Introduction to Number Theory
Lecture Notes

Adam Boocher

2014

1 Introduction

These notes will cover all material presented during class. These lectures have been compiled from a variety of sources, mainly from the recommended books:

- Elementary Number Theory, by Kenneth H. Rosen, 6th Edition, 2011, Pearson. Library: QA241Ros
- A friendly introduction to number theory by J. H. Silverman, Prentice Hall, 2013.Library: QA241Sil

These books are both excellent sources of examples, additional practice problems and I find them to be eminently readable. They are on reserve in the Murray Library.

1.1 A Preview: Pythagorean Triples

The classical theorem of Pythagoras states that if a, b, c are the side lengths of a right triangle, (c being the hypotenuse) then

$$a^2 + b^2 = c^2.$$

In this lecture we will answer the following

Question 1.1. *What are the natural number solutions (a, b, c) to the equation $a^2 + b^2 = c^2$? Such a solution is called a Pythagorean triple.*

Example 1.2. *Some easy-to-remember Pythagorean triples are e.g.* $(3, 4, 5), (5, 12, 13), (8, 15, 17)$.

A first question we might ask if there are infinitely many such triples. However, we see that as soon as we have a single solution, we have found infinitely many:

Remark 1.3. *If (a, b, c) is a Pythagorean triple, and d is any positive integer then so is (da, db, dc).*

Proof. We just check that

$$(da)^2 + (db)^2 = d^2(a^2 + b^2) = d^2(c^2) = (dc)^2.$$

□

Given this fact, we define a **primitive Pythagorean triple (PPT)** to be a Pythagorean triple such that a, b, c have no common factor. This means that there is no number d that divides all of a, b, c. We can now rephrase Question 1.1 as: What is the set of PPTs?

As a first step, let's consider the possible parities of the numbers (the parity of a number refers to whether the number is even or odd). It's straightforward to check that the square of an even number is even, and the square of an odd number is odd. With that in mind, the only possible solutions to $a^2 + b^2 = c^2$ must be of the form

$$\begin{gathered}\text{odd} + \text{odd} = \text{even} \\ \text{odd} + \text{even} = \text{odd} \\ \text{even} + \text{even} = \text{even}\end{gathered}$$

We can rule out the last possibility since that would imply that a, b, c are divisible by 2. We can also rule out the first possibility: Suppose that

$$a = 2x + 1, \quad b = 2y + 1, \quad c = 2z.$$

Then after simplifying we see that

$$4x^2 + 4x + 4y^2 + 4y + 2 = 4z^2.$$

But this is impossible, since the right hand side is divisible by 4 but the left hand side is not!

Hence we can reformulate Question 1.1 as

Question 1.4. *Find all natural number solutions to $a^2 + b^2 = c^2$ with a odd, b even, and a, b, c have no common factors.*

Remark 1.5. *Notice that requiring that a, b, c have no common factor is the same as requiring that no two of them share a common factor. Indeed, if p was a common prime factor, then if p divides, say c and b, then it divides $c^2 - b^2$ and hence it divides a^2. But now by prime factorisation, this means it divides a. (Thanks to those who asked me about this after the lecture!)*

Let's get to work! We can note that

$$a^2 = c^2 - b^2 = (c - b)(c + b)$$

In other words, the product of $(c - b)$ and $(c + b)$ is a perfect square. We will now show that $c - b$ and $c + b$ are relatively prime. Indeed, suppose that they both shared a common prime factor d, then certainly d should divide their sum and difference. Thus d divides

$$(c - b) + (c + b) = 2c, \quad \text{and} \quad (c + b) - (c - b) = 2b.$$

But now b and c have no common factor, so it must be that d divides 2. But d cannot be 2 since $c+b$ is odd! But now as $c-b, c+b$ have no factors in common (i.e. they are *relatively prime*) we see that the only way that their product can be a square is if both factors are squares:[1]

$$c+b=s^2, \quad c-b=t^2,$$

where $s>t\geqslant 1$ are odd integers with no common factors. Then

$$a=\sqrt{(c-b)(c+b)}=\sqrt{s^2t^2}=st.$$

We can now solve for b and c to obtain our first

Theorem 1.6. *Every PPT (a,b,c) with a odd satisfies*

$$a=st, \quad b=\frac{s^2-t^2}{2}, c=\frac{s^2+t^2}{2},$$

where $s>t\geqslant 1$ are chosen to be odd integers with no common factors.

You should notice that we have only completed "half" of this proof. To complete it, we should show that for every such choice of s,t we actually obtain a PPT. (Can you complete the proof?)

This theorem is quite striking at first glance, but it still leaves a bit to be desired as to "why" PPTs should have such a special form. Also, we seemed to have gotten lucky with our even/odd analysis in the proof. Indeed, for many problems in number theory, things won't work out this nicely! However, there is a nice method which extends a bit more generally, which we present here.

1.2 A Geometric Derivation

Notice that if (a,b,c) is a PPT then $(a/c, b/c)$ is a point with rational coordinates on the unit circle $x^2+y^2=1$. As students in the North, we naturally notice that $N=(0,1)$ is a point on the circle. Never matter that one of the coordinates is zero, we can worry about that later. The key insight is to now notice that if (x,y) is another point on the circle with rational coordinates then the slope of the line between these two points will have rational slope. The converse is also true, which we prove now:

Suppose that $P=(x,y)$ is a point on the unit circle such that the line between N and P has a rational slope m. Then $m=(y-1)/x$ or equivalently $y=mx+1$. Since P lies on the unit circle, we can conclude that

$$x^2+(mx+1)^2=1$$

$$(1+m^2)x^2+2mx+1=1$$

$$x((1+m^2)x+2m)=0$$

[1] This actually relies on unique factorization of integers, to be covered in the next lecture.

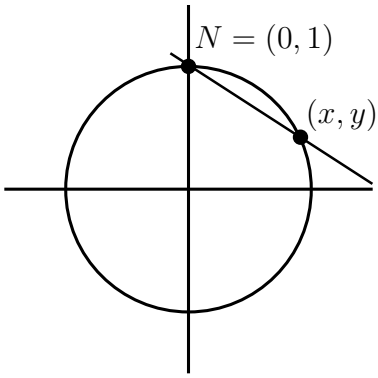

Figure 1: Geometric Method of Finding Pythagorean Triples

This equation describes the set of points that lie on the intersection of the circle and the line. The solution $x = 0$ is the point N, and the solution $x = (-2m)/(1 + m^2)$ describes the point P. Using $y = mx + 1$ we have proven

Theorem 1.7. *Every point on the circle $x^2 + y^2 = 1$ with rational coordinates is of the form*

$$(x, y) = \left(\frac{-2m}{1 + m^2}, \frac{1 - m^2}{1 + m^2} \right)$$

where m is a rational number. (Except for the point $(0, -1)$ which is the limiting value as $m \to \infty$.)

If we write $m = u/v$ and clear denominators, then we see that this formula becomes:

$$(x, y) = \left(\frac{-2uv}{u^2 + v^2}, \frac{v^2 - u^2}{v^2 + u^2} \right)$$

which if we plug into the equation for the unit circle and simplify we get

$$x^2 + y^2 = 1$$

$$(-2uv)^2 + (v^2 - u^2)^2 = (v^2 + u^2)^2$$

$$(uv)^2 + \left(\frac{v^2 - u^2}{2} \right)^2 = \left(\frac{v^2 + u^2}{2} \right)^2.$$

Comparing with 1.6 we see that we've found exactly the same points!

There are lots of other questions we might want to answer. For example, if c is given, do there exist a and b so that $a^2 + b^2 = c$? If so, how many such a and b are there? For example

$$33^2 + 56^2 = 65^2 \quad \text{and} \quad 16^2 + 63^2 = 65^2.$$

It turns out that the a highbrow way to view this question (and others) involves passing to the so-called ring of Gaussian integers - which involves imaginary numbers. We will return to this topic at the end of the course once we have a larger toolkit.

Main Points from Lecture 1:

- Know how to find infinitely many PPTs
- Be able to use the geometric method of using lines with rational slope to find rational points. Memory of this method is important. Memory of the actual formulas is not.
- Have familiarity with basic divisibility arguments.

2 The Primes

Notation: $\mathbb{Z}$ = ring of integers $\{0, \pm 1, \pm 2, \dots\}$;
$\mathbb{N}$ = set of positive integers $\{1, 2, 3, \dots\}$;
$\mathbb{Q}$ = field of rational numbers $\{n/m : m \in \mathbb{N}, n \in \mathbb{Z}\}$;
$\mathbb{R}$ = field of real numbers.

2.1 Prime Numbers

A positive integer $p > 1$ is called *prime* if $p \neq mn$ for all $m, n \in \mathbb{N}$ with $m > 1$ and $n > 1$. Otherwise (i.e., if $c > 1$ can be written as $c = mn$ for some $m, n \in \mathbb{N}$ with $m > 1$ and $n > 1$) then c is called *composite*.

Theorem 2.1 (Fundamental Theorem of Arithmetic)**.** *Every $n \in \mathbb{N}$ has a unique representation as a product of primes:*

$$n = p_1^{e_1} \cdots p_k^{e_k}, \qquad \text{where } k \geq 0 \text{ and each } e_k \in \mathbb{N}.$$

[Convention: empty products (here, for $n = 1$, are $= 1$, and empty sums are $= 0$.]

This theorem was proved in Year 1 (in Proofs and Problem-solving). See Martin Liebeck: A concise introduction to Pure Mathematics, Chapman and Hall 2000. A visualisation of this theorem can be seen at http://www.datapointed.net/visualizations/math/factorization/animated-diagrams/

Remark 2.2. *We've known some version or other of the FTA for most of our lives, and as such it probably seems like a rather obvious, and daresay, even boring fact. However, it's really quite a striking feature of the natural numbers. Later in the course we will encounter other number systems (i.e. rings) in which unique factorization into primes does not hold. For an excursion in this direction, take a look at Silverman's discussion on the $\mathbb{E}-$world, in which he talks about the set of even numbers.*

Proposition 2.3. *Suppose $p, n, m \in \mathbb{N}$ with p prime and p dividing nm (i.e., $pr = nm$ for some $r \in \mathbb{N}$). Then either $p \mid n$ ("p divides n") or $p \mid m$ (or possibly both).*

Proof. Apply Theorem 2.1 to $pr = nm$ by writing r, n and m as a product of primes. Then pr and nm give two ways of writing the same number as a product of primes. Because this representation is unique by Theorem 2.1, p must appear in the representation of nm as a product of primes. But this representation was obtained by multiplying the representation of n and the representation of m as products of primes. So p must appear in at least one of these representations. Hence $p \mid n$ or $p \mid m$. □

This result is false for composite numbers, as e.g., $6 = 2 \cdot 3$, and so $6 \mid 2 \cdot 3$, but $6 \nmid 2$ and $6 \nmid 3$. More generally, if $c = nm$ with $n > 1$, $m > 1$ (so c composite) then $c \mid nm$ but $c \nmid n$ and $c \nmid m$. In general, if you see a non-example like this, remember it! I find this to be the best way to remember the hypotheses of theorems.

Example 2.4. *If p is a prime number and p divides $2n$ then either p divides 2 or p divides n. We used this fact in the first lecture when we were discussing Pythagorean triples.*

Theorem 2.5. *If n is composite then it must be divisible by some prime $\leqslant \sqrt{n}$.*

Proof. If all prime factors of n are $> \sqrt{n}$ then clearly all factors of n are $> \sqrt{n}$. Thus since n is composite, we have some factorisation $n = ab > \sqrt{n}\sqrt{n} = n$, a contradiction. □

This gives us a reasonable algorithm to enumerate, the first few primes. Suppose we wanted to enumerate all primes less than 100. We could write the first 100 numbers down, and then cross off all multiples of 2, 3, 5, and 7. By the previous theorem, any numbers remaining must be prime, since $\sqrt{100} = 10$. This is called the Sieve of Eratosthenes. For an animation and more information click: http://en.wikipedia.org/wiki/Sieve_of_Eratosthenes

2.2 Distribution of the primes

A whole course could be devoted to the distribution of the prime numbers. Basically the main motivating question asks: How are the primes interspersed among the natural numbers? As a first step, we know from Euclid that there are infinitely many primes:

Theorem 2.6. *(Euclid) There are infinitely many prime numbers.*

Proof. Suppose that there were only finitely many primes $p_1, \ldots, p_k$. Then consider the integer $N = p_1 \cdots p_k + 1$. This number is not divisible by any of of the p_i (it has remainder 1 upon division). However, by Theorem 2.1 it must be divisible by some prime p. Since this p is none of our p_i we have a contradiction. We must not have written down all the primes. □

We also have a proof (quite easily) that there are infinitely many odd primes. In other words, there are infinitely many prime numbers of the form $2k + 1$. Recall from the first lecture that the dichotomy between even and odd numbers can be generalised. On the first homework assignment you will prove that there are infinitely many prime numbers of the form $4k + 3$. In fact, a much stronger statement is true:

Theorem 2.7 (Dirichlet's Theorem)**.** *If a and b are positive integers not divisible by the same prime then there are infinitely many primes of the form $ak + b$.*

The proof of this theorem is difficult, and is beyond the scope of this course. It is just the tip of the iceberg concerning questions about the distribution of the primes. My favourite theorem of this type, for which we *do* have an elementary proof is the following. It provides yet another proof that there are infinitely many primes.

