trivial square roots, we can easily factorize N (how?).

In order to describe the solutions of a polynomial $f(x)$ over $\mathbb{Z}_n$ for any n, we need to find the roots of $f(x)$ over $\mathbb{Z}_{p^k}$ for prime powers p^k. We shall leave this for later.

6.2 Wilson's Theorem

Since the only square roots of 1 modulo p are ± 1 for a prime p, for any element $a \in \mathbb{Z}_p^*$, we have $a \neq a^{-1}$ unless $a=\pm 1$. Thus in the list $2,3,...,p-2$ we have each element and its inverse exactly once, hence $(p-1)!=-1 \pmod p$. On the other hand when p is composite, $(p-1)!$ is divisible by all the proper factors of p so we have:

Theorem: For an integer $p>1$ we have $(p-1)!=-1 \pmod p$ if and only if p is prime.

At first glance, this seems like a good way to tell if a given number is prime but unfortunately there is no known fast way to compute $(p-1)!$.

Wilson's Theorem can be used to derive similar conditions:

Theorem: For an odd integer $p > 1$, let $r = (p-1)/2$. Then

$$(-1)^r (r!)^2 = -1 \pmod{p}$$

if and only if p is prime.

Proof:

$$\begin{aligned}(p-1)! &= 1(p-1)2(p-2)...((p-1)/2)((p+1)/2) \\ &= 1(-1)2(-2)...r(-r) \\ &= (-1)^r (r!)^2\end{aligned}$$

and the result follows from Wilson's Theorem.

7 Units and the Totient Function

If $y \in \mathbb{Z}_n$ is invertible (that is, if y^{-1} exists), then we say y is a *unit*. The set of units of $\mathbb{Z}_n$ is denoted by $\mathbb{Z}_n^*$, or $\mathbb{Z}_n^\times$.

We know y is a unit if and only if y and n are coprime. So the size of $\mathbb{Z}_n^*$ is precisely the number of integers in $[1..n-1]$ that are coprime to n.

We write $\phi(n)$ for the number of elements of $\mathbb{Z}_n^*$. The function $\phi(n)$ is called the *Euler totient function*. Actually, it turns out to be convenient to have $\phi(1) = 1$, so we prefer to define $\phi(n)$ as the number of integers in $[1..n]$ coprime to n. (This agrees with our original definition except when $n = 1$.)

Examples: $\phi(6) = 2$ since among $[1..6]$ only 1 and 5 are coprime to 6, and thus are the only units in $\mathbb{Z}_6$.

Let p be a prime. Then every nonzero element $a \in \mathbb{Z}_p$ is coprime to p (and hence a unit) thus we have $\phi(p) = p-1$.

How about powers of primes? If p is a prime, then the only numbers not coprime to p^k are the multiples of p, and there are $p^k/p = p^{k-1}$ of these. Hence

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

Now let m, n be coprime, and let $x \in \mathbb{Z}_{mn}$. Let $a = x \pmod{m}$ and $b = x \pmod{n}$ (we reduce x modulo p and q). Then by the Chinese Remainder Theorem, x is a unit if and only if a and b are. Thus $\mathbb{Z}_{pq}^* = \mathbb{Z}_p^* \times \mathbb{Z}_q^*$.

Looking at the size of these sets gives this fact: for p, q coprime, we have

$$\phi(pq) = \phi(p)\phi(q).$$

(Thus ϕ is multiplicative.)

Putting this together with the previous statement $\phi(p^k) = p^k - p^{k-1}$ for prime p, we get that for any integer $n = p_1^{k_1}...p_m^{k_m}$ (where we have factorized n into primes) we have

$$\phi(n) = (p_1^{k_1} - p_1^{k_1-1})...(p_m^{k_m} - p_m^{k_m-1})$$

Often this formula is rewritten as

$$\phi(n) = n\left(1 - \frac{1}{p_1}\right)...\left(1 - \frac{1}{p_m}\right)$$

8 Modular Exponentiation

Suppose we are asked to compute 3^5 modulo 7. We could calculate $3^5 = 243$ and then reduce 243 mod 7, but a better way is to observe $3^4 = (3^2)^2$. Since $3^2 = 9 = 2$ we have $3^4 = 2^2 = 4$, and lastly

$$3^5 = 3^4 \times 3 = 4 \times 3 = 5 \pmod{7}.$$

The second way is better because the numbers involved are smaller.

This trick, known as repeated squaring, allows us to compute $a^k \bmod n$ using only $O(\log k)$ modular multiplications. (We can use the same trick when exponentiating integers, but then the multiplications are not modular multiplications, and each multiplication takes at least twice as long as the previous one.)

8.1 The Discrete Log Problem

Let us examine the behaviour of the successive powers of 3 modulo 7.

$$3^1 = 3$$

$$3^2 = 2$$

$$3^3 = 6$$

$$3^4 = 4$$

$$3^5 = 5$$

$$3^6 = 1$$

Note we compute each power by multiplying the previous answer by 3 then reducing modulo 7. Beyond this, the sequence repeats itself (why?):

$$3^7 = 3$$

$$3^8 = 2$$

and so on.

At a glance, the sequence $3,2,6,4,5,1$ seems to have no order or structure whatsoever. In fact, although there are things we can say about this sequence (for example, members three elements apart add up to 7), it turns out that so little is known about the behaviour of this sequence that the following problem is difficult to solve efficiently:

(The discrete log problem) Let p be a prime, and g,h be two elements of $\mathbb{Z}_p^*$. Suppose $g^x = h \pmod{p}$. Then what is x?

Example: One instance of the discrete log problem: find x so that $3^x = 6 \pmod{7}$. (Answer: $x = 3$. Strictly speaking, any $x = 3 \pmod{6}$ will work.)

Recall when we first encountered modular inversion we argued we could try every element in turn to find an inverse, but this was too slow to be used in practice. The same is true for discrete logs: we could try every possible power until we find it, but this is impractical.

Euclid's algorithm gave us a fast way to compute inverses. However no fast algorithm for finding discrete logs is known. The best discrete log algorithms are faster than trying every element, but are not polynomial time.

8.2 Nonunits

Why don't we bother studying the behaviour of nonunits under exponentiation?

First consider when $n = p^k$ for some prime p. Then $a \in \mathbb{Z}_n$ is nonunit exactly when $\gcd(a,n) > 1$, which in this case means $a = dp$ for some d.

We have $a^k = d^k p^k = 0$, thus in at most k steps we hit zero, which is uninteresting (at least for our purposes).

In general, write $n = p_1^{k_1} ... p_n^{k_m}$. By the Chinese Remainder Theorem we have

$$\mathbb{Z}_n = \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_m^{k_m}}$$

Thus an element $a \in \mathbb{Z}_n$ corresponds to some element $(a_1, ..., a_m)$ on the right-hand side, and a is a nonunit if at least one of the a_i is a multiple of p_i. From above, this means in at most k_i steps, the *ith* member will reach zero, so in general, for some k, each a_i^k is zero or a unit, hence we can restrict our study to units.

Note we have again followed an earlier suggestion: we handle the prime power case first and then generalize using the Chinese Remainder Theorem.

9 The Order of a Unit

The discrete log problem may be hard, but we do know some facts about the powers of a unit $a \in \mathbb{Z}_n^*$. Firstly, $a^k = 1$ for some k: since there are finitely many units, we must have $a^x = a^y$ for some $x < y$ eventually, and since a^{-1} exists we find $a^{y-x} = 1$.

Let $a \in \mathbb{Z}_n^*$. The smallest positive integer x for which $a^x = 1 \pmod{n}$ is called the *order* of a. The sequence $a, a^2, \ldots$ repeats itself as soon as it reaches $a^x = 1$. (since $a^{x+k} = a^k$), and we have $a^k = 1$ precisely when k is a multiple of x.

Example: The powers of $3 \pmod{7}$ are $3, 2, 6, 4, 5, 1$ so the order of $3 \pmod{7}$ is 6.

The following theorems narrow down the possible values for the order of a unit.

9.1 Fermat's Little Theorem

Theorem: Let p be a prime. Then $a^p = a \pmod{p}$ for any $a \in \mathbb{Z}_p$.

This theorem is often equivalently stated as $a^{p-1} = 1$ for nonzero a.

Proof: We first show an identity sometimes referred to as the freshman's dream: for a prime p, we have

$$(x+y)^p = x^p + y^p \pmod{p}.$$

This is immediate from the binomial expansion, because p divides every term except for two terms, that is

$$(x+y)^p = x^p + y^p + p(\ldots) = x^p + y^p \pmod{p}$$

By induction we have

$$(x_1 + \ldots + x_n)^p = x_1^p + \ldots + x_n^p$$

Thus if we write $a = 1 + \ldots + 1$ (where there are a 1s), we have

$$a^p = (1 + \ldots + 1)^p = 1^p + \ldots + 1^p = 1 + \ldots + 1 = a.$$

Fermat's Theorem gives an alternative way to compute inverses. For $a \in \mathbb{Z}_p^*$, a^{-1} can be computed as a^{p-2}, since we have $a \cdot a^{p-2} = 1$ by the theorem.

9.2 Euler's Theorem

Theorem: If $a \in \mathbb{Z}_n^*$ then $a^{\phi(n)} = 1 \pmod{n}$.

This reduces to Fermat's Little Theorem when n is prime.

Proof: Let $m = \phi(n)$, and label the units $u_1, \ldots, u_m$. Consider the sequence $au_1, \ldots, au_m$ (we multiply each unit by a). If $au_i = au_j$, then multiplying by a^{-1} (which exists since a is a unit) shows $u_i = u_j$, hence the members of the sequence are distinct. Furthermore products of units must also be units, hence $au_1, \ldots, au_m$ must be $u_1, \ldots, u_m$ in some order.

Multiplying all the units together gives

$$au_1 \ldots au_m = u_1 \ldots u_m.$$

Rearranging yields

$$a^{\phi(n)} = a^m = (u_1 \ldots u_m)(u_1 \ldots u_m)^{-1} = 1.$$

Similarly Euler's Theorem also gives an alternative way to compute inverses. For $a \in \mathbb{Z}_n^*$, a^{-1} can be computed as $a^{\phi(n)-1}$. This is efficient as we may use repeated squaring, but Euclid's algorithm is still faster (why?) and does not require one to compute $\phi(n)$.

These theorems do not tell us the order of a given unit $a \in \mathbb{Z}_n^*$ but they do narrow it down: let x be the order of a. If we know $a^y = 1$ by Euclid's algorithm we can find m, n such that

$$d = mx + ny$$

where $d = \gcd(x, y)$. Then

$$a^d = a^{mx+ny} = (a^x)^m (a^y)^n = 1$$

thus since $d \leq x$ we must have $d = x$ (since x is the smallest positive integer for which $a^x = 1$), and hence x must be a divisor of y. Thus by Euler's Theorem, the order of a divides $\phi(n)$.

These theorems are special cases of Lagrange's Theorem from group theory. (Fermat and Euler died long before group theory was discovered.)

9.3 Multiplication and Order

Let x be the order of $a \in \mathbb{Z}_n^*$, and y be the order of $b \in \mathbb{Z}_n^*$. What is the order of ab?

Suppose $(ab)^k = 1$. Raising both sides to x shows

$$b^{kx} = 1^k b^{kx} = (a^x)^k b^{kx} = (ab)^{kx} = 1.$$

Since b has order y we see that $y|kx$.

Suppose x, y are coprime. Then we must have $y|k$. Similarly $x|k$, hence k must be a multiple of xy. On the other hand, we have $(ab)^{xy} = 1$. Hence the order of ab is precisely xy: for elements with coprime orders, the order of their product is equal to the product of their orders.