Theorem 2.8. *The sum of the reciprocals of the primes diverges*

$$\sum_{p \text{ prime}} \frac{1}{p} \to \infty.$$

We will need the following tools for the proof:

0. The Fundamental Theorem of Arithmetic.
1. $1/(1-x) = 1 + x + x^2 + \cdots + x^n + \cdots$.
2. $\log(1-x) = -x - x^2/2 - x^3/3 - \cdots - x^n/n - \cdots$.
3. The harmonic series $1 + 1/2 + 1/3 + 1/4 + \cdots$ diverges.

(Tools 1-3 were introduced in Calculus, (recall that the series $\sum 1/n^s$ converges if and only if $s > 1$.)

Proof. Let n be a natural number. We define the following product:

$$\lambda(n) = \prod_{p \leqslant n} \left(\frac{1}{1 - \frac{1}{p}} \right)$$

Our first goal is to prove that $\lambda(n)$ diverges. To see this, note that we can rewrite each of the factors as an infinite sum using Tool 1.

$$\lambda(n) = \left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\right)\left(1 + \frac{1}{3} + \frac{1}{3^2} + \cdots\right)\left(1 + \frac{1}{5} + \frac{1}{5^2} + \cdots\right)\cdots$$

When we expand this product, we will obtain all fractions of the form

$$\frac{1}{2^{a_1}3^{a_2}\cdots p_k^{a_k}}$$

where all prime factors $\leqslant n$ appear. By the fundamental theorem of arithmetic, we'll get all the numbers $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}$ (and many many more). Therefore

$$\lambda(n) > 1 + 1/2 + 1/3 + \ldots + 1/n.$$

Hence as $n \to \infty$, $\lambda(n) \to \infty$ by Tool 3. In particular this means that $\log \lambda(n) \to \infty$ as $n \to \infty$.

Our second (and final goal) is to relate $\lambda(n)$ with the series $\sum 1/p$. To do this, we take logs. By basic properties of logs of products and reciprocals, we obtain:[2]

$$\log(\lambda(n)) = \sum_{p \leqslant n} -\log(1 - 1/p).$$

Using Tool 2, we obtain:

$$\begin{aligned} -\log(1-1/p) &= \frac{1}{p} + \frac{1}{2p^2} + \frac{1}{3p^3} + \frac{1}{4p^4} + \cdots \\ &< \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \frac{1}{p^4} + \cdots \\ &= \frac{1}{p}(1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \frac{1}{p^4} + \cdots) \\ &= \frac{1}{p} \cdot \frac{1}{1-1/p} = \frac{1}{p} \cdot \frac{p}{p-1} \\ &\leqslant \frac{2}{p}. \end{aligned}$$

But then this means that

$$\log \lambda(n) < \sum_{p \leqslant n} \frac{2}{p}.$$

Since the left hand side diverges, we have that $\sum_{p \leqslant n} \frac{2}{p}$ diverges. Division by two yields the result. □

This Theorem is somewhat surprising, since for instance the sum $1+1/4+1/9+1/16+\cdots$ converges, yet $\sum 1/p$ diverges. Hence it might be appropriate to say that "There are more prime numbers than square numbers." However this sentence is pure nonsense without a proper definition. Along a different track, one way of measuring how many prime numbers there are is to consider the fraction of natural numbers that the primes comprise.

Consider the function

$$\pi(n) = \#\{primes\ p \leqslant n\}.$$

E.g. $\pi(5) = 3$, $\pi(100) = 25$, and $\pi(5000) = 669$. In the limit it looks like

$$\lim \frac{\pi(n)}{n} = 0.$$

In other words, the primes have density 0 in the natural number. For an awesome elementary proof of this fact, check out http://www.math.udel.edu/~idmercer/primes-density.pdf.

It is interesting to ask how quickly this ratio $\pi(n)/n$ approaches zero.

[2] If you're reading these notes, now is a good time to grab a pen and paper and track the following derivations carefully. The concepts aren't difficult, but unfortunately the notation can make this seem a bit intimidating.

Theorem 2.9. *When n is large, the number of primes less than x is approximately equal to $x/\ln x$. In other words*

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln(x)} = 1.$$

This theorem was conjectured by many in the late 18th century, and was first noticed by observing tables of prime numbers made by hand. It was first proved independently by Hadamard and de la Valée-Poussin in 1896. Their proofs use complex analysis. An elementary proof was discovered by Paul Erdös and Atle Selberg in 1949. A good book for the mathematician's bookshelf is Proofs from the Book, which is a book containing many beautiful and elegant proofs of theorems - many of which are mathematical. It seems a full pdf is posted on one of the authors' websites: http://www.math.boun.edu.tr/instructors/ozturk/proofs.pdf. For an excellent story, check out The First 50 Million Prime Numbers http://people.mpim-bonn.mpg.de/zagier/files/doi/10.1007/BF03039306/fulltext.pdf, or watch the video of Barry Mazur talking about the Riemann Hypothesis http://fora.tv/2014/04/25/Riemann_Hypothesis_The_Million_Dollar_Challenge. We don't have time in this course to go much more into the theory of the distribution of the primes, but there are many accessible introductions to this area. Come talk to me if you're interested in these topics and I can help point you in the right direction.

Finally, it wouldn't be a lecture about the distribution of the primes if we didn't include at least one (or two) open conjectures.

Conjecture 2.10. *(Twin Prime Conjecture) There are infinitely many prime numbers p such that $p+2$ is also prime. These are called twin primes.*

Conjecture 2.11. *(Goldbach's Conjecture) Every even integer larger than 2 can be written as a sum of two primes.*

Extra credit (and fame and fortune) goes to anyone who can solve one of these!

Main Points from Lecture 2:

- Statement of the Fundamental Theorem of Arithmetic and fluency in using uniqueness to prove statements such as Prop 2.3.
- Proof of the infinitude of primes and its variants.
- The statement that $\sum 1/p$ diverges.

3 gcd, lcm and the Euclidean Algorithm

The *greatest common divisor* (gcd) of $n, m \in \mathbb{N}$ is, as the name suggests, the largest integer that divides both of them. Such an integer always exists, as $1 \mid n$ and $1 \mid m$ (so common

divisors exist) and clearly no common divisor can exceed $\min(n, m)$ (so we are taking the maximum over a nonempty finite set. To recap: $1 \leqslant \gcd(n, m) \leqslant \min(n, m)$.

In some texts the gcd is called the hcf ("highest common factor"). We will often denote the gcd of a and b simply by (a, b). In particular, if $(a, b) = 1$ then we say that a and b are relatively prime.

The *least common multiple* (lcm) of $n, m \in \mathbb{N}$ is the smallest integer that both n and m divide. Again, such an integer exists, as n and m both divide nm. So clearly

$$\max(n, m) \leqslant \operatorname{lcm}(n, m) \leqslant nm.$$

Example 3.1. *One way of computing the* gcd *is to factorize. For example* $24 = 2^3 \cdot 3$ *and* $84 = 2^2 \cdot 3 \cdot 7$. *Hence the greatest common factor is* $2^2 \cdot 3 = 12$.

As we'll see shortly, there is a method for computing the gcd *that doesn't involve factoring. This is a good thing as factoring is quite slow.*

Proposition 3.2. *Given* $m, \in \mathbb{N}$ *let* $p_1, \ldots, p_k$ *be the primes dividing* m *or* n *(i.e.,* mn*), and write*

$$m = \prod_{i=1}^{k} p_i^{e_i}, \qquad n = \prod_{i=1}^{k} p_i^{f_i},$$

where $e_i \geq 0$, $f_i \geq 0$. *Then*

(i) $\gcd(m, n) = \prod_{i=1}^{k} p_i^{\min(e_i, f_i)}$;

(ii) $\operatorname{lcm}(m, n) = \prod_{i=1}^{k} p_i^{\max(e_i, f_i)}$;

(iii) $\gcd(m, n) \cdot \operatorname{lcm}(m, n) = mn$;

(iv)

$$\gcd\left(\frac{n}{\gcd(n, m)}, \frac{\operatorname{lcm}(n, m)}{n}\right) = 1.$$

Proof. (i) Suppose that $d \mid m$. Then $dd' = m$ say. so any prime dividing d will divide $dd' = m$. Hence d is of the form

$$d = \prod_{i=1}^{k} p_i^{e_i'}, \qquad \text{where } e_i' \geq 0.$$

Since $d \mid m$, clearly $e_i' \leq e_i$.

If also $d \mid n$, then $e_i' \leq f_i$. so $e_i' \leq \min(e_i, f_i)$. But $\prod_{i=1}^{k} p_i^{\min(e_i, f_i)}$ divides both m and n, so it is their gcd.

(ii) Similary if $m \mid \ell$ and $\ell = \prod_{i=1}^{k} p_i^{f_i'} \cdot \ell'$ say, where ℓ' is a product of primes different from $p_1, \ldots, p_k$, then also $m \mid \prod_{i=1}^{k} p_i^{f_i'}$, so we can take $\ell' = 1$ (i.e., ignore it!). Hence

$$\prod_{i=1}^{k} p_i^{e_i} \mid \prod_{i=1}^{k} p_i^{f_i'},$$

so that $e_i \leq f_i'$. Similarly $n \mid \ell$ gives $f_i \leq f_i'$. Hence $f_i' \geq \max(e_i, f_i)$. But clearly m and n both divide $\prod_{i=1}^k p_i^{\max(e_i,f_i)}$, so this must equal $\operatorname{lcm}(m,n)$.

(iii) If e and f are any two real numbers then

$$\min(e,f) + \max(e,f) = e + f,$$

since one of $\min(e,f)$ and $\max(e,f)$ is e and the other is f. Hence

$$\gcd(m,n) \cdot \operatorname{lcm}(m,n) = \prod_{i=1}^k p_i^{\min(e_i,f_i)+\max(e_i,f_i)} = \prod_{i=1}^k p_i^{e_i+f_i} = mn.$$

(iv) We have using (iii) that

$$\gcd\left(\frac{n}{\gcd(n,m)}, \frac{\operatorname{lcm}(n,m)}{n}\right) = \gcd\left(\frac{n}{\gcd(n,m)}, \frac{m}{\gcd(n,m)}\right) = \frac{\gcd(n,m)}{\gcd(n,m)} = 1.$$

□

Proposition 3.3. *Suppose that $(a,b) = d$. Then $(a/d, b/d) = 1$.*

Proof. Suppose that $c = (a/d, b/d)$ is the gcd. Then $a/d = c \cdot k$ and $b/d = c \cdot l$. Clearing fractions we see that $a = ckd$ $b = cld$. But then cd is a common factor of a and b. Since d was the $gcd(a,b)$ we must have that $cd = d$ and hence $c = 1$. □

Proposition 3.4. *If a, b, c are integers then $(a + cb, b) = (a, b)$.*

Proof. Let $d = (a + cb, b)$ and $e = (a, b)$. Since e divides a and b it surely divides $a + cb$ and b, so $e \mid d$. Conversely, since $d \mid b$, it follows that if $d \mid a + cb$ then $d \mid (a + cb) - cb$ and hence $d \mid a$. Thus d divides a and b whence $d \mid e$. Since d and e divide one another, they must be equal. □

Question 3.5. *If a and b are integers, then what is the set of values that $ax + by$ can take on as x, y range through all integers? We call such numbers* linear combinations of a and b.

This question is related to the postage stamp question which asks if you have postage stamps of values a and b, what are the possible values of total postage that you can make. Note in this case, we are only allowed to nonnegative combinations of a and b, whereas in the Question we allow all integers.

Example 3.6. *What integers are of the form $8x + 12y$?*

First notice that any integer of this form is definitely a multiple of 4*, as it is the* gcd *of* 8 *and* 12*. Further, notice that if we could somehow write $4 = 8x_0 + 12y_0$ then we would be able to write* any *multiple of* 4 *as*

$$4k = 8(kx_0) + 12(ky_0).$$

In this case, it's easy to see that we can indeed write $4 = 8(-1) + 12(1)$. Thus our answer is that

$$\{8x + 12y, \ |x, y \in \mathbb{Z}\} = \{4k \mid k \in \mathbb{Z}\} = 4\mathbb{Z}.$$

In general the following is true:

Theorem 3.7. *The set of linear combinations of two numbers a and b is equal to the set of multiples of (a, b)*

$$\{ax + by, \ |x, y \in \mathbb{Z}\} = \{(a, b)k \mid k \in \mathbb{Z}\} = (a, b)\mathbb{Z}.$$

Proof. As in the example, it is clear that any combination must be a multiple of (a, b) since this divides bother ax and by. What remains to be shown is that (a, b) can indeed be written as a linear combination of a and b. To see this, we argue by contradiction. Suppose that d is the smallest positive integer that is a linear combination of a and b.[3] Say that $d = am + bn$. Now by the division algorithm, we have that

$$a = dq + r, \ 0 \leqslant r < d.$$

Now $r = a - dq = a - (am + bn)q = (1 - m)a + nqb$ is yet another linear combination of a and b. But by assumption, d was the smallest positive such number. Hence $r = 0$, and $a = qd$. Thus $d \mid a$. Similarly $d \mid b$. Hence d is a common divisor of a and b and now the first sentence of this proof shows that $d = (a, b)$. □

Corollary 3.8. *If $(a, b) = 1$ then there exists $m, n \in \mathbb{Z}$ with $am + bn = 1$.*

Theorem 3.7 says something quite useful: That the smallest positive integer which can be written as a linear combination of a and b is (a, b). In the next section, we exploit this to create an algorithm to compute (a, b).

3.1 Finding the gcd without factoring - The Euclidean Algorithm

Given $a, b \in \mathbb{N}$ with $a > b$ (say). Our goal is to compute $g = (a, b)$. We can first divide b into a to get

$$b = qa + r \qquad (q \in \mathbb{N}, 0 \leq r < a).$$

Now notice that

$$(a, b) = (a, b - aq) = (a, r)$$

where the first equality comes from Proposition 3.4.

We can continue with a and r: $a = q_2 + r_2$ say. Then also $g = \gcd(r, r_2)$. Continue in this way until the remainder is 0. Then $g = \gcd(r_k, 0) = r_k$ (the last nonzero remainder).

Example 3.9. *Use the Euclidean algorithm to compute $(51, 87)$:*

$$\begin{aligned} 87 &= 51 + 36 \\ 51 &= 36 + 15 \\ 36 &= 15 \cdot 2 + 6 \\ 15 &= 6 \cdot 2 + 3 \\ 6 &= 3 \cdot 2 + 0. \end{aligned}$$

[3] Why does such a number exist?

Thus $(51, 87) = 3$*, the last nonzero remainder. If we want to write* 3 *as a linear combination of* 51 *and* 87 *we can just step backwards through this:*

$$\begin{aligned} 3 &= 15 - 6(2) \\ &= 15 - (36 - 15 \cdot 2) \cdot 2 = 15(5) - 36(2) \\ &= (51 - 36)(5) - 36(2) = 51(5) - 36(7) \\ &= 51(5) - (87 - 51)(7) = 51(12) - 87(7). \end{aligned}$$

Main Points from Lecture 3:

- How to compute the gcd of two numbers from a factorization or from the Euclidean algorithm
- The gcd of a, b is an integer combination of a and b.
- All integer combinations of a, b are multiples of the gcd.

4 Linear Diophantine Equations

In this lecture we will learn how to solve equations of the form $ax + by = c$ where a, b, c are integers, and we seek integer solutions (x, y). Note that describing the set of real solutions is easy: we can just solve for one variable, say $y = (c - ax)/b$ and then all solutions are of the form $(x, (c - ax)/b)$. However it's not so clear exactly when both of these numbers will be integers. We'll see that the answer comes quickly with the help of the Euclidean algorithm.

Let's work out a few examples to see the salient points:

$$12x + 18y = 10$$

As in our work with linear combinations, we see that the left hand side is always divisible by $(12, 18) = 6$. But the right hand side is not. Therefore this equation has no solution. Hence we have

If $ax + by = c$ has an integer solution then c must be an integer multiple of (a, b).

With that in mind let's try an example that has a chance of having solutions:

$$12x + 18y = 42.$$

We can divide through by $(12, 18) = 6$ to obtain $2x+3y = 7$, and now we can spot a solution in our heads: $2(2) + 3(1) = 7$. It is instructive to find another solution in a more systematic way: Notice that the Euclidean algorithm produces a solution $2(-1) + 3(1) = 1$. We can multiply everything by 7 to obtain $2(-7) + 3(7) = 7$. What is the relationship between our two solutions $(x, y) = (2, 1)$ and $(x, y) = (-7, 7)$? The answer is the following Theorem

Theorem 4.1. *The equation $ax + by = c$ has an integer solution if and only if c is divisible by $d = (a, b)$. If this is the case, then there are infinitely many solutions. If (x_0, y_0) is one particular solution, then all solutions are of the form*

$$x = x_0 - (b/d)n, \quad y = y_0 + (a/d)n$$

where n is an integer.

Proof. By the discussion preceding this theorem, it is clear that a solution exists only if $d \mid c$. In this case a solution always exists as the Euclidean algorithm will always yield a solution to $d = as + bt$. Multiplying both sides by c/d will yield a solution. To see that there are infinitely many solutions, let's check that $(x_0 - (b/d)n, y_0 + (a/d)n)$[4] is indeed a solution:

$$a(x_0 - (b/d)n) + b(y_0 + (a/d)n) = (ax_0 + by_0) - (ab/d)n + (ba/d)n = (c) + 0 = c.$$

Now suppose that (x_1, y_1) is an arbitrary solution. This means that $ax_1 + by_1 = c$. Notice

$$a(x_0 - x_1) + b(y_0 - y_1) = c - c = 0.$$

This means that $a(x_0 - x_1) = -b(y_0 - y_1)$. Let us now divide through by d to obtain

$$\frac{a}{d}(x_0 - x_1) = -\frac{b}{d}(y_0 - y_1).$$

Now since a/d and b/d are relatively prime, we see that $a/d \mid (y_0 - y_1)$ hence $y_0 - y_1 = (a/d)n$. Substituting and canceling, we obtain:

$$\frac{a}{d}(x_0 - x_1) = -\frac{b}{d}\frac{a}{d}n$$

$$(x_0 - x_1) = -\frac{b}{d}n.$$

In summary we have shown that $x_1 = x_0 + \frac{b}{d}n$ and $y_1 = y_0 - \frac{a}{d}n$. The signs are different here as we stated them in the statement of the theorem, but since n is allowed to be positive or negative, this is the same solution set as we required. □

Many times we will be interested in knowing the natural number solutions to an equation. In this case it is possible that there may be no solutions, or only finitely many.[5]

Example 4.2. *(A puzzle my dad asked me when I was a kid) A farmer wishes to buy* 100 *animals and spend exactly* \$100*. Cows are* \$10*, sheep are* \$3 *and pigs are* \$0.50*. Is this possible?*

Solution: The system of equations is

$$c + s + p = 100, \quad 10c + 3s + 0.50p = 100.$$

[4] This is an ordered pair, NOT a gcd

[5] Can you think of an equation with no natural number solutions?

Subsituting $p = 100 - c - s$ *we obtain*

$$10c + 3s + 0.50(100 - c - s) = 100$$

$$20c + 6s + 100 - c - s = 200$$

$$19c + 5s = 100.$$

As $(19, 5) = 1$ *this equation will have infinitely many integer solutions. We can find one by the Euclidean algorithm.*

Scratchwork:

$$19 = 5(3) + 4, \quad 5 = 4 + 1$$

$$1 = 5 - 4 = 5 - (19 - 5(3)) = 19(-1) + 5(4)$$

$$100 = 19(-100) + 5(400).$$

Hence $c = -100, s = 400$ *is one integer solution. By the Theorem, all solutions are of the form*

$$c = -100 - 5n, \quad s = 400 + 19n.$$

Since we are looking for positive integer solutions, we see that $-100-5n > 0$ *and* $400+19n > 0$. *This yields* $-20 > n$ *and* $-21 \leqslant n$, *hence* $n = -21$ *gives the unique solution in positive integers. This yields*

$$c = 5, \quad s = 1, \quad p = 94.$$

4.1 Multivariate linear equations over $\mathbb{Z}$

Given integers $a_1, a_2, \ldots, a_n, b$, how do we find all $(x_1, \ldots, x_n)$:

$$a_1x_1 + \cdots + a_nx_n = b? \tag{1}$$

Important easy cases

1. $g = \gcd(a_1, \ldots, a_n) \nmid b$. Then the LHS of (1) is divisible by g, but the RHS is not, so (1) has no solution in integers.

2. $a_1 = 1$. Then $x_2, x_3, \ldots, x_n$ can be chosen to be *any* integers, with (1) then determining x_1. Clearly this gives *all* solutions of (1) in this case.

Example for 1. The equation $6x_1 + 8x_2 = 11$ has no integer solution, as the LHS is even while the RHS is odd.

Example for 2. The general integer solution of $x_1 + 7x_2 + 9x_3 = 3$ is $(x_1, x_2, x_3) = (3 - 7x_2 - 9x_3, x_2, x_3)$ for x_2, x_3 arbitrary in $\mathbb{Z}$.

General strategy for solving (1)**:** Make linear changes of variables to successively reduce the maximum modulus of coefficients of (1). Keep doing this until either

- get case where $\gcd(a_1, a_2, \ldots, a_n) \nmid b$, so no solution, as in 1. above;
- get a coefficient $= 1$, and so can solve as in 2. above.

This is best illustrated by an example.

Example. Solve $3x + 4y + 5z + 6w = 7$ for integers x, y, z, w.

Solution. Write equation as $3(x + y) + y + 5z + 6w = 7$, and put $u = x + y$. So $3u + y + 5z + 6w = 7$, and $x = u - y$.
Now choose u, z, w arbitrarily in $\mathbb{Z}$. Then $y = 7-3u-5z-6w$ and $x = u-y = -7+4u+5z+6w$. Thus the general solution is $(x, y, z, w) = (-7+4u+5z+6w, 7-3u-5z-6w, z, w)$.

Solution algorithm for solving (1) in integers:

- Pick the a_i of smallest modulus. If $|a_i| = 1$, can solve (1) as in 2. above.
- Otherwise, when smallest modulus of a_i is ≥ 2: For convenience assume $a_1 > 0$ and it is the a_i of smallest modulus. If all the a_i divisible by a_1 and $a_1 \nmid b$, then no solution by 1. above. If all the a_i divisible by a_1 and $a_1 \mid b$, then simply divide the equation by a_1. Now the new a_1 is $= 1$, so can solve it by 2. above.

  Otherwise,, choose an a_1 *not* divisible by a_1 – assume it is a_2. Write $a_2 = qa_1 + a_2'$, where $0 < a_2' < a_1$, and put $u = x_1 + qx_2$. Then (1) becomes

$$a_1x_1 + (qa_1 + a_2')x_2 + a_3x_3 \cdots + a_nx_n = b,$$

  or

$$a_1u_1 + a_2'x_2 + a_3x_3 \cdots + a_nx_n = b. \tag{2}$$

  This new equation (2) has smallest coefficient $a_2' < a_1$. So we can repeat the process. Keep repeating until we get either 1. (so no solution) or 2. (so can write down solution). In the latter case we use the linear equations generated (e.g., $u_1 = x_1 + qx_2$) to get expressions for the original variables.

4.2 Review of Congruences

We now briefly review properties of congruences. A solid mastery of the basics will be necessary for the course. At the end of this section will be many problems designed to give you practice working with congruences. Please let me know if you have any questions either before or after class (or in an email). The material in this section is found in Rosen 4.1 (Introduction to Congruences)

The congruence $a \equiv b \bmod n$ means that the difference $(a - b)$ is divisible by n. In other words, a is equal to a multiple of n plus b. In other words, $a = nq + b$.