More generally, let $d = \gcd(x,y)$. Then $x|ky$ and $y|kx$ implies that k must be a multiple of $(x/d)(y/d)$. If z is the least common multiple of x and y, which we can compute by $z = xy/d$, then $(ab)^z = 1$. All we can say at the moment is that the order of ab is a multiple of xy/d^2 and divides xy/d.

9.4 The RSA Problem

Suppose we are given positive integers e, N, and $a^e \pmod N$ for some unit a. How can we recover a?

One strategy is to find an integer d such that $a^{de} = a \pmod N$. By Euler's Theorem, d will satisfy this equation if $de = k\phi(N) + 1$ for some k. In other words, we compute

$$d = e^{-1} \pmod{\phi(N)}$$

and compute $(a^e)^d$ to recover a.

However it is not known how to compute $\phi(N)$ from N without factoring N, and it is not known how to factor large numbers efficiently.

10 Primality Tests

Given an integer n, how can we tell if n is prime? Assume n is odd, since the even case is trivial.

The most obvious idea is to look for factors of n, but no efficient factoring algorithm is known.

10.1 The Fermat Test

By Fermat's Theorem, if n is prime, then for any a we have $a^{n-1} = 1 \pmod n$. This suggests the Fermat test for a prime: pick a random $a \in [1..n-1]$ then see if $a^{n-1} = 1 \pmod n$. If not, then n must be composite.

However, equality may hold even when n is not prime. For example, take $n = 561 = 3 \times 11 \times 17$. By the Chinese Remainder Theorem

$$\mathbb{Z}_{561} = \mathbb{Z}_3 \times \mathbb{Z}_{11} \times \mathbb{Z}_{17}$$

thus each $a \in \mathbb{Z}_{561}^*$ corresponds to some

$$(x,y,z) \in \mathbb{Z}_3^* \times \mathbb{Z}_{11}^* \times \mathbb{Z}_{17}^*.$$

By Fermat's Theorem, $x^2 = 1$, $y^{10} = 1$, and $z^{16} = 1$. Since 2, 10, and 16 all divide 560, this means $(x,y,z)^{560} = (1,1,1)$, in other words, $a^{560} = 1$ for any $a \in \mathbb{Z}_{561}^*$.

Thus no matter what a we pick, 561 always passes the Fermat test despite being composite so long as a is coprime to n. Such numbers are called Carmichael numbers, and it turns out there are infinitely many of them.

If a is not coprime to n then the Fermat test fails, but in this case we can recover a factor of n by computing $\gcd(a,n)$.

10.2 The Miller-Rabin Test

We can do better by recalling n is prime if and only if the solutions of $x^2 = 1 \pmod{n}$ are $x = \pm 1$.

So if n passes the Fermat test, that is, $a^{n-1} = 1$, then we also check $a^{(n-1)/2} = \pm 1$, because $a^{(n-1)/2}$ is a square root of 1.

Unfortunately, numbers such as the third Carmichael number 1729 still fool this enhanced test. But what if we iterate? That is, so long as it's possible, we continue halving the exponent until we reach a number besides 1. If it's anything but -1 then n must be composite.

More formally, let 2^s be the largest power of 2 dividing $n-1$, that is, $n-1 = 2^s q$ for some odd number q. Each member of the sequence

$$a^{n-1} = a^{2^s q}, a^{2^{s-1} q}, \ldots, a^q.$$

is a square root of the preceding member.

Then if n is prime, this sequence begins with 1 and either every member is 1, or the first member of the sequence not equal to 1 is -1.

The Miller-Rabin test picks a random $a \in \mathbb{Z}_n$. If the above sequence does not begin with 1, or the first member of the sequence that is not 1 is also not -1 then n is not prime.

It turns out for any composite n, including Carmichael numbers, the probability n passes the Miller-Rabin test is at most $1/4$. (On average it is significantly less.) Thus the probability n passes several runs decreases exponentially.

If n fails the Miller-Rabin test with a sequence starting with 1, then we have a nontrivial square root of 1 modulo n, and we can efficiently factor n. Thus Carmichael numbers are always easy to factor.

Exercise: What happens when we run the Miller-Rabin test on numbers of the form pq where p, q are large primes? Can we break RSA with it?

Given n, find s so that $n-1 = 2^s q$ for some odd q. Then we implement a single Miller-Rabin test as follows:

1. Pick a random $a \in [1..n-1]$.
2. If $a^q = 1$ then n passes.
3. Otherwise, for $i = 0, \ldots, s-1$ see if $a^{2^i q} = -1$. If so, n passes.
4. Otherwise n is composite.

We also perform a few trial divisions by small primes before running the Miller-Rabin test several times.

Strictly speaking, these tests are *compositeness tests* since they do not prove the input is prime, but rather prove that an input is composite.

There exist deterministic polynomial-time algorithms for deciding primality (see Agrawal, Kayal and Saxena), though at present they are impractical.

11 Generators

A unit $g \in \mathbb{Z}_n^*$ is called a *generator* or *primitive root* of $\mathbb{Z}_n^*$ if for every $a \in \mathbb{Z}_n^*$ we have $g^k = a$ for some integer k. In other words, if we start with g, and keep multiplying by g eventually we see every element.

Example: 3 is a generator of $\mathbb{Z}_4^*$ since $3^1 = 3, 3^2 = 1$ are the units of $\mathbb{Z}_4^*$.

Example: 3 is a generator of $\mathbb{Z}_7^*$. From before the powers of 3 are $3, 2, 6, 4, 5, 1$ which are the units of $\mathbb{Z}_7^*$.

Example: 3 is not a generator of $\mathbb{Z}_{11}^*$ since the powers of 3 $\pmod{11}$ are $3, 9, 5, 4, 1$ which is only half of $\mathbb{Z}_{11}^*$.

Theorem: Let p be a prime. Then $\mathbb{Z}_p^*$ contains exactly $\phi(p-1)$ generators. In general, for every divisor $d | p-1$, $\mathbb{Z}_p^*$ contains $\phi(d)$ elements of order d.

Proof: by Fermat's Theorem we know the equation

$$x^{p-1} - 1 = 0 \pmod{p}$$

has $p-1$ distinct solutions, namely every element of $\mathbb{Z}_p^*$. Let q^k be a prime power dividing $p-1$. Then we can factorize the above

$$x^{p-1} - 1 = (x^{q^k} - 1)g(x) = 0 \pmod{p}$$

where $g(x)$ is some degree $p-1-q^k$ polynomial. From our notes on polynomials we know that $(x^{q^k}-1)$ has at most q^k roots and $g(x)$ has at most $p-1-q^k$ roots, and since their product has $p-1$ different roots, we see that there are exactly q^k distinct solutions to

$$x^{q^k} - 1 = 0 \pmod{p}.$$

Any solution a must have order dividing q^k. Suppose such an a has order less than q^k. Then its order must divide q^{k-1}, and a is a solution of $x^{q^{k-1}} - 1 = 0$.

By a similar argument $x^{q^{k-1}} - 1$ (which is a factor of $x^{q^k} - 1$) has exactly q^{k-1} distinct roots, so there must be exactly $q^k - q^{k-1}$ roots of $x^{q^k} - 1$ that are not roots of $x^{q^{k-1}}$ and hence are of order q^k. Recall $q^k - q^{k-1} = \phi(q^k)$.

Factorize $p-1$ into primes:

$$p - 1 = q_1^{k_1} ... q_n^{k_n}$$

For each $q_i^{k_i}$ we find an element a_i with order $q_i^{k_i}$. Since the orders of each a_i are coprime, the order of their product is equal to the product of their orders, that is $a_1 ... a_n$ has order $q_1^{k_1} ... q_n^{k_n} = p-1$ and thus is a generator.

We have $\phi(q_i^{k_i})$ choices for each a_i thus we have exactly $\phi(q_1^{k_1}) ... \phi(q_n^{k_n}) = \phi(p-1)$ different generators.

A similar argument proves the theorem for the divisors of $p-1$.

Alternative Proof: We can use a counting argument and basic facts about cyclic groups instead.

Any element $a \in \mathbb{Z}_p^*$ must have order dividing $p-1$ (by Fermat). Then if a has order d, the solutions of

$$x^d - 1 = 0 \pmod{p}.$$

are precisely $a, a^2, ..., a^d = 1$ and there are no other elements of order d since $x^d - 1$ has at most d roots.

It is easy to show that a^k has order d if and only if $\gcd(k,d) = 1$, thus either there are no elements of order d or there are exactly $\phi(d)$ elements of order d.

Now $\sum_{d|p-1} \phi(d) = p-1$ (which we can prove using multiplicative functions or cyclic groups) and if any of the $\phi(d)$ were replaced with 0 on the left-hand side, the equality would fail. Hence there must be exactly $\phi(d)$ elements of order d for each $d|p-1$ (since each one of the $p-1$ elements of $\mathbb{Z}_p^*$ must have some order).

What about powers of primes, or composite numbers in general?

12 Cyclic Groups

$\mathbb{Z}_n^*$ is an example of a *group*. We won't formally introduce group theory, but we do point out that a group only deals with one operation. The $*$ in $\mathbb{Z}_n^*$ stresses that we are only considering multiplication and forgetting about addition.

Notice we rarely add or subtract elements of $\mathbb{Z}_n^*$. For one thing, the sum of two units might not be a unit. We performed addition in our proof of Fermat's Theorem, but this can be avoided by using our proof of Euler's Theorem instead. We did need addition to prove that $\mathbb{Z}_n^*$ has a certain structure, but once this is done, we can focus on multiplication.

Let us see what can be said from studying multiplication alone.

When $\mathbb{Z}_n^*$ has a generator, we call $\mathbb{Z}_n^*$ a *cyclic group*. If g is a generator we write $\mathbb{Z}_n^* = \langle g \rangle$.

A *subgroup* of $\mathbb{Z}_n^*$ is a non-empty subset H of $\mathbb{Z}_n^*$ such that if $a, b \in H$, then $ab \in H$. Thus any subgroup contains 1, and also the inverse of every element in the subgroup. (Why? Hint: our definition of a subgroup only works when every element has a finite order; the real definition is different!)

Examples: Any $a \in \mathbb{Z}_n^*$ can be used to generate cyclic subgroup $\langle a \rangle = \{a, a^2, ..., a^d = 1\}$ (for some d). For example, $\langle 2 \rangle = \{2, 4, 1\}$ is a subgroup of $\mathbb{Z}_7^*$. Any group is always a subgroup of itself. $\{1\}$ is always a subgroup of any group. These last two examples are the *improper* subgroups of a group.

12.1 Lagrange's Theorem

We prove Lagrange's Theorem for $\mathbb{Z}_n^*$. The proof can easily be modified to work for a general finite group.

Our proof of Euler's Theorem has ideas in common with this proof.

Theorem: Let H be a subgroup of $\mathbb{Z}_n^*$ of size m. Then $m|\phi(n)$.

Proof: If $H = \mathbb{Z}_n^*$ then $m = \phi(n)$. Otherwise, let $H = \{h_1, ..., h_m\}$. let a be some element of $\mathbb{Z}_n^*$ not in H, and consider the set $\{h_1a, ..., h_ma\}$ which we denote by Ha. Every element in this set is distinct (since multiplying $h_ia = h_ja$ by a^{-1} implies $h_i = h_j$), and furthermore no element of Ha lies in H (since $h_i = h_ja$ implies $a = h_j^{-1}h_i$ thus $a \in H$, a contradiction).

Thus if every element of $\mathbb{Z}_n^*$ lies in H or Ha then $2m = \phi(n)$ and we are done. Otherwise take some element b in $\mathbb{Z}_n^*$ not in H or Ha. By a similar argument, we see that $Hb = \{h_1b, ..., h_mb\}$ contains exactly m elements and has no elements in common with either H or Ha.

Iterating this procedure if necessary, we eventually have $\mathbb{Z}_n^*$ as the disjoint union of the sets $H, Ha, Hb, ...$ where each set contains m elements. Hence $m|\phi(n)$.

Corollary: Euler's Theorem (and Fermat's Theorem). Any $a \in \mathbb{Z}_n^*$ generates a cyclic subgroup $\{a, a^2, ..., a^d = 1\}$ thus $d|\phi(n)$, and hence $a^{\phi(n)} = 1$.

12.2 Subgroups of Cyclic Groups

Theorem: All subgroups of a cyclic group are cyclic. If $G = \langle g \rangle$ is a cyclic group of order n then for each divisor d of n there exists exactly one subgroup of order d and it can be generated by $a^{n/d}$.

Proof: Given a divisor d, let $e = n/d$. Let g be a generator of G. Then $\langle g^e \rangle = \{g^e, g^{2e}, ..., g^{de} = 1\}$ is a cyclic subgroup of G of size n/d.

Now let $H = \{a_1, ..., a_{d-1}, a_d = 1\}$ be some subgroup of G. Then for each a_i, we have $a_i = g^k$ for some k. By Lagrange's Theorem the order of a_i must divide d, hence $g^{kd} = 1$.

Since the order of g is n, we have $kd = mn = mde$ for some m. Thus $k = em$ and $a_i = (g^e)^m$, that is each a_i is some power of g^e, hence H is one of the subgroups we previously described.

12.3 Counting Generators

Theorem: Let G be cyclic group of order n. Then G contains exactly $\phi(n)$ generators.

Proof: Let g be a generator of G, so $G = \{g, ..., g^n = 1\}$. Then g^k generates G if and only if $g^{km} = g$ for some m, which happens when $km = 1 \pmod{n}$, that is k must be a unit in $\mathbb{Z}_n$, thus there are $\phi(n)$ values of k for which g^k is a generator.

Example: When $\mathbb{Z}_n^*$ is cyclic (i.e. when $n = 2, 4, p^k, 2p^k$ for odd primes p), $\mathbb{Z}_n^*$ contains $\phi(\phi(n))$ generators.

We can now prove a theorem often proved using multiplicative functions:

Theorem: For any positive integer n

$$n = \sum_{d|n} \phi(d).$$

Proof: Consider a cyclic group G of order n, hence $G = \{g, ..., g^n = 1\}$. Each element $a \in G$ is contained in some cyclic subgroup. The theorem follows since there is exactly one subgroup H of order d for each divisor d of n and H has $\phi(d)$ generators.

12.4 Group Structure

In an abstract sense, for every positive integer n, there is only one cyclic group of order n, which we denote by C_n. This is because if g is a generator, then $C_n = \{g, g^2, ..., g^n = 1\}$ which completely determines the behaviour of C_n.

Example: Both $\mathbb{Z}_3^*$ and $\mathbb{Z}_4^*$ are cyclic of order 2, so they both behave exactly like C_2 (when considering multiplication only). This is an example of a *group isomorphism*.

Example: For $n = 2^k p_1^{k_1} ... p_m^{k_m}$ for odd primes p_i, by the Chinese Remainder Theorem we have

$$\mathbb{Z}_n^* = \mathbb{Z}_{2^k}^* \times \mathbb{Z}_{p_1^{k_1}}^* \times ... \times \mathbb{Z}_{p_m^{k_m}}^*$$

Recall each $\mathbb{Z}_{p_i^{k_i}}^*$ is cyclic, and so are $\mathbb{Z}_2^*$ and $\mathbb{Z}_4^*$. Also recall for $k > 2$ we have that $3 \in \mathbb{Z}_{2^k}^*$ has order 2^{k-2} and no element has a higher order. Using some group theory this means the group structure of $\mathbb{Z}_n^*$ is

$$C_{2^{k-1}} \times C_{p_1^{k_1} - p_1^{k_1-1}} \times ... \times C_{p_m^{k_m} - p_m^{k_m-1}}$$

when $k = 1, 2$ and

$$C_2 \times C_{2^{k-2}} \times C_{p_1^{k_1} - p_1^{k_1-1}} \times ... \times C_{p_m^{k_m} - p_m^{k_m-1}}$$

when $k > 2$.

13 Quadratic Residues

Let $a \in \mathbb{Z}_n$. We say a is a *quadratic residue* if there exists some x such that $x^2 = a$. Otherwise a is a *quadratic nonresidue*.

Efficiently distinguishing a quadratic residue from a nonresidue modulo $N = pq$ for primes p, q is an open problem. This is exploited by several cryptosystems, such as Goldwassser-Micali encryption, or Cocks identity-based encryption. More general variants of this problem underlie other cryptosystems such as Paillier encryption.

Let p be an odd prime, as the case $p = 2$ is trivial. Let g be a generator of $\mathbb{Z}_p^*$. Any $a \in \mathbb{Z}_p^*$ can be written as g^k for some $k \in [0..p-2]$.

Say k is even. Write $k = 2m$. Then $(g^m)^2 = a$, so a is a quadratic residue. Exactly half of $[0..p-2]$ is even (since p is odd), hence at least half of the elements of $\mathbb{Z}_p^*$ are quadratic residues.

Suppose we have $b^2 = a$. Then $(-b)^2 = a$ as well, and since $b \neq -b$ (since $p > 2$) every quadratic residue has at least two square roots (in fact, we know from studying polynomials there can be at most two), thus at most half the elements of $\mathbb{Z}_p^*$ are quadratic residues. (Otherwise there are more square roots than elements!)

Thus exactly half of $\mathbb{Z}_p^*$ are quadratic residues, and they are the even powers of g.

Given $a = g^k$, consider the effect of exponentiating by $(p-1)/2$. If k is odd, say $k = 2m+1$, we get

$$a^{(p-1)/2} = g^{(2m+1)(p-1)/2} = g^{(p-1)m} g^{(p-1)/2} = g^{(p-1)/2}$$

The square of $g^{(p-1)/2}$ is $g^{p-1} = 1$, so $g^{(p-1)/2}$ is 1 or -1. But g has order $p-1$ (g is a generator) thus we must have $g^{(p-1)/2} = -1$.

If k is even, say $k = 2m$, then $a^{(p-1)/2} = g^{(p-1)m} = 1$.

In other words, we have proved *Euler's Criterion*, which states a is a quadratic residue if and only if $a^{(p-1)/2} = 1$, and a is a quadratic nonresidue if and only if $a^{(p-1)/2} = -1$.

Example: We have -1 is a quadratic residue in $\mathbb{Z}_p$ if and only if $p = 1 \pmod 4$.

13.1 The Legendre Symbol

We define the *Legendre symbol* for odd primes p and integers a by

$$\left(\frac{a}{p}\right) = a^{(p-1)/2}.$$

The symbol can also be written as $(a|p)$.

Thus:

- $\left(\frac{a}{p}\right) = 1$ if a is a quadratic residue of p

- $\left(\frac{a}{p}\right) = -1$ if a is a quadratic nonresidue of p
- $\left(\frac{a}{p}\right) = 0$ if $a = 0$ modulo p

For any integer b

$$\left(\frac{a}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right)$$

and for any r coprime to p

$$\left(\frac{ar^2}{p}\right) = \left(\frac{a}{p}\right)$$

14 Gauss' Lemma

There is a less obvious way to compute the Legendre symbol. Among other things, we can use it to easily find $\left(\frac{2}{p}\right)$. Before stating the method formally, we demonstrate it with an example.

Let $p = 17$, and $a = 7$. There are 16 nonzero elements $[1..16]$. Consider the first half $[1..8]$ and multiply them all by 7 to get 7, **14**, 4, **11**, 1, 8, **15**, 5. We've singled out 14, 11 and 15 because they are greater than $p/2$ (that is, 9 or higher).

So exactly 3 of them are greater than $p/2$. Gauss' Lemma states that if we take this 3 and raise -1 to this power, then we have $\left(\frac{a}{p}\right)$, that is:

$$\left(\frac{7}{17}\right) = (-1)^3 = -1.$$

Theorem (Gauss' Lemma): Let p be an odd prime, q be an integer coprime to p. Take the least residues of $\{q, 2q, ..., q(p-1)/2\}$, that is, reduce them to integers in $[0..p-1]$. Let u be the number of members in this set that are greater than $p/2$. Then

$$\left(\frac{q}{p}\right) = (-1)^u.$$

Proof: Let $b_1, ..., b_t$ be the members of the set less than $p/2$, and $c_1, ..., c_u$ be the members greater than $p/2$. Then $u + t = (p-1)/2$. Consider the sequence

$$0 < b_1, ..., b_t, p - c_1, ..., p - c_u < p/2$$

Each of these are distinct: clearly $b_i \neq b_j$ and $c_i \neq c_j$ whenever $i \neq j$ (since q is invertible), and if $b_i = p - c_j$, then let $b_i = rq, c_j = sq$. Then $r + s = 0$, which is a contradiction since

$$0 < r, s < p/2.$$

Hence they must be precisely the numbers $1, ..., (p-1)/2$ in some order, thus

$$\begin{aligned} q(2q)...(q(p-1)/2) &= b_1...b_t c_1...c_u \\ &= (-1)^u b_1...b_t(p-c_1)...(p-c_u) \\ &= (-1)^u \left(\frac{p-1}{2}\right)! \end{aligned}$$

Dividing both sides by $((p-1)/2))!$ completes the proof.

For example, let p be an odd prime and take $q = 2$. The sequence $2, 2\times 2, ..., 2(p-1)/2$ consists of positive least residues. We have $p = 8x + y$ for some integer x and $y \in \{1, 3, 5, 7\}$. By considering each case we see that the number of elements greater than $p/2$ is even when $p = 1, 7 \pmod 8$ and odd when $p = 3, 5 \pmod 8$. We restate this as follows.

Theorem: Let p be an odd prime.

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = (-1)^{\lfloor \frac{p+1}{4} \rfloor}$$

Proof: See the last paragraph, and note that $(p^2-1)/8$ is even when $p = \pm 1 \pmod 8$ and odd otherwise. Similarly for $\lfloor (p+1)/4 \rfloor$.

Example: By Gauss' Lemma, $\left(\frac{3}{p}\right) = 1$ if $p = \pm 1 \pmod{12}$ and -1 otherwise (that is, $p = \pm 5 \pmod{12}$). Combining this with the above result for $-1 \pmod p$ we have $\left(\frac{-3}{p}\right) = 1$ if $p = 1 \pmod 6$ and -1 otherwise ($p = -1 \pmod 6$).

15 Quadratic Reciprocity

The law of *quadratic reciprocity*, noticed by Euler and Legendre and proved by Gauss, helps greatly in the computation of the Legendre symbol.

First, we need the following theorem:

Theorem: Let p be an odd prime and q be some odd integer coprime to p. Let $m = \lfloor q/p \rfloor + \lfloor 2q/p \rfloor ... \lfloor ((p-1)/2)q/p \rfloor$.

Then $m = u \pmod 2$, where as in Gauss' Lemma, u is the number of elements of $\{q, 2q, ..., q(p-1)/2\}$ which have a residue greater than $p/2$.

Proof: For each $i = 1, ..., (p-1)/2$, the equation $iq = p\lfloor iq/p \rfloor + r_i$ holds for some $0 < r_i < p$. Reading the proof of Gauss' Lemma, we see these are precisely the b_i and c_i.