Example 4.3.

$$15 \equiv 1 \bmod 7$$

$$-3 \equiv 14 \bmod 17$$

$$10 \equiv 0 \bmod 5$$

I find it helpful to thing of negative numbers as being "less than a multiple of n". For example $30 \equiv -4 \bmod 17$ because "30 is 4 less than a multiple of 17."

There are multiple ways to represent numbers using congruences, and we call each set of equivalences a *congruence class*. For example

$$\cdots -4 \equiv 1 \equiv 6 \equiv 11 \equiv 16 \cdots \bmod 5$$

Is the congruence class of the 1 mod 5.

Definition 4.4. *A complete system of residues modulo n is a set of integers such that every integer is congruent modulo n to exactly one integer in the set. A least positive reside for an integer a is the smallest positive integer b such that $a \equiv b \bmod n$.*

Example 4.5. *Modulo 5, a complete system of residues if $\{0, 1, 2, 3, 4\}$. Another is $\{-2, -1, 0, 1, 2\}$. Yet another is $\{0, 1, 2, 3, 19\}$.*

Arithmetic with congruences behaves extremely well, as we summarize here:

Theorem 4.6. *If a, b, c, d and n are integers with $n > 0$ and $a \equiv b \bmod n$ and $c \equiv d \bmod n$ then*

$$a + c \equiv b + d \bmod n.$$

$$a - c \equiv b - d \bmod n.$$

$$ac \equiv bd \bmod n.$$

The proofs of these exercises follow from the definition of modular arithmetic. The third property is a special case of Problem 2c on the first Homework. We present the proof here.

Proof. If $a \equiv b \bmod n$ then $a = kn + b$ for some integer k. Similarly, $c = \ell n + d$. Thus

$$ac = (kn + b)(\ell n + d) = k\ell n^2 + b\ell n + dkn + db$$

and therefore $ac - bd = n(k\ell n + b\ell + dk)$ is a multiple of n. Hence $ac \equiv bd \bmod n$. □

Example 4.7. *Compute: $93 \cdot 17$ modulo 6. Since $93 \equiv 3 \bmod 6$ and $17 \equiv -1 \bmod 6$ we conclude that $93 \cdot 17 = -3 \bmod 6$.*

Finally, we discuss exponents and their role in modular arithmetic. It is not true that we can reduce the exponents modulo n in computations:

$$2^{10} \equiv 1024 \equiv 4 \bmod 5$$

$$2^0 \equiv 1 \bmod 5.$$

However, two algorithms exist which can help us readily compute high powers of numbers modulo n.

In general, the method that works best is successive squaring. Simply compute successive squares, reducing modulo n when necessary. Then use these numbers to compute the desired power. This is best illustrated by an example.

Example 4.8. *Compute the least positive residue modulo 7 of* 2^{37}*. We compute powers,*
$2^2 \equiv 4$
$2^4 \equiv 4^2 \equiv 2$
$2^8 \equiv 2^2 \equiv 4$
$2^{16} \equiv 4^2 \equiv 2$
$2^{32} \equiv 2^2 \equiv 4$
Thus $2^{37} = 2^{32} \cdot 2^4 \cdot 2^1 = 4 \cdot 2 \cdot 2 \equiv 2 \bmod 7$.

4.3 Lots of Practice Problems with Congruences

(The Starred Problems will appear on the next Homework to be Handed-In)

1. Show that the following congruences hold:

$$13 \equiv 1 \bmod 2, \quad 111 \equiv -9 \bmod 40, \quad 69 \equiv 62 \bmod 7.$$

2. Show that if a is an odd integer then $a^2 \equiv 1 \bmod 8$. (Try to find two proofs, one using modular arithmetic and one that doesn't)

3. Find the least positive residue of $1! + 2! + 3! + \cdots 100!$ modulo 7

4. Show by mathematical induction that if n is a positive integer then $4^n \equiv 1 + 3n \bmod 9$.

5. Find the least positive residue modulo 47 of 2^{200}.

6. $\star$ Show that for every integer n there are infinitely many Fibonacci numbers f_k such that m divides f_k. (Hint: Show that the sequence of least positive residues modulo n of the fibonacci numbers is a repeating sequence.)

7. $\star$ If a, b, c, m are integers such that $m > 0$, $d = (c, m)$ and $ac \equiv bc \bmod m$ then $a = b \bmod m/d$.

Main Points from Lecture 4:

- How to solve systems of linear diophantine equations using the Euclidean algorithm
- Basic properties of congruences

5 Solving Linear Congruences

5.1 Warmup! The Extended Euclidean Algorithm

So far we have only used the Euclidean algorithm in the classical way: Given a and b, use the algorithm to find their gcd. We can then back-substitute to find a solution to the equation $ax + by = \gcd(a, b)$. We now present a way that does this all at once called the extended Euclidean Algorithm.

The basic idea is simple: Given a, b our goal is to find the gcd and also a solution to the equation $ax + by = \gcd(a, b)$. We illustrate with an example: $\gcd(91, 77)$. Notice that the following equations hold obviously.

$$E1: \ 91(1) + 77(0) = 91, \quad (1, 0, 91)$$

$$E2: \ 91(0) + 77(1) = 77, \quad (0, 1, 77)$$

We have written the coefficients to the right. Now notice what happens when we subtract the second equation from the first $E3 = E1 - E2$.

$$E3: \ 91(1) - 77(1) = 14, \quad (1, -1, 14)$$

Now we can set $E4 = E2 - 5 \cdot E3$:

$$E4: \ 91(-5) + 77(6) = 7, \quad (-5, 6, 7).$$

$$E5: \ 91(11) + 77(-13) = 0.$$

What sort of magic is this? If you look at the numbers on the right side of the equation, they are simply the remainders that come up in the Euclidean algorithm. Hence 7 is the last nonzero remainder so it is the gcd. Hence we have found $91(-5)+77(6) = 7$. This algorithm can be done rapidly if we ignore writing the equations and just work with the vectors.

Example 5.1. *Compute* $\gcd(561, 306)$ *using the extended Euclidean Algorithm: We begin with the vectors* $v_0 = (1, 0, 561)$ *and* $v_1 = (0, 1, 306)$ *and just subtract one from the other sucessively:*

$$\begin{aligned} v_0 &= (1, 0, 561) \\ v_1 &= (0, 1, 306) \\ v_2 &= (1, -1, 255), \quad (v_2 = v_0 - v_1) \\ v_3 &= (-1, 2, 51), \quad (v_3 = v_1 - v_2) \\ v_4 &= (6, -11, 0), \quad (v_4 = v_2 - 5v_3). \end{aligned}$$

Thus the gcd is 51 and $561(-1) + 2(306) = 51$.

For completeness the full algorithm is detailed below:

Proposition 5.2. *Given m, n and $g = \gcd(m, n)$, there exist integers u, w:*

$$um + wn = g \qquad \textit{Bézout's Identity}\ .$$

These integers can be found as follows:
Put $v_0 = (1, 0, m)$, $v_1 = (0, 1, n)$, and if $m = qn + r$ then put

$$v_2 = v_0 - qv_1 = (1, -q, m - qn) = (1, -q, r).$$

Apply the same procedure to v_1 and v_2: if $n = q_2 r + r_2$ put $v_3 = v_1 - q_2 v_2$. Continue in this way until the third component of the current v is 0. Then the previous v is (u, w, g) with $um + wn = g$.

Proof. Clearly v_0 and v_1 lie on the plane $mx + ny = z$ in $\mathbb{R}^3$. Further, if v_i and v_{i+1} lie on this plane, then so does $v_{i+2} = v_i - q_* v_{i+1}$ (for some $q_* \in \mathbb{N}$). Hence, by induction, all of $v_0, v_1, v_2, \ldots$ lie on this plane. In particular (u, w, g) , which is one of the v_i's, lies on this plane. Hence $um + wn = g$. □

We now solve congruences of the form

$$ax \equiv c \bmod n.$$

Recall that from the definition this means that $ax - c = ny$ for some integer y. Rewriting we can think of this as a linear diophantine equation

$$ax - ny = c.$$

(Notice the roles of the letters is slightly different here than it was before.) Hence for a solution to exist, if $d = (a, n)$, it must be the case that $d \mid c$. Further, if one solution (x_0, y_0) exists then there are infinitely many solutions, given by Theorem 4.1:

$$x = x_0 + (n/d)t, \quad y = y_0 + (a/d)t.$$

Since we are solving a congruence, however, it makes sense to talk about the congruence classes which are solutions. In other words, we want to know how many incongruent solutions there are to the equation modulo n.

Theorem 5.3. *If (a, n) divides c then the congruence $ax \equiv c \bmod n$ has exactly d incongruent solutions modulo n.*

Proof. Let x_0 be a solution to the congruence. By the discussion above, we know that all solutions are of the form $x_0 + (n/d)t$ where $t \in \mathbb{Z}$. We now see how many of these are incongruent modulo n. Suppose that we have

$$x_1 = x_0 + (n/d)t_1, \quad x_2 = x_0 + (n/d)t_2.$$

Then $x_1 - x_2 = (n/d)(t_1 - t_2)$. Hence x_1 and x_2 are congruent modulo n if and only if $(n/d)(t_1 - t_2)$ is a multiple of n. This occurs exactly when there exists an integer ℓ such that $(n/d)(t_1 - t_2) = \ell n$. Simplifying we see $(t_1 - t_2) = \ell d$, which is equivalent to

$$t_1 \equiv t_2 \bmod d.$$

Summing up, the solutions x that are inequivalent modulo n are exactly the ones that have corresponding values of t that are inequivalent modulo d. There are d such classes for t, which proves the theorem. □

Note the special case when $d = (a, n) = 1$.

Corollary 5.4. *If $(a, n) = 1$ then the congruence $ax = c \bmod n$ has a unique solution.*

An Algorithm: To solve a congruence of the form $ax \equiv c \bmod n$ we can proceed algorithmically:

First we check the necessary condition that $d = (a, n)$ divides c.

If so, then we expect there to be d distinct solutions modulo n. To find one of these, write

$$d = ax_0 + ny_0$$

Then going modulo n, we see that x_0 is a solution to the congruence.

The set of all solutions is then $\{x_0, x_0 + (n/d), x_0 + 2(n/d), \ldots, x_0 + (d-1)(n/d)\}$. We might represent this as the set $x_0 + t(n/d), 0 \leqslant t \leqslant d - 1$.

Example 5.5. *To find all solutions to $9x \equiv 12 \bmod 15$, we first check that $d = (9, 15) = 3$ indeed divides 12. By the Theorem there will be 3 inequivalent solutions. By the Euclidean algorithm, we see that*

$$d = 15(-1) + (2)(9).$$

Hence $9(2) \equiv 3 \bmod 15$ and thus $9(8) \equiv 12 \bmod 15$. Thus $x = 8$ is a solution. All solutions will therefore be of the form $8 + (15/3)t = 8 + 5t$ for $t = 0, 1, 2$. Hence the congruence classes of the solutions are $3, 8, 13$.

Main Points from Lecture 5:

- Using the Extended Euclidean algorithm to write $\gcd(a, b)$ as an integer combination of a and b.
- The number of solutions to the congruence $ax \equiv c \bmod n$ is $d = (a, n)$.
- All solutions are of the form $x_0 + t(n/d)$ for $t = 0, 1, \ldots, d - 1$.

6 Modular Inverses and the Chinese Remainder Theorem

A solution to the congruence $ax \equiv 1 \bmod n$ is called an *inverse* of a modulo n. By Corollary 5.4, this solution is unique. Inverses are incredibly useful, because if you have one, then it allows you to easily solve all other congruences. Indeed, if you know that $ax_0 \equiv 1 \bmod n$ then $a(x_0 b) \equiv b \bmod n$.

Example 6.1. *Find all solutions to* $7x \equiv 1 \bmod 31$. *We use the Euclidean algorithm to determine that*

$$(31)(-2) + 7(9) = 1$$

and hence $7(9) \equiv 1 \bmod 31$. *Thus* 9 *and* 7 *are inverses modulo* 31.

An important special case is when $n = p$ is a prime number. In this case, every nonzero number a has a unique inverse. For example, we list the inverses modulo 11 in the following table

a	1	10	2	3	5	7
a^{-1}	1	10	6	4	9	8

Notice in this table we have chosen the representatives $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ to represent the nonzero congruence classes modulo p. However, if we chose $\{-1, -2, -3, -4, -5, 1, 2, 3, 4, 5\}$, the table would be:

a	1	-1	2	3	5	-4
a^{-1}	1	-1	-5	4	-2	-3

Theorem 6.2. *A number* a *is equal to its own inverse modulo* p *if and only if* $a \equiv \pm 1 \bmod p$.

Proof. Since $1^2 = (-1)^2 = 1$, we see that ± 1 are their own inverses. To see that there are no others, notice that $a^2 \equiv 1 \bmod p$ means that $a^2 - 1$ is a multiple of p. But then p must divide $(a+1)(a-1)$ meaning that p must divide either $a+1$ or $a-1$. Hence $a \equiv \pm 1 \bmod p$. □

Theorem 6.3 (Chinese Remainder Theorem)**.** *Given* $m_1, \ldots, m_k \in \mathbb{N}$ *with* $\gcd(m_i, m_j) = 1 \quad (i \neq j)$ *("pairwise coprime"), and* $a_1, \ldots, a_k \in \mathbb{Z}$, *then the system of congruences*

$$\begin{aligned} x &\equiv a_1 \bmod m_1 \\ x &\equiv a_2 \bmod m_2 \\ &\vdots \\ x &\equiv a_k \bmod m_k \end{aligned}$$

has a solution $x \in \mathbb{Z}$.

Proof. In fact x can be constructed explicitly. For $i = 1, \ldots, k$ define m_i^* to be the inverse modm_i of $m_1 \ldots m_{i-1} m_{i+1} \ldots m_k$, so that

$$m_1 \ldots m_{i-1} m_i^* m_{i+1} \ldots m_k \equiv 1 \bmod m_i.$$

Then $x = \sum_{i=1}^{k} a_i m_1 \ldots m_{i-1} m_i^* m_{i+1} \ldots m_k \equiv a_i \bmod m_i$ for $i = 1, \ldots, k$, because every term except the ith is divisible by m_i. □

Remark 6.4. *Notice that if $n \in \mathbb{N}$ with $n = \prod_{i=1}^{k} p_i^{e_i}$ where p_i are distinct primes. Then $x \equiv a \bmod n$ is the unique solution to the system of congruences*

$$x \equiv a \bmod p_i^{e_i}, \quad i = 1, \ldots, k.$$

Indeed, $(x - a)$ is divisible by n if and only if it is divisible by $p_i^{e_i}$ for all i.

Then, if x_0 is one solution to this set of congruences, it's easy to see (how?) that the general solution is $x = x_0 + \ell m_1 \cdots m_k$ for any integer ℓ. In particular, there is always a choice of ℓ giving a unique solution x in the range $0 \leq x < m_1 \cdots m_k$ of the set of congruences.

6.1 Examples and Exercises

Example 6.5. *This example comes from the ancient Chinese puzzle (third century C.E.) in* Master Sun's Mathematical Manual. *Find a number that leaves a remainder 1 when divided by 3, a remainder of 2 when divided by 5 and a remainder of 3 when divided by 7.*

This system of equations is

$$\begin{aligned} x &\equiv 1 \bmod 3 \\ x &\equiv 2 \bmod 5 \\ x &\equiv 3 \bmod 7. \end{aligned}$$

We have $k = 3$ equations, so following the solution in the theorem, we form all products of $k - 1 = 2$ moduli and compute inverses.

$$\begin{aligned} (m_1) &\equiv (5 \cdot 7)^{-1} \bmod 3 \\ (m_2) &\equiv (3 \cdot 7)^{-1} \bmod 5 \\ (m_3) &\equiv (3 \cdot 5)^{-1} \bmod 7 \end{aligned}$$

We can check that these numbers are $(m_1', m_2', m_3') = (2, 1, 1)$. Hence

$$x = (1) \cdot 2 \cdot 5 \cdot 7 + (2) \cdot 3 \cdot 1 \cdot 7 + (3) \cdot 3 \cdot 5 \cdot 1 = 70 + 42 + 45 = 157.$$

Is a solution. Furthermore, since $2 \cdot 5 \cdot 7 = 105$ all solutions are of the form $157 + 105n$. In particular, the smallest positive solution is $x = 52$.

There is also an iterative way to find a solution

Example 6.6.

$$\begin{aligned} x &\equiv 1 \bmod 5 \\ x &\equiv 2 \bmod 6 \\ x &\equiv 3 \bmod 7. \end{aligned}$$

The first equation says $x = 5t + 1$*, and hence the second says* $5t + 1 \equiv 2 \bmod 6$*. This is the same as*

$$5t \equiv 1 \bmod 6$$

we can multiply both sides by 5 *(the inverse of* 5*) to obtain*

$$t \equiv 5 \bmod 6$$

. Hence $t = 6s + 5$*, so that* $x = 30s + 26$*. Finally we sub in to obtain*

$$30s + 26 \equiv 3 \bmod 7$$

$$2s \equiv 5 \bmod 7$$

$$s \equiv 6 \bmod 7.$$

Thus $s = 6$*, and* $x = 30(6) + 26 = 206$*.*

This second method allows an algorithm for solving systems of congruences even in the case when the m_i are not relatively prime (when a solution exists. See Exercises 15-20 in Rosen 4.3) For this course it is important to know the statement and proof of the Chinese Remainder Theorem. For solving practical problems, either method is acceptable.

Remark 6.7. *Notice that we can in fact solve any system of congruences of the form* $ax = b \bmod m$ *using the methods above, provided that* a *has an inverse modulo* m*. The first step is just to multiply both sides of the congruence by* a^{-1}*.*

6.2 Exercises

1. Find all solutions to

   - $$x \equiv 4 \bmod 11$$
     $$x \equiv 3 \bmod 17$$

   - $$x \equiv 0 \bmod 2$$
     $$x \equiv 0 \bmod 3$$
     $$x \equiv 1 \bmod 5$$
     $$x \equiv 6 \bmod 7.$$

2. Show that if $(a, b) = 1$ and c is an integer, then there exists an integer n such that $(an + b, c) = 1$.

3. Solve the system:

$$x \equiv 4 \bmod 6$$
$$x \equiv 13 \bmod 15$$

Note that the moduli are NOT relatively prime.

Main Points from Lecture 6:

- The inverse of a exists modulo n if and only if $(a, n) = 1$.
- If $ax + ny = 1$ then x is the inverse of a modulo n.
- Method and proof of the Chinese Remainder Theorem

7 Solving Polynomial Equations

Let's begin with a one sentence summary of what the Chinese Remainder Theorem says from last time:

Knowing a number x modulo N is equivalent to knowing x modulo each of its relatively prime factors.

For example, knowing that $x \equiv 27 \bmod 30$ is the same as knowing

$$x \equiv 1 \bmod 2, \quad x \equiv 0 \bmod 3, \quad x \equiv 2 \bmod 5.$$

Example 7.1. *How many solutions does the equation $x^2 \equiv 1 \bmod pq$ where p and q are distinct odd primes?*

Solution: By the CRT we know that this is equivalent to finding solutions of the form $x^2 \equiv 1 \bmod p$ and $x^2 \equiv 1 \bmod q$. These each have exactly two solutions: $+1, -1$, (we exclude the case $p = 2$ because in this case $1 = -1$. Hence in total there are four possible systems of congruences, so in total there are 4 solutions.

For example the square roots of 1 modulo 77 are equal to $\{1, 34, 43, 76\}$

More generally, we can show that if $N = p_1 \cdots p_k$ is a product of distinct odd primes then $x^2 \equiv 1 \bmod N$ has 2^k distinct solutions modulo N. We say that 1 has 2^k square roots.

This motivates a question: If we know $x \bmod p$, what can we say about $x \bmod p^2$?

Example 7.2. *Solve the polynomial congruence $2x^3 + 7x - 4 \equiv 0 \bmod 200$.*

Solution: Notice that $200 = 2^3 \cdot 5^2 = 8 \cdot 25$ This problem is equivalent to solving the system of equations

$$2x^3 + 7x - 4 \equiv 0 \bmod 8$$

$$2x^3 + 7x - 4 \equiv 0 \bmod 25.$$

We can check that $x \equiv 4 \bmod 8$ *(just by trial and error) and later today we'll show that* $x \equiv 16 \bmod 25$. *These two linear equations combine by the CRT to show that the solution is* $x \equiv 116 \mod 200$.

The CRT provides a very effective way of chopping up a problem into smaller pieces by turning the problem "Solve $f(x) \equiv 0 \mod n$" into a system of problems "Solve $f(x) \equiv 0 \bmod p_i^{e_i}$" if $n = \prod p_i^{e_i}$. In this section we will develop a method for solving polynomial equations of the form $f(x) = 0 \bmod p^e$.

Continuing the example, notice that to solve the equation $2x^3 + 7x - 4 \bmod 5$ we only need to test $0, 1, 2, 3, 4$. This is reasonably quick. And we see that all solutions satisfy $x \equiv 1 \bmod 5$. However, we'd like to avoid check all the numbers $0, \ldots, 24$ to solve this equation modulo 25. Notice that any solution x to

$$2x^3 + 7x - 4 \equiv 0 \bmod 25$$

is also a solution modulo 5. Hence $x \equiv 1 \bmod 5$. Hence $x = 5t + 1$. Substituting we see that

$$2(5t+1)^3 + 7(5t+1) - 4 \equiv 0 \bmod 25$$

$$2(\cancel{(5t)^3} + \cancel{3 \cdot (5t)^2} + 3 \cdot 5t + 1) + 35t + 7 - 4 \equiv 0 \bmod 25.$$

$$2(3 \cdot 5t + 1) + 35t + 7 - 4 \equiv 0 \bmod 25.$$

$$65t + 5 \equiv 0 \bmod 25.$$

$$15t + 5 \equiv 0 \bmod 25.$$

(Notice that everything on the left was divisible by 5). We can eliminate a factor of 5 by Exercise 4.3.7. Hence

$$3t + 1 \equiv 0 \bmod 5.$$

which has $t \equiv 3 \bmod 5$ is its unique solution. Hence $x \equiv 16 \bmod 25$ is the unique solution to our original equation. We say that $x \equiv 16 \bmod 25$ is a "lift" of the solution $x \equiv 5 \bmod 5$.

Theorem 7.3. *Suppose that* $f(x)$ *is a polynomial with integer coefficients and* k *is an integer with* $k \geqslant 2$. *Suppose further that* r *is a solution of the congruence* $f(x) \equiv 0 \bmod p^{k-1}$. *Then,*

1. *if* $f'(r) \neq 0 \bmod p$, *then there is a unique integer* t, $0 \leqslant t < p$, *such that* $f(r + tp^{k-1}) \equiv 0 \bmod p^k$ *given by*
$$t \equiv -f'(r)^*(f(r)/p^{k-1}) \bmod p,$$
where $f'(r)^*$ *is the inverse of* $f'(r)$ *modulo* p;

2. *if* $f'(r) \equiv 0 \bmod p$ *and* $f(r) \equiv 0 \bmod p^k$, *then* $f(r + tp^{k-1}) \equiv 0 \bmod p^k$ *for all integers* t.

3. *if* $f'(r) \equiv 0 \bmod p$ *and* $f(r) \neq 0 \bmod p^k$, *then* $f(x) \equiv 0 \bmod p^k$ *has no solutions with* $x \equiv r \bmod p^{k-1}$.

7.1 A Fireside Chat With Hensel's Lemma

It's fair to say that the statement of Hensel's Lemma is a bit intimidating - but that doesn't mean that the concept is difficult. Hopefully the following chat between Hensel's Lemma and some guy named Earle will help with the concept.

Earle: So what's the deal with you, anyway?

HL: Well, I provide a method of telling how many solutions you have modulo p^2 given solutions modulo p.

Earle: Is that it?

HL: Well, I can inductively be used to find solutions modulo p^3, p^4, and on and on. In an advanced course, you might wonder if there's a limit term at p^∞, and the answer is YES, and that concerns the p-adics and ...

Earle: Whoa, let's worry about that in the advanced course. So tell me, in layman's terms what your lemma does.

HL: Well suppose you've got a polynomial, and all you know is the following table of numbers

x	0	1	2	3	4
$f(x)$	10	1	6	-7	25

Then what are the solutions to $f(x) \equiv 0 \bmod 5$?

Earle: Well you just have to check the equivalence classes, so it looks like $x \equiv 0$ and $x \equiv 4$ are the only solutions.

HL: Good. Now what can you say about the solutions to $f(x) \equiv 0 \bmod 25$?

Earle: Well I don't know. I mean I know that x has to be 0 or 4 modulo 5. So that means x has to be either $0, 5, 10, 15, 20$ or $4, 9, 14, 19, 24$.