Then summing these equations modulo 2 gives

$$\begin{aligned} q(p^2-1)/8 = &\quad pm + b_1 + ... + b_t + c_1 + ... + c_u \\ = &\quad pm + b_1 + ... + b_t + up + (p - c_1) + ... + (p - c_u) \\ = &\quad pm + up + 1 + 2 + ... + (p-1)/2 \\ = &\quad pm + up + (p^2-1)/8 \end{aligned}$$

That is,

$$(q-1)(p^2-1)/8 = p(m+u)$$

Since p and q are odd, we have $m = u \pmod 2$.

Theorem (Law of Quadratic Reciprocity): Let p, q be distinct odd primes. If $p = q = -1 \pmod 4$, then

$$\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$$

otherwise

$$\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right).$$

which we can state as

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

Proof: From above, we need only show

$$m + n = (p-1)(q-1)/4$$

where $m = \lfloor q/p \rfloor + \lfloor 2q/p \rfloor ... \lfloor ((p-1)/2)q/p \rfloor$, and n is similarly defined by swapping p and q.

Eisenstein found an elegant geometrical proof. Consider the line L from $(0,0)$ to (p,q), and the rectangle R with corners at $(0,0)$ and $(p/2, q/2)$. How many lattice points lie strictly in R?

Simply computing the area of a rectangle gives $(p-1)(q-1)/4$. Alternatively, we can count the number of points above and below L inside R, since no lattice points can lie on L in R since p, q are coprime.

Consider the points below L on the line $x = 1$. They have y-coordinates of $1, 2, ..., \lfloor q/p \rfloor$. When $x = 2$, there are $\lfloor 2q/p \rfloor$ points, and so on, giving a total of m points below the L. Similarly there are n points above L in R, proving the result.

We can restate the proof algebraically. Consider the numbers $py - qx$ for $x = 1, ..., (p-1)/2$ and $y = 1, ..., (q-1)/2$. There are a total of $(p-1)(q-1)/4$ numbers, not necessarily distinct. None are zero, since p, q are coprime. The result follows after observing that n of them are positive, and m are negative.

Example:

$$\left(\frac{31}{103}\right) = -\left(\frac{103}{31}\right) = -\left(\frac{10}{31}\right) = -\left(\frac{2}{31}\right)\left(\frac{5}{31}\right) = -(-1)\left(\frac{5}{31}\right)$$

since $2 = -5 \pmod 8$. Next,

$$\left(\frac{5}{31}\right) = -\left(\frac{31}{5}\right) = -\left(\frac{1}{5}\right) = -1.$$

Hence 31 is a quadratic nonresidue modulo 103.

This method is flawed because it relies on factoring, so we might think we should stick to our original modular exponentiation for computing the Legendre symbol. But it turns out all is well once we extend the Legendre symbol.

15.1 The Jacobi Symbol

The *Jacobi symbol* $\left(\frac{a}{b}\right)$ is defined for all odd positive integers b and all integers a. When b is prime, it is equivalent to the Legendre symbol. If $b = 1$, define $\left(\frac{a}{1}\right) = 1$. Lastly, for other values of b, factor b into primes: $b = p_1^{k_1} \ldots p_n^{k_n}$ and define

$$\left(\frac{a}{b}\right) = \left(\frac{a}{p_1}\right)^{k_1} \ldots \left(\frac{a}{p_n}\right)^{k_n}$$

Thus for odd positive integers b, b_1, b_2 we have

$$\left(\frac{a}{b_1 b_2}\right) = \left(\frac{a}{b_1}\right)\left(\frac{a}{b_2}\right)$$

Other properties of the Legendre symbol carry over. By inducting on the number of primes in the factorization of b, one can show:

$$\left(\frac{-1}{b}\right) = (-1)^{(b-1)/2}$$

$$\left(\frac{2}{b}\right) = (-1)^{(b^2-1)/8}$$

$$\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = (-1)^{\frac{a-1}{2}\frac{b-1}{2}}$$

The last property allows us to compute the Jacobi symbol without factoring.

Example:

$$\left(\frac{31}{103}\right) = -\left(\frac{103}{31}\right) = -\left(\frac{-21}{31}\right) = -\left(\frac{-1}{31}\right)\left(\frac{21}{31}\right)$$
$$= \left(\frac{31}{21}\right) = \left(\frac{-11}{21}\right) = \left(\frac{-1}{21}\right)\left(\frac{11}{21}\right) = \left(\frac{21}{11}\right) = \left(\frac{-1}{11}\right) = -1.$$

Hence 31 is a quadratic nonresidue modulo 103.

16 Carmichael Numbers

Recall Carmichael numbers are composite numbers that almost always fool the Fermat primality test. We can show that Carmichael numbers must have certain properties.

First we show they cannot be of the form $n = pq$ where p, q are distinct primes with $p > q$. By the Chinese Remainder Theorem we have $\mathbb{Z}_n = \mathbb{Z}_p \times \mathbb{Z}_q$. Then

$$n - 1 = pq - 1 = q(p-1) + q - 1.$$

Suppose a is not a multiple of p. By Fermat

$$a^{n-1} = (a^{p-1})^q a^{q-1} = a^{q-1} \pmod{p}$$

Then if a passes the Fermat test, we must have $a^{q-1} = 1 \pmod{p}$ and hence $a^d = 1$, where $d = \gcd(p-1, q-1)$. Since $\mathbb{Z}_p^*$ is cyclic, there are exactly d choices for a that satisfy $a^d = 1$.

Since $p > q$, the greatest common divisor of $p-1$ and $q-1$ is at most $(p-1)/2$. Thus at most half the choices for a can fool the Fermat test.

Next suppose n is not squarefree, that is $n = p^k r$ for some prime p and $k \geq 2$. Then by the Chinese Remainder Theorem, $\mathbb{Z}_n = \mathbb{Z}_{p^k} \times \mathbb{Z}_r$. Any $a \in \mathbb{Z}_{p^k}$ satisfies $a^{\phi(p^k)} = 1$ by Euler's Theorem, so if $a^{n-1} = 1$ as well, then we must have $a^d = 1$ where

$$d = \gcd(n-1, \phi(p^k)) = \gcd(p^k r - 1, p^{k-1}(p-1)).$$

Since $\mathbb{Z}_{p^k}$, is cyclic, exactly d elements $a \in \mathbb{Z}_{p^k}$ satisfy $a^d = 1$. As p cannot divide $p^k r - 1$, the largest possible value for d is $p-1$, giving an upper bound of $(p-1)/(p^k - 1) \leq 1/4$ probability that n will pass the Fermat test.

Hence if n is a Carmichael number, then n is squarefree and is the product of at least three distinct primes.

16.1 Solovay-Strassen Test

Recall our first suggestion for improving the Fermat test was to check a candidate x satisfies $x^{(n-1)/2} = \pm 1$ after checking $x^{n-1} = 1$, since there are no nontrivial square roots of unity modulo a prime.

We can improve this by checking instead that

$$x^{(n-1)/2} = \left(\frac{x}{n}\right).$$

This is known as the Solovay-Strassen test. Recall that the Jacobi symbol can be evaluated quickly using quadratic reciprocity.

Why does this help? From above we know that if n is not squarefree then n fails the Fermat test with probability at least $3/4$. So we need only consider the case when n is squarefree but composite, say $n = pr$ where p is prime and r is an odd number greater than 1.

By the Chinese Remainder Theorem any $x \in \mathbb{Z}_n$ can be written as $(a,b) \in \mathbb{Z}_p \times \mathbb{Z}_r$. For any nonzero $a \in \mathbb{Z}_p$ we have $a^{p-1} = 1$ by Fermat, so if $a^{n-1} = 1$ then we have $a^d = 1$ where $d = \gcd(p-1, n-1)$.

If $d < p-1$ then d is at most $(p-1)/2$. Since $\mathbb{Z}_p^*$ is cyclic, at most $(p-1)/2$ elements of $a \in \mathbb{Z}_p$ satisfy $a^d = 1$. That is, the probability a passes the Fermat test is at most $1/2$.

On the other hand, if $d = p-1$, then this implies $p-1|n-1$. Since $n-1 = pr-1 = (p-1)r + r - 1$ we have $p-1|r-1$, so write $r-1 = s(p-1)$. Then

$$a^{pr-1} = a^{(p-1)r} a^{(p-1)s} = 1.$$

But we have

$$\left(\frac{x}{n}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{r}\right)$$

As r contains at least one odd prime factor, the sign of $(b|r)$ is positive or negative with the same probability, that is, there is a $1/2$ chance that x fails the test.

In other words, for any composite n, including Carmichael numbers, the probability n passes the Solovay-Strassen test is at most $1/2$.

This was the first algorithm discovered for finding large primes. The Miller-Rabin test surpasses the Solovay-Strassen test in every way: the probability a composite number n passes is only $1/4$, and no Jacobi symbol computations are required. Moreover, any a that exposes the compositeness of n in the Solovay-Strassen test also triggers the Miller-Rabin test.

17 Multiplicative Functions

An *arithmetical function*, or 'number-theoretic function' is a complex-valued function defined for all positive integers. It can be viewed as a sequence of complex numbers.

Examples: $n!, \phi(n), \pi(n)$ which denotes the number of primes less than or equal to n.

An arithmetical function is *multiplicative* if $f(mn) = f(m)f(n)$ whenever $\gcd(m,n) = 1$, and *totally multiplicative* or *completely multiplicative* if this holds for any m, n. Thus $f(1) = 1$ unless f is the zero function, and a multiplicative function is completely determined by its behaviour on the prime powers.

Examples: We have seen that the Euler totient function ϕ is multiplicative but not totally multiplicative (this is one reason it is convenient to have $\phi(1) = 1$). The function n^2 is totally multiplicative. The product of (totally) multiplicative functions is (totally) multiplicative.

Theorem: Let $f(n)$ be a multiplicative function, and define

$$F(n) = \sum_{d|n} f(d).$$

Then $F(n)$ is also a multiplicative function.

Proof: Let m,n be positive integers with $\gcd(m,n)=1$. Then

$$\begin{aligned} F(mn) &= \sum_{d|mn} f(d) \\ &= \sum_{r|m,s|n} f(rs) \\ &= \sum_{r|m,s|n} f(r)f(s) \\ &= \sum_{r|m} f(r) \sum_{s|n} f(s) \\ &= F(m)F(n) \end{aligned}$$

since $r|m$ and $s|n$ implies $(r,s)=1$.

We just wrote $F(n)$ in terms of $f(d)$. Can we do the reverse? That is, write $f(n)$ in terms of $F(d)$? Yes; this is known as Möbius inversion.

Since ϕ is multiplicative, we have that

Theorem:

$$\sum_{d|n} \phi(n) = n$$

This theorem can also be proved using basic facts about cyclic groups.

Examples: The divisors of 12 are $1,2,3,4,6,12$. Their totients are $1,1,2,2,2,4$ which sum to 12.

The function $f(n)=1$ is (totally) multiplicative. Let $d(n)$ be the number of divisors of n. Then since $d(n)=\sum_{d|n} 1$ we see that $d(n)$ is multiplicative.

The function $f(n)=n$ is (totally) multiplicative. Let $\sigma(n)$ be the sum of divisors of n. Then since $\sigma(n)=\sum_{d|n} n$ we see that $\sigma(n)$ is multiplicative. In general, we can apply this trick to any power of n.

17.1 Perfect Numbers

A positive integer n is a *perfect number* if $\sigma(n)=2n$. The first perfect numbers are $6,28,496,8128$.

It is not known if any odd perfect numbers exist.

Let n be an even perfect number, so $n=2^{q-1}m$ for some $q>1$ and odd m. Since σ is multiplicative,

$$2^q m = 2n = \sigma(n) = \sigma(2^{q-1})\sigma(m) = (2^q-1)\sigma(m)$$

hence $2^q|\sigma(m)$, so write $\sigma(m)=2^q s$ for some s and hence $(2^q-1)s=m$.