HL: Do you know anything else?

Earle: Well from the given information I guess I know that $f(4) = 25$ so $f(4) \equiv 0 \bmod 25$. So that's one solution. And I know that $f(0) = 10$ so 0 is NOT a solution modulo 25.

HL: Good. That's really all you can say. But now what if I told you that $f'(0) = 2$?

Earle: Oh, then the theorem says that $f(x) \equiv 0 \bmod 25$ should have a unique solution... or something, right? But I don't remember the formula, there's a t and a $r + tp^{k-1}$ blech.

HL: Mostly right. It says that there is a unique solution with x congruent to 0 modulo 5. So only one of the numbers $0, 5, 10, 15, 20$ is going to be a solution. And check out Corollary 8.2 for a more convenient way to work. This wasn't written on the board during class, but it's very helpful. It says that a solution mod 25 will just be given by

$$r_2 = r - f'(r)^* f(r)$$

In other words, take the root from the previous step and then subtract off a correction term.

Earle: This is like Newton's method, isn't it?

HL: It is indeed! So

$$r_2 = 0 - ((2)^*)(10)$$

Earle: Wait a second, when you take the inverse of 2, is that mod 5 or mod 25?

HL: Well it turns out that it won't matter, but in the statement of the lemma, this inverse business is ALWAYS just modulo p. So yea, you just want the inverse mod 5.

Earle: Ok, so $r_2 = 0 - (3)(10) = -30$ which is -5 modulo 25 which is 20 modulo 25. Pretty cool. I don't even have to check those other candidates among $0, 5, 10, 15, 20$. I know that 20 has got to be the solution.

HL: Ok, now what if I told you that $f'(4) = 0$.

Earle: Oh that'd be a sad day.

HL: Not so much. My Theorem says that if $f'(4)$ is zero - then either ALL of the lifts of 4 are solutions, or NONE of the lifts are solutions.

Earle: Oh, so I could just check one to see whether or not it was a solution.

HL: Yea, and you may as well just check 4 itself.

Earle: Ok $f(4) = 25 \equiv 0 \bmod 25$ so that means we have one solution, so they must all be solutions! So all of $4, 9, 14, 19, 24$ are solutions.

HL: Yep. And summing up, that means that 20 and $4, 9, 14, 19, 24$ are the solutions modulo 25.

Earle: That was pretty easy. Can we do another step?

HL: Sure. To lift that 0 mod 5 solution we can just do another round of that formula:

$$r_3 = r_2 - f'(r)^*(f(r_2))$$

where r was the original solution mod 5.

Earle: Wait wait, this is the part that really confuses me. First of all didn't you mean to put an r_2 into that $f'(r)^*$ term?

HL: Well I could have, but it won't make a difference. Remember, r_2 and r are the same mod 5. So that means that $f'(r)$ and $f'(r_2)$ are the same modulo 5. And since we take our inverses mod 5, this is all that matters.

Earle: Oh ok. Ohhh so that inverse term... it's gonna be 3 again, cause at every step I'm just applying $f'(0)^*$? So

$$r_3 = 20 - 3f(20)$$

HL: Looks good to me.

This concludes our fireside chat with Hensel's Lemma. I hope this has been helpful.

8 The Proof of Hensel's Lemma and Example

To prove Hensel's Lemma we will need the following Lemma from Taylor series. Notice that this Taylor series is finite since the derivatives of a polynomial are eventually all zero. Indeed, if $f(x)$ is a polynomial of degree n then the $(n+1)$-st derivative is always zero.

Lemma 8.1. *If $f(x)$ is a polynomial of degree n then*

$$f(a+b) = f(a) + f'(a)b + f''(a)b^2/2! + \cdots f^{(n)}(a)b^n/n!$$

where the coefficients $(f(a), f'(a), f''(a)/2!, \ldots)$ are polynomials in a with integer coefficients.

Proof. All but the integrality of the coefficients follows from the Taylor expansion about the point $x = a$. To see that the coefficients are integers, just note that the k-th derivative of x^m is

$$m(m-1)(m-2)(m-3)\cdots(m-k+1)x^{m-k}$$

Notice that this is divisible by $k!$ since

$$\frac{m(m-1)(m-2)(m-3)\cdots(m-k+1)}{k!} = \binom{m}{k}.$$

□

Proof of Hensel's Lemma. Recall that we are assuming that r is a solution of $f(r) \equiv 0 \bmod p^{k-1}$. We seek solutions modulo p^k that are congruent to r modulo p^{k-1}. In other words, we are looking for solutions modulo p^k of the form $r + tp^{k-1}$. We will seek precise conditions for t. Notice that the previous lemma says

$$f(r + tp^{k-1}) = f(r) + f'(r)tp^{k-1} + \frac{f''(r)}{2!}t^2p^{2k-2} + \cdots .$$

Notice that all terms but the first two are zero modulo p^k (since $k \geqslant 2$). Hence

$$f(r + tp^{k-1}) = f(r) + f'(r)tp^{k-1} \bmod p^k.$$

Since we are assuming $r + tp^{k-1}$ is a solution modulo p^k the left hand side is zero and we can conclude that

$$f'(r)tp^{k-1} \equiv -f(r) \bmod p^k.$$

But we are assuming that $f(r) \equiv 0 \bmod p^{k-1}$ so by homework 4.3.7 again, we see that

$$f'(r)t \equiv -f(r)/p^{k-1} \bmod p.$$

Now we just examine cases. If $f'(r)$ is nonzero modulo p then this equation must have a unique solution for t

$$t \equiv -f'(r)^* f(r)/p^{k-1} \bmod p$$

where the * denotes inverse modulo p. This establishes part (1) of the theorem.

If $f'(r) \equiv 0 \bmod p$. Then the equation is of the form

$$0t \equiv -f(r)/p^{k-1} \bmod p.$$

If the right hand side is nonzero, this has no solutions. If the right hand side is zero, then any value of t is a solution. This proves (2) and (3). □

One Corollary of this theorem provides a particularly easy method for computing "lifts" of solutions modulo p.

Corollary 8.2. *Suppose that r is a solution to $f(r) \equiv 0 \bmod p$ where p is prime. If $f'(r) \neq 0 \bmod p$ then there is a unique solution r_k modulo p^k for each $k = 2, 3, \ldots$ such that*

$$r_k = r_{k-1} - f(r_{k-1})f'(r)^*.$$

where $f'(r)^*$ is the inverse of $f'(r)$ modulo p.

Proof. We see from the hypotheses of Hensel's lemma that we are in Case (1). Hence r lifts to a unique solution r_2 modulo p^2 with $r_2 = r + tp$ with $t = -f'(r)^*(f(r)/p)$. Hence

$$r_2 = r - f'(r)^*(f(r)).$$

Now since $r_2 \equiv r \bmod p$. It follows that $f'(r_2) \equiv f'(r) \neq 0 \bmod p$. Using Hensel's lemma again, we see that the unique solution modulo p^3 is then $r_3 = r_2 - f(r_2)f'(r)^*$. Continuing this way we see that we can obtain solutions modulo p^k for all k. □

Example 8.3. *Find the solutions of*

$$x^3 + x^2 + 29 \equiv 0 \bmod 25.$$

Solution: Let $f(x) = x^3 + x^2 + 29$. *Then the solutions modulo* 5 *are* $x \equiv 3 \bmod 5$. *Since* $f'(x) = 3x^2 + 2x$, *we have* $f'(3) \equiv 3 \neq 0$ *modulo* p. *Also* $f(3) = 15$ *Hence the unique solution modulo* 25 *is*

$$r_2 \equiv 3 - 15(3)^{-1} \equiv 3 - 15(2) \equiv -27 \equiv 23$$

is the unique solution modulo 25.

Rosen has a few more examples worked in detail. I emailed scanned copies of this section in the book. If you'd like me to resend it to you, please let me know.

8.1 Exercises

1. Find all solutions to $x^2 + 4x + 2 = 0 \bmod 7^3$.
2. Find all solutions to $x^2 + x + 34 = 0 \bmod 81$.
3. How many incongruent solutions are there to $x^5 + x - 6 \equiv 0 \bmod 144$?

Main Points from Lecture 7 and 8:

- How to apply the Chinese Remainder Theorem to solving equations modulo N via factorization.
- The method and proof of Hensel's Lemma.

9 The field $\mathbb{F}_p$

For the next two weeks we will be studying in detail the integers modulo p, where p is prime. The set of congruence classes modulo p forms a field, which we now review:

9.1 Fields

A *field* F is a set supplied with two binary operations '+' and '×' (i.e., maps from $F \times F$ to F), and containing special elements 0 and 1 such that

- F is an abelian group under +, with 0 as its identity element, and $-a$ as the (additive) inverse of $a \in F$;
- $F^\times = F \setminus \{0\}$ is an abelian group under ×, with 1 as the identity element, and a^{-1} the (multiplicative) inverse of $a \in F^\times$;

- The Distributive Law holds: for all $a, b, c \in F$ we have

$$a \times (b + c) = a \times b + a \times c.$$

  This describes how $+$ and $\times$ interact in F.

As a consequence of these rules, we can show (won't prove)

Proposition 9.1. *For $a, b \in F$ we have*

- $a \times 0 = 0$*;*
- $a \times (-b) = -(a \times b)$*;*
- *Cancellation Law: if $a \times b = 0$ then $a = 0$ or $b = 0$ (or both).*

Examples of fields are: the complex numbers $\mathbb{C}$, the real numbers $\mathbb{R}$, the rational numbers $\mathbb{Q}$, and the finite fields $\mathbb{F}_p$ for p prime – see below.

9.1.1 Construction of $\mathbb{F}_p$

Start with the integers $\mathbb{Z}$ and a prime p, and define an equivalence relation on $\mathbb{Z}$ by saying that two integers a and b are equivalent if $a \equiv b \bmod p$. This defines an equivalence relation on $\mathbb{Z}$. The elements of $\mathbb{F}_p$ are the equivalence classes under this relation. Taking equivalence class representatives to be $0, 1, 2, 3, \ldots, p-1$, we can effectively regard $\mathbb{F}_p$ as the set $\{0, 1, 2, 3, \ldots, p-1\}$. Addition, negation, multiplication and reciprocals are performed modulo p, so that the result can always be chosen to be in $\{0, 1, 2, 3, \ldots, p-1\}$.

For example, in $\mathbb{F}_7$, $3 + 4 = 0$ as in $\mathbb{Z}$ we have $3 + 4 = 7 \equiv 0 \bmod 7$. Hence also $-3 = 4$ and $-4 = 3$ in $\mathbb{F}_7$. Further, because $3 \times 5 = 15 \equiv 1 \bmod 7$, we have $3^{-1} = 5$ and $5^{-1} = 3$ in $\mathbb{F}_7$.

9.2 Solving equations in $\mathbb{F}_p$

We now restrict our congruences to a prime modulus p, and consider the solutions of equations $f(x) = 0$ for $f(x) \in \mathbb{F}_p[x]$ and $x \in \mathbb{F}_p$. This is equivalent, for $f(x) \in \mathbb{Z}[x]$, of solving $f(x) \equiv 0 \bmod p$ for $x \in \{0, 1, 2, \ldots, p-1\}$.

Theorem 9.2. *A nonzero polynomial $f \in \mathbb{F}_p[x]$ of degree n has at most n roots x in $\mathbb{F}_p$.*

Proof. Use induction: for $n = 1$, $f(x) = ax + b$ say, with $a \neq 0$, whence $f(x) = 0$ has a solution $x = -a^{-1}b$ in $\mathbb{F}_p$.

Now assume $n \geqslant 1$ and that the result holds for n. Take $f(x) \in \mathbb{F}_p[x]$ of degree $n + 1$. If $f = 0$ has no roots $x \in \mathbb{F}_p$ the the result is certainly true. Otherwise, suppose $f(b) = 0$ for some $b \in \mathbb{F}_p$. Now divide $x - b$ into $f(x)$, (i.e., one step of the Euclidean algorithm for polynomials) to get $f(x) = (x - b)f_1(x) + r$ say, where f_1 is of degree n, and $r \in \mathbb{F}_p$. Putting $x = b$ shows that $r = 0$. Hence $f(x) = (x - b)f_1(x)$, where f_1 has, by the induction hypothesis, at most n roots $x \in \mathbb{F}_p$. So f has at most $n+1$ roots $x \in \mathbb{F}_p$, namely b and those of $f_1 = 0$. Hence the result is true for $n + 1$ and so, by induction, true for all $n \geq 1$. □

Note that the proof, and hence the result, holds equally well when $\mathbb{F}_p$ is replaced by *any* field F.