Thus two of the divisors of m are s and $m=(2^q-1)s$. But these already sum to $2^q s$, hence m can have no other divisors, implying that $s=1$ and $m=2^q-1$ is prime. The converse is clear, thus n is an even perfect number if and only if

$$n = 2^{q-1}(2^q-1)$$

with 2^q-1 prime.

This implies q is prime as $d|q$ implies $2^d-1|2^q-1$. The converse is unsurprisingly false. A number of the form 2^q-1 is called a *Mersenne number*, and if it is prime, then it is called a *Mersenne prime*.

Mersenne conjectured that for $q \leq 257$ the only primes q which yielded primes 2^q-1 were $2,3,5,7,13,17,19,31,67,127,257$. He made five mistakes: $67,257$ should not be on the list, and he missed $61,89,107$.

Modulo 10, the powers of 2 cycle through 2, 4, 8, 6, hence $n=2^{q-1}(2^q-1)$ cycles through $2(4-1),4(8-1),8(6-1),6(2-1)$. The third of these is 0, implying 5 divides n which means n cannot be perfect, because 5 is not one below a power of 2. The other possibilities imply every even perfect number ends with 6 or 8.

Applying a similar but more exhausting calculation modulo 100 for $q>2$, we find even perfect numbers other than 6 must end with 28 or an odd digit followed by 6.

17.2 Fermat Numbers

What about numbers of the form 2^r+1?

It is not hard to see that if 2^r+1 is prime then r must be a power of 2. Numbers of the form $2^{2^m}+1$ are called *Fermat numbers*, and are called *Fermat primes* if prime. Fermat conjectured that all Fermat numbers are prime, and he was right for $m=0,\ldots,4$ (which give the primes $3,5,17,257,65537$), but wrong for $m=5$ (which is divisible by 641) and other values of m. In fact no other Fermat primes are known.

It can be shown that a regular n-gon that can be constructed using a ruler and compass if and only if

$$n=2^k p_1 \ldots p_m$$

where each p_i is a distinct Fermat prime and k is some nonnegative integer. In 1796, Gauss proved the sufficiency of this condition (though not the necessity) when he was only 19.

18 Möbius Inversion

Suppose for some (not necessarily multiplicative) number-theoretic function f

$$F(n)=\sum_{d|n} f(d).$$

Can we make $f(n)$ the subject of this equation?

We'll see that we can find a function μ such that

$$f(n)=\sum_{d|n}\mu(n/d)F(d)=\sum_{d|n}\mu(d)F(n/d).$$

and we call this process 'Möbius inversion'.

We have

$$\begin{aligned}\sum_{d|n}\mu(d)F(n/d) &= \sum_{d|n}\mu(d)\sum_{r|\frac{n}{d}} f(r)\\ &= \sum_{dr|n}\mu(d)f(r)\\ &= \sum_{r|n} f(r)\sum_{d|(n/r)}\mu(d)\end{aligned}$$

If we want this equal to $f(n)$ we need μ to satisfy

$$\sum_{d|m}\mu(d)=\begin{cases}0, & m>1\\ 1, & m=1\end{cases}$$

A little thought leads to this unique solution, known as the 'Möbius function':

$$\mu(n)=\begin{cases}1 & n=1\\ 0 & p^2|n \text{ for some prime } p\\ (-1)^r & n=p_1\ldots p_r \text{ for distinct primes } p_i\end{cases}$$

Notice μ is multiplicative, which implies $f(n)$ is multiplicative if $F(n)$ is. In summary,

Theorem:

$$F(n)=\sum_{d|n} f(d)$$

if and only if

$$f(n)=\sum_{d|n}\mu(n/d)F(d)$$

and $f(n)$ is multiplicative if and only if $F(n)$ is multiplicative.

Example: From before $n = \sum_{d|n} \phi(n)$. Write $n = p_1^{k_1} ... p_m^{k_m}$. Then

$$\begin{aligned}
\phi(n) &= \sum_{d|n} \mu(d) \frac{n}{d} \\
&= n \sum_{d|n} \mu(d) \frac{1}{d} \\
&= n \left(1 - \sum_i \frac{1}{p_i} + \sum_{i \neq j} \frac{1}{p_i p_j} - ... \right) \\
&= n \left(1 - \frac{1}{p_1} \right) ... \left(1 - \frac{1}{p_m} \right)
\end{aligned}$$

which is another way to derive the formula for ϕ.

Gauss encountered the Möbius function over 30 years before Möbius when he showed that the sum of the generators of $\mathbb{Z}_p^*$ is $\mu(p-1)$. More generally, if $\mathbb{Z}_n^*$ has a generator, then the sum of all the generators of $\mathbb{Z}_n^*$ is $\mu(\phi(n))$. This can be seen by considering the sums of the roots of polynomials of the form $x^d - 1$ where $d | \phi(n)$.

19 Generators: The General Case

We previously studied generators of $\mathbb{Z}_n^*$ for prime n. How do we generalize to any n?

We follow the most obvious strategy: first consider prime powers, then use the Chinese Remainder Theorem to generalize.

19.1 Powers of Two

First we see that 1 is a generator for $\mathbb{Z}_2^*$ and 3 is a generator for $\mathbb{Z}_4^*$. A quick check reveals $\mathbb{Z}_8^*$ has no generator: the square of any odd number is 1 modulo 8.

Next suppose $\mathbb{Z}_{2^k}^*$ has a generator g for some $k > 3$. Then for each $a \in \mathbb{Z}_8^*$, we have $g^x = a \pmod{2^k}$ for some x. This equation still holds modulo 8 since $8 | 2^k$. But this is a contradiction since it would imply g is a generator of $\mathbb{Z}_8^*$.

Thus if n is a power of 2, $\mathbb{Z}_n^*$ has a generator if and only if $n = 2$ or $n = 4$.

We examine the behaviour of units under exponentiation modulo a power of two more closely. By induction we can show that

$$(1+2n)^{2^{t-3}} = 1 + 2^{t-2}(n - n^2 + 2n^4) \pmod{2^t}$$

(first explicitly compute for $t = 4, 5$ before proving the inductive step).

From here one can show that if $a = \pm 3 \pmod 8$ then the order of a modulo 2^t is 2^{t-2}, otherwise the order is strictly smaller.

19.2 Squares of Odd Primes

Let p be an odd prime. Let g be a generator of $\mathbb{Z}_p^*$. We try to find a generator of $\mathbb{Z}_{p^2}^*$.

Intuitively, such a generator ought to relate to the generators of $\mathbb{Z}_p^*$. So consider the problem of finding the order of $g + kp$ in $\mathbb{Z}_{p^2}^*$ for any k.

Let t be the order of $g + kp$ in $\mathbb{Z}_{p^2}^*$. From $(g+kp)^t = g^t = 1 \pmod p$ we deduce $(p-1) | t$. Since $t | \phi(p^2) = p(p-1)$, there are two possibilities. Either $t = p - 1$ or $t = p(p-1)$. In the latter case, $g + kp$ generates $\mathbb{Z}_{p^2}^*$.

But the former case $(g+kp)^{p-1} = 1 \pmod{p^2}$ occurs if and only if

$$(g+kp)^p = g + kp \pmod{p^2}.$$

Considering the binomial expansion,

$$g^p - g = kp \pmod{p^2}.$$

As p is not invertible, we go back to first principles and find

$$k = (g^p - g)/p \pmod{p}$$

(the division is an integer division). In other words, $g + kp$ is a generator of $\mathbb{Z}_{p^2}^*$ except for one value of k. Thus each of the $\phi(p-1)$ generators of $\mathbb{Z}_p^*$ can be used to construct $p-1$ generators of $\mathbb{Z}_{p^2}^*$.

Alternative proof: we show that at least one of g or $g+p$ is a generator:

$$\begin{aligned}((g+p)^{p-1} - 1) - (g^{p-1} - 1) &= (g+p)^{p-1} - g^{p-1} \\ &= (g+p-g)((g+p)^{p-2} + (g+p)^{p-3}g + ... + g^{p-2}) \\ &= p((p-1)g^{p-2} + p(...))\end{aligned}$$

by considering the binomial expansion. Thus:

$$((g+p)^{p-1} - 1) - (g^{p-1} - 1) = p(p-1)g^{p-2} = -pg^{p-2} \pmod{p^2}$$

Since this is nonzero, we see that $g+p$ and g cannot both be roots of the polynomial $x^{p-1} - 1$, hence at least one of them has order $p(p-1)$.

19.3 Powers of Odd Primes

Let g be a generator of $\mathbb{Z}_{p^2}^*$. We show g is a generator for all powers of p.

First, a lemma. Suppose for some $s \geq 1$:

$$g^t = 1 + kp^s \pmod{p^{s+1}}$$

Then for any $r \geq 0$, by considering the binomial expansion and inducting on r, we have:

$$g^{tp^r} = 1 + kp^{s+r} \pmod{p^{s+r+1}}.$$

We use this lemma to show g also generates $\mathbb{Z}_p^*$ (which implies the generators we constructed above are the only generators of $\mathbb{Z}_{p^2}^*$, though we could also show this via a counting argument). If g has order t modulo p, then:

$$g^t = 1 + kp$$

for some k, and setting $s = 1, r = 1$ in the lemma yields:

$$g^{tp} = 1 \pmod{p^2}$$

Since g generates $\mathbb{Z}_{p^2}^*$, the exponent tp must be a multiple of $\phi(p^2) = p(p-1)$. Therefore t is a multiple of $p-1 = \phi(p)$ so g generates $\mathbb{Z}_p^*$. (We can generalize to show any generator for a prime power is also a generator for any lower power.)

Using the lemma is somewhat gratuitous as we could have reasoned as above when we showed a generator of a power of 2 must be a generator of lower power of 2, but on the other hand it's satisfying to use one lemma to dispatch all the cases.

Now for the other powers of p. We just established g generates $\mathbb{Z}_p^*$, that is:

$$g^{p-1} = 1 + kp$$

for some k, which must be coprime to p as g generates $\mathbb{Z}_{p^2}^*$.

Applying the lemma with $t = p-1, s = 1$ gives:

$$g^{(p-1)p^r} = 1 + kp^{1+r} \pmod{p^{r+2}}$$

When $r = 1$, this is:

$$g^{\phi(p^2)} = 1 + kp^2 \pmod{p^3}$$

Let t be the order of g modulo p^3 so $t | \phi(p^3)$. We also have:

$$g^t = 1 \pmod{p^2}$$

which means $\phi(p^2) | t$ because g generates $\mathbb{Z}_{p^2}^*$. Thus t is either $\phi(p^2)$ or $\phi(p^3)$. It cannot be the former since k is coprime to p, thus it must be the latter and thus g generates $\mathbb{Z}_{p^3}^*$.

We iterate this argument for higher powers of p.

19.4 The General Case

We first consider odd n. Write $n = p_1^{k_1}...p_m^{k_m}$. By the Chinese Remainder Theorem we have

$$\mathbb{Z}_n^* = \mathbb{Z}_{p_1^{k_1}}^* \times ... \times \mathbb{Z}_{p_m^{k_m}}^*$$

Each $x \in \mathbb{Z}_n^*$ corresponds to some element $(x_1,...,x_n)$ of the right-hand side. Now each x_i satisfies

$$x_i^{\phi(p_i^{k_i})} = 1 \pmod{p_i^{k_i}}$$

so if we take the lowest common multiple of the orders

$$\lambda(n) = \text{lcm}(\phi(p_1^{k_1}),...,\phi(p_m^{k_m}))$$

we find $(x_1,...,x_n)^\lambda = (1,...,1)$, thus x has order dividing $\lambda(n)$. On the other hand, if we choose each g_i to be a generator of $\mathbb{Z}_{p_i^{k_i}}$ then $(g_1,...,g_n)$ has order $\lambda(n)$.