Question. Where in the above proof was the fact that we were working over a field used? There were two places. Once in the base case when $n = 1$ and then once again when we concluded that if $(x - b)f_1(x) = 0$ then either one of the factors must equal zero.

Remark 9.3. *Note that this theorem is not true if we work modulo a composite number. For instance, the polynomial $x^2 - 1$ has 4 roots modulo 15. It's instructive to really think the above proof through using this polynomial to see where it breaks down. Notice also that*

$$x^2 - 1 = (x - 1)(x + 1) = (x - 4)(x + 4)$$

has two different factorizations!

9.3 Some Special Congruences - Wilson's Theorem, Fermat's Theorem and Euler's Theorem

We now prove some special congruences that will be useful for the rest of the course.

Theorem 9.4 (Wilson's Theorem)**.** *If p is prime then $(p - 1)! \equiv -1 \bmod p$.*

Proof. Recall that by Theorem 6.2 the only numbers that are their own inverse modulo p are 1 and $p - 1$. Hence if we rearrange the terms

$$(p - 1)! = 1 \cdot 2 \cdots (p - 1) = 1 \cdot (p - 1) \cdot (2 \cdot 2^{-1}) \cdots (a \cdot a^{-1})$$

where on the right we have paired each number a with its inverse. It's clear that this product is equal to $(p - 1) \cdot (1 \cdots 1) \equiv -1 \bmod p$. □

Theorem 9.5 (Fermat's Little Theorem)**.** *If p is a prime then for all integers a, $a^p \equiv a \bmod p$. If further, $a \not\equiv 0 \bmod p$ then $a^{p-1} \equiv 1 \bmod p$.*

Proof. Notice that the second statement follows from the first by multiplying both sides by a^{-1} (which exists if and only if $a \not\equiv 0 \bmod p$.) To prove the first statement we argue by induction. Clearly the statement if true if $a = 0$. Now notice that by the binomial theorem

$$(a + 1)^p = a^p + \binom{p}{1} a^{p-1} + \binom{p}{2} a^{p-2} + \cdots + \binom{p}{p-1} a + 1.$$

Recall that $\binom{p}{k}$ is divisible by p for $1 \leqslant k \leqslant p - 1$ (Why?). Hence

$$(a + 1)^p \equiv a^p + 1 \bmod p.$$

Now by induction we have

$$(a + 1)^p \equiv a + 1.$$

This proves the result for all positive integers, and since this is a statement about congruences, this takes care of all the equivalence classes (including the class of negative ones). □

In class, we used Fermat's Theorem to provide another proof of Wilson's Theorem. See if you can fill in the details! The rough outline is that Fermat's Theorem says that each nonzero element of $\mathbb{F}_p$ is a solution to the equation $x^{p-1} - 1 = 0$. Now use the fact that you've found $p - 1$ roots of this equation, and a little factorization to finish the job!

Main Points from Lecture 9:

- Definition and basic properties of a field
- Definition of $\mathbb{F}_p$
- The number of roots in $\mathbb{F}_p$ of a polynomial of degree n is at most n
- Statement and two proofs of Wilson's Theorem
- Statement (and your favorite proof) of Fermat's Little Theorem

10 Primitive Roots and the Structure of $\mathbb{F}_p$

10.1 A Warmup for Things to Come:

We start today with defining the Euler phi-function: $\varphi(n)$ is defined to be the number of positive integers not exceeding n that are coprime to n.

n	1	2	3	4	5	6	7	8	9	10	11	12
$\varphi(n)$	1	1	2	2	4	2	6	4	6	4	10	4

Figure 2: Some Values of $\varphi(n)$

We will study this function in more detail next week, but for today we note one beautiful property of φ:

Theorem 10.1. *If $n \in \mathbb{N}$ then*

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

Proof. This proof is pretty intuitive. We want to show that a sum of a bunch of numbers is equal to n. A good way to show something like this is to establish a bijection between the numbers $\{1, \ldots, n\}$ and the things you are trying to count. For us, we are going to do the following: To each integer $1 \leqslant a \leqslant n$, compute (a, n). This is certainly some number d that divides n. Now the following equation is obvious:

$$\sum_{d|n} \#\{a \mid (a, n) = d\} = n$$

Indeed, every number from 1 to n appears in exactly one of the sets. On the other hand, it's easy to see that

$$\#\{a \mid (a,n) = d\} = \varphi(n/d).$$

(Indeed, if $(a,n) = d$ then $(a/d, n/d) = 1$. Conversely, if $(b, n/d) = 1$ then $(bd, n) = d$.) Therefore:

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

Now whether we sum $\varphi(d)$ or $\varphi(n/d)$ we should get the same thing. □

In class we worked out exactly what this proof says for $n = 12$. This is one of those proofs for which working through it with an example in mind is very helpful.

10.2 $\mathbb{F}_p$ and its groups under $+$ and $\times$

We are now going to explore the structure of the field $\mathbb{F}_p$. Notice that by definition, fields have two groups floating around. First there is the additive group, $(F, +)$. For $\mathbb{F}_p$ this group is fairly simple to describe. It is a *cyclic group* of order p.

Definition 10.2. *Let G be a group. We say that an element $g \in G$ generates G if the set of powers of g and g^{-1} is equal to all of G. If such a g exists, we say that G is cyclic and we write $G = \langle g \rangle$.*

Definition 10.3. *If $g \in G$, we say that the order of g is the smallest positive integer n such that $g^n = 1$.*

It is clear that a group is cyclic if and only if it has an order with order equal to the number of elements in G. In class you will work out several examples to and determine the number of elements of each order under the operations of $+$ and $*$

Some orders of elements modulo 7

n	0	1	2	3	4	5	6
order under $+$	1	7	7	7	7	7	7

,

n	1	2	3	4	5	6
order under $\times$	1	3	6	3	6	2

Some orders of elements modulo 11

n	0	1	2	3	4	5	6	7	8	9	10
order under $+$	1	11	11	11	11	11	11	11	11	11	11

n	1	2	3	4	5	6	7	8	9	10
order under $\times$	1	10	5	5	5	10	10	10	5	2

Some orders of elements modulo 13

n	0	1	2	3	4	5	6	7	8	9	10	11	12
order under $+$	1	13	13	13	13	13	13	13	13	13	13	13	13

n	1	2	3	4	5	6	7	8	9	10	11	12
order under $\times$	1	12	3	6	4	12	12	4	3	6	12	2

Hopefully the pattern in the table for the operation $+$ is clear. As an exercise, prove that if p is prime, then the order (under plus) of an element $1 \leqslant a \leqslant p-1$ is precisely equal to p. However, under $\times$ the situation is clearly a bit more subtle. In these examples, it's true that there is an element of order $p-1$ in each case.

10.3 $\mathbb{F}_p^\times$ is cyclic!

We denote the group of nonzero elements of $\mathbb{F}_p$ by $\mathbb{F}_p^\times$. We now state the rather surprising fact:

Theorem 10.4. *$\mathbb{F}_p^\times$ is a cyclic group.*

Definition 10.5. *We call an element $x \in \mathbb{F}_p^\times$ a primitive root if x is a generator for $\mathbb{F}_p^\times$. In other words, if $x, x^2, x^3, \ldots, x^{p-1}$ are all distinct numbers modulo p. We may also say that x is a primitive root modulo p.*

It is easy to see that an element x is primitive if and only if its order is equal to $p-1$.

For example, if $p = 7$ then we could try a few numbers:

$$\langle 1 \rangle = \{1, 1^2, 1^3, \ldots, \} = \{1\}$$

$$\langle 2 \rangle = \{2, 2^2, 2^3\} = \{1, 2, 4\}$$

$$\langle 3 \rangle = \{3, 3^2, 3^3, 3^4, 3^5, 3^6\} = \{1, 2, 3, 4, 5, 6\}$$

So 3 is a generator for the multiplicative group $\mathbb{F}_p^\times$ when $p = 7$. However, the numbers 1 and 2 failed to generate everything.

As the example above indicates, 3 is a primitive root modulo 7. As another example, if $p = 23$ then 5 is the smallest positive integer which is a primitive root. As we'll see, the theory of these roots can be quite mysterious. The following lemma will be useful:

Lemma 10.6. *Let G be a group and $g \in G$. If for integers m, n, we have $g^m = 1$ and $g^n = 1$ then $g^{\gcd(m,n)} = 1$.*

Proof. Note that since g^{-1} exists, it makes sense to talk about both positive at negative powers of g. Thus since $\gcd(m, n)$ can be written as $mx + ny = \gcd(m, n)$, we see that

$$a^{\gcd(m,n)} = (a^m)^x \cdot (a^n)^y = 1.$$

□

Lemma 10.7. *Let G be a finite group, of cardinality N. If $g \in G$ then the order of g divides N.*

Proof. Let m denote the order of g. Suppose that m does not divide N. Then $\gcd(m, N) < m$. But then by the previous proposition $g^{\gcd(m,N)} = 1$. This is a contradiction since by assumption, we assumed that m was the smallest positive integer k such that $g^k = 1$. □

Let $N(d)$ denote the number of elements of order d in $\mathbb{F}_p^\times$. Let's go back to the table above with $p = 13$. We see that we have

$$N(12) = 4,\ N(6) = 2,\ N(4) = 2,\ N(3) = 2,\ N(2) = 1.$$

These values are exactly the same as the values of the Euler φ function!

Theorem 10.8. *Let if $d \in \mathbb{N}$ then let $N(d)$ denote the number of elements of order d in $\mathbb{F}_p^\times$. Then $N(d) = \varphi(d)$.*

Proof. We proceed in two steps. First we will show that $N(d) \leqslant \varphi(d)$.

Let's do it! So given d, if $N(d) = 0$ we are done, since $0 \leqslant \varphi(d)$. So now suppose that $N(d) > 0$. This means that there is some element $a \in \mathbb{F}_p^\times$ of order d. In other words, the elements

$$\{a, a^2, a^3, \ldots, a^d\}$$

are all distinct modulo p. Now these guys are all roots of the equation $x^d - 1$. (Why?) So now if x is some element of order d then it must be a solution to $x^d - 1$, and hence it is one of the elements a^k. However, an element of the form a^k has order d if and only if $(k, d) = 1$. (Why?)

Thus we have shown that the elements of order d are precisely those elements a^k with $(k, d) = 1$. In other words, there are $\varphi(d)$ of them. (Remark: We have shown that if $N(d) > 0$ then $N(d) = \varphi(d)$. This is nice, but unfortunately doesn't help with the proof since we use different means from here onwards).

We are now ready for the second step. Showing that $N(d) = \varphi(d)$. We will use the fact that $N(d) \leqslant \varphi(d)$. Notice that every element in $\mathbb{F}_p^\times$ has some order. And there are $p - 1$ elements. And the order of elements must divide $p - 1$. Thus

$$p - 1 = \sum_{d|p-1} N(d)$$

And by Step 1, we have

$$p - 1 \leqslant \sum_{d|p-1} \varphi(d)$$

But now by Theorem 10.1, we know that the right hand side is $p - 1$. But this must mean that the inequality $\leqslant$ is actually an equality. So this means that $N(d) = \varphi(d)$ for all d. □

Proof of Theorem 10.4. Since $N(p - 1) = \varphi(p - 1) > 0$ we have that there is always an element of order $p - 1$. Thus $\mathbb{F}_p^\times$ is cyclic. □

10.4 Taking nth roots in $\mathbb{F}_p^\times$

Take an odd prime p and g a fixed primitive root mod$p$. Then for any $B \in \mathbb{F}_p^\times$ we define the *index* (old-fashioned word) or *discrete logarithm* (current jargon) of B, written $\operatorname{ind} B$ or $\log_p B$, as the integer $b \in \{0, 1, \ldots, p-2\}$ such that $B = g^b$ in $\mathbb{F}_p$. Clearly the function $\log_p$ depends not only on p but also on the choice of the primitive root g.

Proposition 10.9. *Given $n \in \mathbb{N}$ and $B \in \mathbb{F}_p^\times$, the equation $X^n = B$ in $\mathbb{F}_p^\times$ has a solution $X \in \mathbb{F}_p^\times$ iff $\gcd(n, p-1) \mid \log_p B$.*

When $\gcd(n, p-1) \mid \log_p B$ then the number of distinct solutions X of $X^n = B$ in $\mathbb{F}_p^\times$ is $\gcd(n, p-1)$.