Hence a generator exists in $\mathbb{Z}_n^*$ if and only if $\lambda(n) = \phi(n)$. Since each $\phi(p_i^{k_i})$ is even, $\lambda(n) = \phi(n)$ can only when $m = 1$, that is, n must be a prime power.

Now suppose $n = 2^k q$ where q is odd. Again by the Chinese Remainder Theorem we have $\mathbb{Z}_n^* = \mathbb{Z}_{2^k}^* \times \mathbb{Z}_q^*$. If $k > 2$ then $\mathbb{Z}_n^*$ has no generator since $\mathbb{Z}_8^*$ doesn't. If $k = 2$, since $\mathbb{Z}_4^*$ has order 2, the lowest common multiple of $\phi(4)$ and $\phi(q)$ is less than $\phi(4q)$, thus no generator exists. Lastly if $k = 1$ then if g is a generator for $\mathbb{Z}_q$ then $\lambda(n) = \phi(n)$ thus a generator exists ($(1,g)$ is a generator).

In summary, $\mathbb{Z}_n^*$ has a generator precisely when $n = 2,4,p^k,2p^k$ for odd primes p and positive integers k.

We can use the above to tighten Euler's Theorem. Write $n = 2^k p_1^{k_1}...p_m^{k_m}$ for odd distinct primes p_i, and define

$$\lambda(n) = \text{lcm}(\phi(2^k),\phi(p_1^{k_1}),...,\phi(p_m^{k_m}))$$

for $k \leq 2$ and

$$\lambda(n) = \text{lcm}(\phi(2^{k-1}),\phi(p_1^{k_1}),...,\phi(p_m^{k_m}))$$

for $k > 2$.

Then $a^{\lambda(n)} = 1$ for all $a \in \mathbb{Z}_n^*$, and furthermore $\lambda(n)$ is the smallest positive integer satisfying this condition because there exists $a \in \mathbb{Z}_n^*$ with order $\lambda(n)$.

19.5 Quadratic Residues

We can apply our new knowledge to study quadratic residues in more general settings.

If $\mathbb{Z}_n^*$ has a generator, then $\phi(n)$ plays the same role as $p-1$ in the odd prime case for quadratic residues.

For example, let us consider when -1 is a quadratic residue.

For odd primes p, $\phi(p^k) = p^k - p^{k-1}$, which is 0 or 2 mod 4 depending on whether p is 1 or 3 mod 4, so -1 is a quadratic residue in $\mathbb{Z}_p^k$ if and only if $p = 1 \pmod 4$.

For $p = 2$, if n is a quadratic residue modulo 2^k then it must also be a quadratic residue for all lower powers of 2, which implies, for example, -1 is a quadratic residue only when $k = 1$.

Write $n = \prod 2^k p_i^{k_i}$ for odd distinct primes p_i. By the Chinese remainder theorem, -1 is a quadratic residue if and only if $k \leq 1$ and each $p_i = 1 \bmod 4$.

20 Cyclotomic Equations

We try to solve the cyclotomic equation $x^p - 1 = (x-1)(x^{p-1} + x^{p-2} + ... + 1) = 0$ algebraically. (Transcendentally, the roots are $e^{2\pi ik/p}$ for $k = 0,...,p-1$.)

It can be easily shown that if $\gcd(m,n)=1$, then a primitive *mth* root of unity times a primitive *nth* root of unity is a primitive *mnth* root of unity, thus we need only consider prime powers. But then if α is a primitive *pth* root of unity, then $\sqrt[k]{\alpha}$ is a primitive p^k*th* root of unity, so we need only consider the case where p is prime.

In general we can use Gauss' method, but let us see how far elementary methods lead us.

$p=3$: we merely solve the quadratic $x^2+x+1=0$ to obtain

$$x=\frac{-1\pm i\sqrt{3}}{2}$$

$p=5$: we could solve the quartic $x^4+x^3+x^2+x+1=0$ but since it is palindromic we make the variable substitution $y=x+1/x$, and solve

$$y^2+y-1=0$$

to find

$$y=\frac{-1\pm\sqrt{5}}{2}$$

and $x^2-yx+1=0$ implies

$$x=\frac{y\pm\sqrt{y^2-4}}{2}$$

giving the four solutions

$$x=\frac{\sqrt{5}-1\pm\sqrt{-2\sqrt{5}-10}}{4},\frac{-\sqrt{5}-1\pm\sqrt{2\sqrt{5}-10}}{4}$$

$p=7$: the palindrome yields a cubic which can be solved for x.

$p=11$: the palindrome yields a quintic. Now elementary methods fail us and we need resort to Gauss' method as Vandermonde did.

21 The Heptadecagon

In 1796, a teenage Gauss proved that a regular 17-gon can be constructed using a straight-edge and compass by showing that a primitive 17th root of unity can be found by solving a succession of quadratic equations over the rationals.

Factorizing $x^{17}-1=0$ yields:

$$(x-1)(1+x+...+x^{16})=0$$

Let $\zeta=e^{2\pi i/17}$ be a primitive 17th root of unity. Since $\zeta\neq 1$, we must have:

$$\zeta+...+\zeta^{16}=-1$$

Since 3 is a generator of $\mathbb{Z}_{17}^*$, the primitive 17th roots of unity can be written in the sequence

$$\zeta^{3^0},\zeta^{3^1},...,\zeta^{3^{15}}$$

Define x_1 to be the sum of every second member of the sequence, and x_2 to be the sum of the other members, that is,

$$x_1=\zeta^{3^0}+\zeta^{3^2}+...+\zeta^{3^{14}}$$

$$x_2=\zeta^{3^1}+\zeta^{3^3}+...+\zeta^{3^{15}}$$

Then $x_1+x_2=-1$. By construction, x_1 and x_2 are Gaussian periods which means it is easy to compute $x_1x_2=-4$ (or use brute force(!)), thus x_1,x_2 are roots of a quadratic equation with integer coefficients, namely $(-1\pm\sqrt{17})/2$. The solution x_1 is the positive one since only two terms in its sum point to the left on the complex plane.

Next define y_1,y_2 from the elements used to construct x_1 in a similar way:

$$y_1=\zeta^{3^0}+\zeta^{3^4}+\zeta^{3^8}+\zeta^{3^{12}}$$

$$y_2=\zeta^{3^2}+\zeta^{3^6}+\zeta^{3^{10}}+\zeta^{3^{14}}$$

To save room, let us calculate the powers of 3 $\pmod{17}$:

$$1,3,9,10,13,5,15,11,16,14,8,7,4,12,2,6$$

Thus

$$y_1 = \zeta + \zeta^{13} + \zeta^{16} + \zeta^4$$
$$y_2 = \zeta^9 + \zeta^{15} + \zeta^8 + \zeta^2$$

Then $y_1 + y_2 = x_1$. It turns out $y_1 y_2 = -1$, thus y_1, y_2 are roots of a quadratic equation with coefficients involving the integers and x_1.

Similarly we can define y_3, y_4 from x_2

$$y_3 = \zeta^3 + \zeta^5 + \zeta^{14} + \zeta^{12}$$
$$y_4 = \zeta^{10} + \zeta^{11} + \zeta^7 + \zeta^6$$

and solve a quadratic to obtain their values.

Now define z_1, z_2 from y_1 in this fashion:

$$z_1 = \zeta + \zeta^{16}$$
$$z_2 = \zeta^{13} + \zeta^4$$

We have $z_1 + z_2 = y_1$ and $z_1 z_2 = y_3$, so z_1, z_2 can be found from a quadratic whose coefficients we know. Lastly we either note that both the sum and product of ζ and ζ^{16} are known so they can be found from a quadratic, or use the fact that

$$\zeta + \zeta^{16} = 2\cos(2\pi/17)$$

and simply halve z_1.

We can generalize this procedure to find expressions for any root of unity.

21.1 A Magic Solution

Using the above, we can give an elementary method for finding $\cos(2\pi/17)$ that seems to work magically. If we don't mention generators the solution appears mysterious.

Let $c_m = \cos(2\pi m/17)$. By considering the sums of the roots of unity we have $2(c_1 + ... + c_8) = -1$.

Set

$$a = c_1 c_4, b = c_3 c_5, c = c_2 c_8, d = c_6 c_7.$$

By basic trigonometric identities we have

$$2a = c_3 + c_5, 2b = c_2 + c_8, 2c = c_6 + c_7, 2d = c_1 + c_4.$$

(These correspond to the y_i s above.) Thus $a + b + c + d = -1/4$. Also,

$$ac = (c_3 + c_5)(c_6 + c_7)/4 = (c_1 + ... + c_8)/4 = -1/16.$$

Similarly $bd = -1/16$. We also find $16ab = -1 + 4a + 4b$, along with similar equations for bc, cd, da. Define

$$a + c = 2e, b + d = 2f$$

(Naturally, e, f correspond to x_1, x_2 above.) Then

$$e + f = -1/8, 4ef = ab + bc + cd + ad = -1/4$$

so we can solve a quadratic equation to find e, f. Once we have them, we can solve a quadratic equation to find a, c, and another to find b, d. With these values we can solve for c_1.

[I found this version in a solution that also describes a practical straight-edge-and-compass construction.]

22 Eisenstein's Irreducibility Criterion

Theorem: Let

$$f(x) = a_0 + a_1 x + ... + a_n x^n$$

be a polynomial with integer coefficients. Suppose a prime p divides each of $a_0, a_1, ..., a_{n-1}$ (every coefficient except the leading coefficient), and that p^2 does not divide a_0. Then $f(x)$ has no factor with integer coefficients.

Proof: Suppose $f = gh$. Look at this factorization modulo p.

It turns out $\mathbb{F}_p[x]$ is a unique factorization domain. Modulo p, since $f = a_n x^n$, we find $g = bx^d$ and $h = cx^e$ for some $b, c, d+e = n$.

In other words, p must divide every non-leading coefficient of g and h. In particular, p divides constant terms of g and h, hence p^2 must divide their product, that is, the constant term of f.

We can prove the theorem without introducing UFDs. As above, modulo p we have $f = a_n x^n$, so $f(0) = 0$, thus $g(0)h(0) = 0$ modulo p. Then at least one of $g(0)$ and $h(0)$ is 0. Without loss of generality $g(0) = 0$, so $g = xg_1$ for some polynomial g_1.

Then $a_n x^{n-1} = g_1 h$, and repeating this argument $n-1$ times shows $g = bx^d$ and $h = cx^e$ for some $b, c, d+e = n$. We now argue as before.

We can also prove the theorem more directly. Suppose $f = gh$ for polynomials g, h with integer coefficients. Let

$$g(x) = b_d x^d + ... + b_0$$

and

$$h(x) = c_e x^e + ... + c_0$$

for some $d+e = n$. The conditions imply p divides exactly one of b_0 and c_0. Without loss of generality, say p divides b_0 but not c_0.

Since p divides

$$a_1 = b_1 c_0 + b_0 c_1$$

we deduce p divides b_1. We now know p divides b_0, b_1 but not c_0.

Since p divides

$$a_2 = b_2 c_0 + b_1 c_1 + b_0 c_2$$

we deduce p divides b_2. We now know p divides b_0, b_1, b_2 but not c_0.

Continuing in this manner on $a_3 ... a_d$, we conclude by induction that p divides each of $b_0 ... b_d$.

But this implies p divides $b_d c_e = a_n$, a contradiction.

22.1 Gauss' Lemma

We usually combine Eisenstein's criterion with the next theorem for a stronger statement. (The name "Gauss' Lemma" has been given to several results in different areas of mathematics, including the following.)

Theorem: Let $f \in \mathbb{Z}[x]$. Then f is irreducible over $\mathbb{Z}[x]$ if and only if f is irreducible over $\mathbb{Q}[x]$.

(In other words, Let $f(x)$ be a polynomial with integer coefficients. If $f(x)$ has no factors with integer coefficients, then $f(x)$ has no factors with rational coefficients.)

Proof: Let $f(x) = g(x)h(x)$ be a factorization of f into polynomials with rational coefficients. Then for some rational a the polynomial $ag(x)$ has integer coefficients with no common factor. Similary we can find a rational b so that $bh(x)$ has the same properties. (Take the lcm of the denominators of the coefficients in each case, and then divide by any common factors.)

Suppose a prime p divides ab. Since

$$abf(x) = (ag(x))(bh(x))$$

becomes $0 = (ag(x))(bh(x))$ modulo p, we see $ag(x)$ or $bh(x)$ is the zero polynomial modulo p. (If not, then let the term of highest degree in $ag(x)$ be mx^r, and the term of highest degree in $bh(x)$ be nx^s. Then the product contains the term $mnx^{r+s} \neq 0 \pmod{p}$, a contradiction.)

In other words, p divides each coefficient of $ag(x)$ or $bh(x)$, a contradiction. Hence $ab = 1$ and we have a factorization over the integers.

Example: Let p be a prime. Consider the polynomial

$$f(x) = 1 + x + ... + x^{p-1}.$$

We cannot yet apply the criterion, so make the variable subsitution $x = y + 1$. Then we have

$$g(y) = 1 + (y+1) + ... + (y+1)^{p-1}.$$

Note $f(x)$ is irreducible if and only if $g(y)$ is irreducible.

The coefficient of y^k in $g(y)$ is

$$\sum_{m=k}^{p-1} \binom{p-1}{k} = \binom{p}{k+1}.$$

The last equality can be shown via repeated applications of Pascal's identity:

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

Alternatively, use the fact

$$g(y) = \frac{(y+1)^p - 1}{(y+1) - 1}$$

Thus p divides each coefficient except the leading coefficient, and p^2 does not divide the constant term p, hence $f(x)$ is irreducible over the rationals.

23 Gaussian Periods

Let $\zeta = e^{2\pi i/17}$ be a primitive 17th root of unity. Define x_1, x_2 by

$$x_1 = \zeta^{3^0} + \zeta^{3^2} + ... + \zeta^{3^{14}}$$

$$x_2 = \zeta^{3^1} + \zeta^{3^3} + ... + \zeta^{3^{15}}$$

The exponents of the terms of x_1 are the quadratic residues and those of x_2 are the nonresidues.

These are examples of *Gaussian periods*. We stated before that $x_1 x_2$ is a rational that is easy to compute. Why?

Write $(x_1 - x_2)^2$ as

$$(x_1 - x_2)^2 = a_0 + a_1\zeta^{3^1} + a_2\zeta^{3^2} + ... + a_{16}\zeta^{3^{16}}.$$

for integers a_i. Replacing ζ with ζ^3 merely swaps x_1 and x_2 due to their construction, whence

$$(x_2 - x_1)^2 = a_0 + a_1\zeta^{3^2} + a_2\zeta^{3^3} + ... + a_{16}\zeta^{3^{17}}.$$

Thus looking at the coefficients of the powers of ζ we have $a_{16} = a_1, a_1 = a_2, ..., a_{15} = a_{16}$, that is, they are all equal to some integer a:

$$(x_2 - x_1)^2 = a_0 + a(\zeta + ... + \zeta^{16}) = a_0 - a$$

Thus $x_1 x_2 = ((x_1 + x_2)^2 - (x_1 - x_2)^2)/4$ is some rational number. (Recall $x_1 + x_2 = -1$.)

Each of x_1 and x_2 consists of 8 terms so their product $x_1 x_2$ has 64 terms of the form ζ^k for some k. Since $17 = 1 \pmod 4$ we know -1 is a quadratic residue, hence if r is a quadratic residue so is $-r$. As each k can be viewed as the sum of a residue and a nonresidue modulo 17, this means k is never zero.

Hence if we write:

$$x_1 x_2 = b_0 + b(\zeta + ... + \zeta^{16})$$

we must have $b_0 = 0$, and the 64 terms must be evenly distributed among $\zeta, ..., \zeta^{16}$, that is, $b = 4$. Therefore $x_1 x_2 = -4$. (Thanks to Dennis Westra for bringing this argument to my attention.)

Exercise: what happens when -1 is a quadratic nonresidue?

We can generalize these statements.

23.1 A Loose End

Before we argued that given

$$a_0 + a_1\zeta^{3^1} + ... + a_{16}\zeta^{3^{16}} = a_0 + a_1\zeta^{3^2} + ... + a_{16}\zeta^{3^{17}}$$

we can equate coefficients of the powers of ζ. Let us see why. In the above equation, we can move all terms to one side then divide through by ζ to find b_i such that

$$b_1 + b_2\zeta + ... + b_{16}\zeta^{15} = 0$$

(each b_i is the difference between some a_j and a_k). Then consider the polynomial $g(x) = b_1 + b_2x + ... + b_{16}x^{15}$. Now ζ is a root of g as well as the polynomial

$$f(x) = 1 + x + ... + x^{16}.$$

Hence ζ must also be a root of $d = \gcd(f,g)$. If $g \neq 0$, then the polynomial d is a nonzero polynomial dividing f with degree at most g, which is smaller than the degree of f. This is a contradiction since f has no factors with rational coefficients (by Eisenstein).

Thus g must be the zero polynomial, and we may equate the coefficients of the powers of ζ.

[In abstract algebra, we say all this in one line: $\mathbb{Q}[\zeta]$ is a degree 16 extension of $\mathbb{Q}$ thus $g = 0$.]

24 Roots of Unity

Gauss generalized his method to find an expression using radicals for any root of unity. (Compare with Vandermonde's method.)

Suppose we want to find an expression for a primitive *pth* root of unity ζ for a prime p, and assume we have done so for smaller primes. Let d, D be factors of $p-1$ such that $D = qd$ for some q. Let g be a generator of $\mathbb{Z}_p^*$. Let β be a primitive *qth* root of unity.

For any expression γ containing ζ, define $S\gamma$ to be the same expression with each ζ replaced by ζ^g.

Suppose γ satisfies $S^D\gamma = \gamma$. Then define

$$t = \gamma + \beta S^d\gamma + \beta^2 S^{2d}\gamma + ... + \beta^{q-1}S^{(q-1)d}\gamma$$

Then replacing ζ by ζ^{g^d} in this expression yields $\beta^{-1}t$. Since $t^q = (\beta^{-1}t)^q$ we can equate the coefficients of the powers of ζ as before to argue t^q can be expressed in terms of β.

Example: Take $d = 1, D = 2, p = 17$. Then $q = 2, \beta = -1$. If we take $\gamma = x_1$ as defined earlier discussing the 17-gon, we see $Sx_1 = x_2, Sx_2 = x_1$, thus $S^2\gamma = \gamma$. Then the expression t is simply

$$t = \gamma + \beta S\gamma = x_1 - x_2$$

and we saw before t^2 must be an integer.

To continue, we took $d = 2, D = 4$. Again $q = 2, \beta = -1$ and we can take $\gamma = y_1$ as defined earlier. Then note $S^4\gamma = \gamma$ and we see $(y_1 - y_2)^2$ is an integer.

Now define t_i to be t where each β has been replaced by β^i. Then we have

$$\gamma = \frac{t_1 + ... + t_q}{q}$$

(much cancellation occurs since the sum of the *kth* roots of unity is zero for any $k > 1$). By a similar argument, each t_i^q is known, and thus if we choose *qth* roots correctly, then

$$\gamma = \frac{1}{q}\sum_{i=1}^{q} \sqrt[q]{t_i^q}$$

(the $\sqrt{}$ symbol does not have its usual meaning here because the particular *qth* roots we need may not be real).

Instead of trying every possible root until the resulting γ is correct, we consider the expression $t_i t_1^{q-i}$. If we change each ζ to ζ^{g^d} (that is apply S^d) then t_i changes to $\beta^{-i}t_i$, while from before we know t_1^{q-i} becomes $\beta^{-(q-i)}t_1^{q-i}$, thus their product is unchanged.

Arguing as before, $t_i t_1^{q-i}$ is known for all i, so once we have made a choice for the value of t_1 we can easily find the values for each t_i without guesswork.

Example: Let ζ be a primitive fifth root of unity. We shall derive an expression for ζ in terms of a primitive fourth root of unity.

Set $d=1, D=4, p=5$. Take $g=2$, since 2 generates $\mathbb{Z}_5^*$. Then $q=4, \beta=i$. Set γ to simply ζ, so the t_is are:

$$\begin{aligned} t_1 &= \zeta + i\zeta^2 - \zeta^4 - i\zeta^3 \\ t_2 &= \zeta - \zeta^2 + \zeta^4 - \zeta^3 \\ t_3 &= \zeta - i\zeta^2 - \zeta^4 + i\zeta^3 \\ t_4 &= \zeta + \zeta^2 + \zeta^4 + \zeta^3 = -1 \end{aligned}$$

We compute t_1^4 and choose a fourth root of the result, from which we work out t_2, t_3, t_4. To make the computation easier we notice

$$\begin{aligned} t_2^2 &= (\zeta + \zeta^2 + \zeta^3 + \zeta^4) + 2(-\zeta^3 + 1 - \zeta^4 - \zeta^1 + 1 - \zeta^2) \\ &= (-1) + 2(2 - (-1)) = 5 \end{aligned}$$

(I've omitted shortcuts I took for clarity, e.g. since $2 \in \mathbb{Z}_5^*$ the squares of different powers of ζ will be different powers of ζ, and they will add up to -1.)

Now

$$\begin{aligned} t_1^2 &= (\zeta^2 - \zeta^4 + \zeta^3 - \zeta) + 2(i\zeta^3 - 1 - i\zeta^4 - i\zeta + 1 + i\zeta^2) \\ &= (-t_2) + 2i(-t_2) \end{aligned}$$

Thus $t_1^4 = 5(1+2i)^2$ whence $t_1 = \alpha(\sqrt[4]{5}\sqrt{1+2i})$ where α is a fourth root of unity.

Now that we have found the solutions of t_1, we compute

$$\begin{aligned} t_1^2 t_2 &= -(t_2^2)(1+2i) = -5(1+2i) \\ t_1 t_3 &= (\zeta - \zeta^4)^2 - (i(\zeta^2 - \zeta^3))^2 \\ &= (\zeta^2 + \zeta^3 + \zeta^4 + \zeta) + 2(-1-1) = -5 \\ t_4 &= -1 \end{aligned}$$

(Actually first equation is unnecessary since we already have t_2 in terms of t_1 from before.)

Thus after some algebraic manipulation we find

$$\begin{aligned} t_1 &= \alpha(\sqrt[4]{5}\sqrt{1+2i}) \\ t_2 &= -\alpha^2\sqrt{5} \\ t_3 &= -\alpha^3(\sqrt[4]{5}\sqrt{1-2i}) \\ t_4 &= -1 \end{aligned}$$

Finally, we have all four of the primitive fifth roots of unity:

$$\zeta = \frac{-\alpha^2\sqrt{5} - 1 + \alpha(\sqrt[4]{5})(\sqrt{1+2i} - \alpha^2\sqrt{1-2i})}{4}$$

where $\alpha = \pm 1, \pm i$.

If instead we had chosen $d=1, D=2$, and then $d=2, D=4$ (i.e. mirror the process used for the 17th roots of unity] we have ζ expressed in terms of a primitive square root of unity (i.e. over the rationals, since -1 is rational):

$$\zeta = \frac{\sqrt{5} - 1 \pm \sqrt{-2\sqrt{5} - 10}}{4}, \frac{-\sqrt{5} - 1 \pm \sqrt{2\sqrt{5} - 10}}{4}$$

which can be verified to be the same solutions.

25 Binary Quadratic Forms

Can you write 67 in the form x^2+7y^2? By brute force, we find $67 = 2^2 + 7 \times 3^2$, so the answer is yes. But what if I ask the same question about a larger prime like 1234577?

We'll soon learn $1234577 = x^2 + 7y^2$ only if 7 is quadratic residue modulo 1234577. Using quadratic reciprocity, we find that $(7|1234577) = -(1|7) = -1$, so this time there are no solutions.

A *binary quadratic form* is written $[a,b,c]$ and refers to the expression $ax^2+bxy+cy^2$. We are interested in what numbers can be *represented* in a given quadratic form.

The *divisor* of a quadratic form $[a,b,c]$ is $\gcd(a,b,c)$.

Representations x,y with $\gcd(x,y) = 1$ are *primitive representations*.

If the divisor of a form is 1 then it is a *primitive form*, but we can forget this. For our purposes, primitive representations matter, and primitive forms don't.

Completing the square, we find:

$$4a(ax^2+bxy+cy^2) = (2ax+by)^2 - dy^2$$

where $d = b^2 - 4ac$. We call d the *discriminant*.

If $d = 0$, then the quadratic form is a perfect square. This case is trivial.

If $d < 0$ then $ac > 0$, so a,c have the same sign. From the above equality we see if $a > 0$ then the form is nonnegative for any x,y. We call such a form *positive definite*. Similarly, if $a < 0$ then the form is *negative definite*.

If $d > 0$ then a little experimentation shows the form takes negative and positive values. Such a form is termed *indefinite*.

25.1 Equivalent forms

Given a form, if we swap x and y then the resulting form represents the same numbers. We consider them equivalent. There are less trivial ways to change a form so it represents the same numbers. If we replace x with $x+y$, then $x = u-v, y = v$ represents the same number that $x = u, y = v$ did in the original form.

More generally, let T be a 2x2 matrix with integer entries of determinant ± 1, that is, an *integral unimodular matrix*. We state facts that are easy but tedious to prove.

The quadratic form $[a,b,c]$ can be written as the matrix:

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

Why? Evaluate $\begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}$.

We have $\det A = -4d$; some authors define d to be $\det A$ instead of the discriminant, which generalizes nicely beyond the quadratic case.

Then let:

$$A' = T^T A T = \begin{pmatrix} a' & b'/2 \\ b'/2 & c' \end{pmatrix}$$

for some a',b',c'. We write $[a,b,c] \sim [a',b',c']$; this relation is an equivalence relation.

If $\begin{pmatrix} u \\ v \end{pmatrix} = T \begin{pmatrix} u' \\ v' \end{pmatrix}$ then u,v represents the same integer under A as u',v' does under A', and we call these equivalent representations. Equivalent representations have the same divisor.

Equivalent forms represent the same integers, have the same divisor and discriminant.

[If two forms represent the same integers, are they necessarily equivalent? I don't know.]

25.2 Principal forms

The *principal form* of a discriminant d is $[1,0,-k]$ when $d = 4k$ and $[1,1,k]$ when $d = 4k+1$. Its equivalence class is the 'principal class of forms of discriminant d'.

25.3 Reduced forms

By applying transformations judiciously, we can *reduce* any definite form to $[a,b,c]$ such that $-|a| < b \leq |a| < |c|$ or $0 \leq b \leq |a| = |c|$.

Reduction is like Euclid's algorithm. There exist q,r such that $-b = 2|c|q + r$ with $-|c| < r \leq |c|$, so we apply the integral uniform matrix:

$$\begin{pmatrix} 0 & 1 \\ -1 & \operatorname{sgn}(c)q \end{pmatrix}$$

to transform an unreduced form $[a,b,c]$ to $[c,r,d]$ for some d. Repeating eventually leads to $|c| \leq |d|$.

Thus we have a form $[a,b,c]$ satisfying $-|a| < b \leq |a| \leq |c|$, and we are done unless $|a| = |c|$ and $b < 0$. In this last case, we apply the above procedure one more time (we'll find $q = 0$ and $r = -b$) to get the reduced form $[c,-b,a]$.

```
-- | Matrix multiplication.
mmul a b = [[sum $ zipWith (*) r c | c <- transpose b] | r <- a]

tat t m = mmul (transpose t) $ mmul m t

reduce (a, b, c)
  | b^2 - 4*a*c >= 0 = error "indefinite form"
  | -abs a < b, b <= abs a, abs a <  abs c = (a, b, c)
  |      0 <= b, b <= abs a, abs a == abs c = (a, b, c)
  | otherwise = reduce (div a' 2, b', div c' 2)
  where
  c2       = 2 * abs c
  (q0, r0) = (-b) `divMod` c2
  (q , r) | r0 * 2 > c2 = (q0 + 1, r0 - c2)
          | otherwise   = (q0    , r0)
  [[a', b'],[_, c']] = tat
    [[0, 1], [-1, signum c * q]]
    [[2*a, b], [b, 2*c]]
```

Our reduction algorithm works for indefinite forms with nonzero a and c. However, it turns out we need to refine our definition of "reduced" to get useful results in the indefinite case. We'll skip this part of the theory.

Principal forms are reduced forms.

The above conditions imply $b^2 \leq |ac| \leq |d|/3$ when $ac/=0$. This suggests a brute force algorithm to find all reduced forms of a given discriminant d.

The following function returns all positive definite forms for a given discriminant. The negative definite forms are the same with a and c negated.

```
pos d
  | d >= 0        = error "d must be negative"
  | d `mod` 4 > 1 = error "d must be 0 or 1 mod 4"
  | otherwise     = posBs ++ negBs
  where
  upFrom n = takeWhile (\x -> x^2 <= abs d `div` 3) [n..]
  posBs = [(a, b, c) | b <- upFrom 0, a <- upFrom b, a /= 0,
    let (c, r) = divMod (b^2 - d) (4*a), r == 0, c >= a]
  negBs = [(a, -b, c) | (a, b, c) <- posBs, a /= c, b > 0, a > b]
```

For example:

```
> pos (-39)
[(1,1,10),(2,1,5),(3,3,4),(2,-1,5)]
```

Since there are only finitely many reduced forms of discriminant d and every form is equivalent to some reduced form, the number of equivalence classes of forms with discriminant d is finite. We call this number the *class number* of the discriminant d.

At least that's what I understood from *Number Theory* by John Hunter. I think the class number is actually the number of equivalence classes of positive definite forms when $d < 0$, as there's no point doubling the total by also counting the negative definite forms.

Theorem: The equivalence class of a positive definite binary quadratic contains exactly one reduced form.

Proof. Let $f = [a,b,c]$ be a reduced positive definite binary quadratic form. Again, we complete the square:

$$4af(x,y) = (2ax+by)^2 - dy^2$$

Both a and c are positive so $-a < b \leq a < c$ or $0 \leq b \leq a = c$.

Then if $|y| \geq 2$ then $-dy^2 \geq 12ac$ thus

$$f(x,y) \geq 3c > a+c.$$

And if $|x| \geq 2$ and $|y| = 1$ we find

$$f(x,y) \geq ax^2 - a|x| + c \geq 2a + c > a + c.$$

Among the remaining the cases, we find the four smallest integers primitively represented are $a, c, a+b+c, a-b+c$, corresponding to $(x,y) = (1,0),(0,1),(1,1),(1,-1)$. The smallest three are $a, c, a+c-|b|$ and since these are each smaller than $a+c$, they are the smallest three integers (possibly non-distinct) primitvely represented by $[a,b,c]$. Observe $a \leq c \leq a+c-|b|$.

If $[a',b',c']$ is a reduced form equivalent to $[a,b,c]$, the same reasoning implies the smallest three integers it primitively represents are $a', c', a'+c'-|b'|$ and $a' \leq c' \leq a'+c'-|b'|$.

Thus $a = a'$, $c = c'$ and $|b| = |b'|$. When $a = c$, both b and b' are nonnegative, implying $b = b'$. It remains to prove this holds when $a < c$, which is the least pleasant part of the proof.

Suppose $b = -b'$. Since $[a,b,c] \sim [a,-b,c]$ there exists a 2x2 integer matrix

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}$$

with $ps - qr = 1$ that transforms one to the other, that is,

$$\begin{aligned} a &= ap^2 + bpr + cr^2 \\ -b &= 2apq + b(ps+qr) + 2crs \end{aligned}$$

Then:

$$a = ap^2 + bpr + cr^2 > ap^2 - a|pr| + a^2 \geq 2a|pr| - a|pr| = a|pr|$$

Thus we must have $pr = 0$. If $p = 0$ then $r/=0$, whence $a = ap^2 + bpr + cr^2 > c$, a contradiction. So $r = 0$, which means $ps = 1$.

Then:

$$-b = 2apq + b(ps+qr) + 2crs = 2apq + b$$

We deduce $|b| = a|pq|$, which implies $b = 0$ or $b = a$. Since $[a,-a,c]$ is not reduced, we must have $b = 0$.

Theorem: An integer n is primitively representable by a quadratic form $[a,b,c]$ if and only if $[a,b,c] \sim [n,b',c']$ for some b', c'.

Proof. If $n = ax^2 + bxy + c^2$ and $\gcd(x,y) = 1$ then the extended Euclid's algorithm can find p,q so that $px - qy = 1$. Then the integral unimodular transformation:

$$\begin{pmatrix} x & q \\ y & p \end{pmatrix}$$

to $[a,b,c]$ gives $[n,b',c']$ for some b', c'.

Conversely, $(1,0)$ primitively represents n for $[n,b',c']$, so transforms to a primitive representation of n for $[a,b,c]$.

Theorem: A nonzero integer n is primitively representable by a form of discriminant d if and only if

$$x^2 = d \bmod 4|n|$$

for some x.

Proof. If n is primitive representable, by the previous theorem there exists a form $[n, b', c']$ that primitively represents n. The condition follows immediately from $d = b'^2 - 4nc'$.

Conversely, the condition implies $b'^2 - 4nc' = d$ for some integers b', c', and 1, 0 is a primitive representation of n by the form $[n, b', c']$.

Example: The form $x^2 + 3y^2$ of discriminant -12 cannot represent 2, yet $2^2 = -12 \pmod 8$. The above theorem simplies there must exist some form of discriminant -12 in another equivalence class, and indeed we find $[2, 2, 2]$ has the desired discriminant and can represent 2.

25.4 Sum of two squares

The quadratic form $[1, 0, 1]$ has discriminant -4. From the previous theorem, a nonnegative integer n is primitively represented by some form of discriminant -4 if and only there exists x satisfying $x^2 = -4 \bmod 4n$, which is the same as $x^2 = -1 \bmod n$.

A brief search confirms $[1, 0, 1]$ is only reduced positive definite form of discriminant -4. Therefore, n is primitively represented by some form of discriminant -4 if and only if n is primitively represented by $[1, 0, 1]$, namely, n is the sum of two squares.

Factorize n:

$$n = \prod 2^s p_i^{k_i}$$

where the p_i are distinct odd primes. By the Chinese Remainder Theorem, and considering quadratic residues for prime powers, n is primitively represented by $x^2 + y^2$ if and only if each $p_i = 1 \bmod 4$ and $s \leq 1$.

If we allow non-primitive representations, then n is a sum of two squares provided $k_i = 0 \bmod 2$ whenever $p_i = 3 \bmod 4$.