Proof. Write $B = g^b$, $X = g^x$, so that $g^{nx} = g^b$, giving $nx \equiv b \bmod p-1$. Hence the number of solutions is $\gcd(n, p-1)$. □

For large primes p, the problem of finding the discrete logarithm $\log_p B$ of B appears to be an intractable problem, called the *Discrete Logarithm Problem*. Many techniques in Cryptography depend on this supposed fact. See e.g.,

http://en.wikipedia.org/wiki/Discrete_logarithm

Main Points from Lecture 10:

- Definition and properties of $\varphi(n)$
- The multiplicative group of $\mathbb{F}_p$ is cyclic.
- Definition of primitive roots

11 Multiplicative Functions

Euler's φ function is very useful, as we have seen. A large part of the reason why is because it is a multiplicative function. Before going on for a more detailed study of primitive roots, we study multiplicative functions in general.

11.1 Arithmetic functions - more about φ

Arithmetic functions are functions $f : \mathbb{N} \to \mathbb{N}$ or $\mathbb{Z}$ or maybe $\mathbb{C}$, usually having some arithmetic significance. An important subclass of such functions are the multiplicative functions: such an f is *multiplicative* if

$$f(nn') = f(n)f(n')$$

for all $n, n' \in \mathbb{N}$ with n and n' coprime ($\gcd(n, n') = 1$).

Proposition 11.1. *If f is multiplicative and $n_1, \ldots, n_k$ are pairwise coprime ($\gcd(n_i, n_j) = 1$ for all $i \neq j$) then*

$$f(n_1 n_2 \ldots n_k) = f(n_1)f(n_2)\ldots f(n_k).$$

This is readily proved by induction.

Corollary 11.2. *If n factorises into distinct prime powers as $n = p_1^{e_1} \dots p_k^{e_k}$ then*

$$f(n) = f(p_1^{e_1}) \dots f(p_k^{e_k}).$$

So multiplicative functions are completely determined by their values on prime powers. Some examples of multiplicative functions are

- The identity function: $f(n) = n$;
- The constant function $f(n) = 1$;
- The '1-detecting' function $\Delta(n)$, equal to 1 at $n = 1$ and 0 elsewhere – obviously multiplicative;
- $\tau(n) = \sum_{d|n} 1$, the number of divisors of n;
- $\sigma(n) = \sum_{d|n} d$, the sum of the divisors of n.

Proposition 11.3. *The functions $\tau(n)$ and $\sigma(n)$ are both multiplicative.*

Example 11.4. *Let's check that $\sigma(36) = \sigma(9 \cdot 4) = \sigma(9)\sigma(4)$. The divisors of 36 are*

$$1, 2, 3, 4, 6, 9, 12, 18, 36$$

their sum is 91. On the other hand the divisors of 9 and 4 are respectively

$$1, 3, 9, \text{ and } 1, 2, 4$$

Hence $\sigma(9) = 13$ and $\sigma(4) = 7$. Luckily $13 \cdot 7 = 91$.

A curious thing happened on the way to this result. Notice that 36 had nine divisors and that 9 and 4 each had three apiece. If we were gamblers, we might wager that this is no coincidence. We might wager that any divisor of 36 is a product of a divisor of 9 and a divisor of 4, and that this could be done in a unique way. Stop here and think through why this should be true.

Lemma 11.5. *If a and b are relatively prime then any factor of ab can be written as a product of a factor of a times a factor of b in a unique way. Hence the number of divisors of ab is equal to the number of divisors of a times the number of divisors of b. In terms of the function τ we have proven that $\tau(n)$ is a multiplicative function.*

Notice that the Lemma is not true is if a and b fail to be relatively prime. For instance if $a = 4, b = 12$ then the factor $d = 4$ of 48 can be written in many ways as a product of factors of a and b: $4 = 4 \cdot 1 = 2 \cdot 2 = 1 \cdot 4$. The point is that if a and b are relatively prime, there's only one way to do this!

The proof of this lemma is pretty straightforward and is left as an exercise - think about the prime factors that divide a and b and those that divide ab.

To show that σ is multiplicative, we will prove a stronger result that puts σ and τ in a broader context. This is the operation "hat".

Definition 11.6. *Given an arithmetic function f, define its 'sum over divisors' function $\widehat{f}(n) = \sum_{d|n} f(d)$. This is sometimes also called the summatory function of f.*

Notice that if f is a function, then $\widehat{f}$ is another function, and one that depends on f.

For example $\widehat{f}(12) = f(1) + f(2) + f(3) + f(4) + f(6) + f(12)$, which clearly depends on f.

For instance if $f(n) = n$ then

$$\widehat{f}(n) = \sum_{d|n} f(d) = \sum_{d|n} d = \text{ the sum of all divisors of } n = \sigma(n).$$

We could write this as $\widehat{f} = \sigma$. Or since $f(n) = n$, we could write $\widehat{n} = \sigma$. During class, someone asked why we didn't write $\widehat{d} = \sigma$. This is a good question, and one to think about. The answer is just that the functions $f(d) = d$ and the function $f(n) = n$ are the same function. In some circles they have even a third name - the identity function. It's good to be able to keep these things straight.

For further practice, check that $\widehat{1} = \tau$. Remember, the function $f(n) = 1$ is NOT the identity function, it's the constant function that sends everything to 1.

The following proposition shows the **important** fact that if f is a multiplicative function, then so is $\widehat{f}$. I have modified the proof a bit from the one I presented in class to make it a bit easier to follow.

Proposition 11.7. *Let $F(n) = \widehat{f}(n)$. If f is multiplicative, then F is also multiplicative.*

Proof. The relationship between F and f is that $F(n)$ adds up the values of $f(d)$ on all divisors d of n. Now suppose that $n = ab$ with a and b coprime. We want to show that $F(ab) = F(a)F(b)$, given that $f(ab) = f(a)f(b)$. Notice:

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

$$F(a)F(b) = (\sum_{d|a} f(d))(\sum_{e|b} f(e)) \tag{3}$$

We want to show these two are equal. But by the Lemma, we know that every divisor d of ab can be written uniquely as a product of divisors of a and b. Hence

$$F(ab) = \sum_{\substack{d|a \\ e|b}} f(de) = \sum_{\substack{d|a \\ e|b}} f(d)f(e) \tag{4}$$

where the last equality holds since f is multiplicative. Now it is clear that this is equal to $F(a)F(b)$ because each term of the right hand side of (4) is equal to a term of (3) and vice versa. □

Hence we have shown that $\widehat{f}$ is multiplicative whenever f is. Thus we know that σ and τ are multiplicative since they are the hats of the (obviously) multiplicative functions $f(n) = n$ and $f(n) = 1$.

Remark 11.8. *It's fun sometimes to see what properties multiplicative functions have to have by default. For instance, notice that from the definition we see that $f(1) = f(1 \cdot 1) = f(1)f(1)$. So $(f(1))^2 = f(1)$. This only has two possible solutions in $\mathbb{Z}$ so $f(1)$ is either 1 or 0. If $f(1) = 0$ then we can prove that $f(n) = f(n \cdot 1) = f(n)f(1) = 0$ for all n. So in other words, we have shown that if f is not the zero function, then $f(1) = 1$.*

To close our tour of multiplicative functions, let's study our friend φ. As a warmup,let's compute!

Proposition 11.9. *If p is prime and k is a positive integer then $\varphi(p^k) = p^k - p^{k-1} = p^k(1 - \frac{1}{p})$. In particular $\varphi(p) = p - 1$.*

Proof. There are p^k numbers between 1 and p^k. Of these, the ones that are relatively prime to p^k are the ones who are NOT divisible by p. There are p^{k-1} multiples of p in this range, hence $\varphi(p^k) = p^k - p^{k-1}$. □

Theorem 11.10. *Euler's φ function is multiplicative.*

Notice that this theorem shows that if $n = p_1^{e_1} \cdots p_n^{e_n}$ then $\varphi(n) = \varphi(p_1^{e_1}) \cdots \varphi(p_n^{e_k})$. Each factor is easy to compute by the Proposition above. In fact, since $\varphi(p^e) = p^e(1 - \frac{1}{p})$ we see that

$$\varphi(n) = p_1^{e_1}\left(1 - \frac{1}{p_1}\right) \cdots p_k^{e_k}\left(1 - \frac{1}{p_k}\right) = n \prod_{p|n}\left(1 - \frac{1}{p}\right)$$

which proves:

Corollary 11.11. *If n is an integer, then*

$$\varphi(n) = n \prod_{p|n}\left(1 - \frac{1}{p}\right).$$

Remark 11.12. *Notice that this is a quite a nice formula, but in practice it's probably easiest to just remember the formula $\varphi(p^k) = p^k - p^{k-1}$ and use this to compute φ.*

Example 11.13.

$$\varphi(300) = \varphi(3)\varphi(4)\varphi(25) = (3 - 1)(4 - 2)(25 - 5) = 2 \cdot 2 \cdot 20.$$

Proof of Theorem. Take n and n' coprime, and let

$$\{i : 1 \leq i \leq n, \gcd(i, n) = 1\} = \{a_1 < a_2 < \cdots < a_{\varphi(n)}\},$$

the *reduced residue classes* mod$n$. Similarly, let

$$\{j : 1 \leq j \leq n', \gcd(j, n') = 1\} = \{a'_1 < a'_2 < \cdots < a'_{\varphi(n')}\}.$$

The idea is now that numbers that are relatively prime to nn' are gotten by combining pairs of (a_i, a'_j) using the Chinese Remainder Theorem in a unique way. Hence the number of relatively prime integers to nn' is equal to the product of $\varphi(n)\varphi(n')$.

If $x \in \{1, 2, \ldots, nn'\}$ and $\gcd(x, nn') = 1$ then certainly $\gcd(x, n) = \gcd(x, n') = 1$, so that

$$x \equiv a_i \bmod n \qquad\qquad x \equiv a'_j \bmod n' \tag{5}$$

for some pair a_i, a'_j. Conversely, given such a pair a_i, a'_j we can solve (5) using the CRT to get a solution $x \in \{1, 2, \ldots, nn'\}$ with $\gcd(x, nn') = 1$. Thus we have a bijection between such x and such pairs a_i, a'_j. Hence

$$\#\{\text{such } x\} = \varphi(nn') = \#\{a_i, a'_j\} = \varphi(n)\varphi(n').$$

□

12 Euler's Theorem and More about Primitive Roots

We begin with a definition:

Definition 12.1. *An element $1 \leqslant a \leqslant n$ such that a has an inverse modulo n is called a* ***unit modulo*** *n.*

Proposition 12.2. *The set of units modulo n form a group under multiplication. (With multiplication being done "modulo n") We denote this group by $\mathbb{Z}_n^\times$.*[6]

Proof. We have to show that the operation of multiplication is well-defined on the set of units. I.e. that the product of two units is a unit.[7] This follows since the inverse of ab is the product of the inverses of a and b. We also should show that every element has an inverse (by definition), that 1 is in this set (obvious) and that the multiplication is associative (again obvious). In short there wasn't much to see in this proof. Better ask for our money back. □

How many elements are units modulo n? Well an element a has an inverse if and only if $(a, n) = 1$. There are precisely $\varphi(n)$ of these. Thus the **Group of units modulo** n is a group with $\varphi(n)$ elements. Recalling that if you have a finite group G of order $|G|$ then every element g in the group satisfies $g^{|G|} = e$ where e is the identity, we obtain

Theorem 12.3 (Euler's Theorem)**.** *If n is a positive integer then $a^{\varphi(n)} \equiv 1 \bmod n$ for every a with $(a, n) = 1$.*

[6]There are many notations for this group. Some people write $\mathbb{Z}/n\mathbb{Z}^\times$. Others use $U(\mathbb{Z}_n)$ or $U(\mathbb{Z}/n\mathbb{Z})$. What's important is to remember that this group is not all of the numbers from 1 to n, but only those numbers that are coprime to n.

[7]Some of you might call this "showing that the set is closed under multiplication"

Proof. If $(a, n) = 1$ then a is a unit modulo n. Hence it is in the group G of units modulo n. But this means that it satisfies $a^{|G|} = 1$ in this group. Since $|G| = \varphi(n)$ we are done. □

Remark 12.4. *Euler's Theorem is a generalization of Fermat's Little Theorem, since* $\varphi(p) = p - 1$.

An Application: We can use Euler's Theorem to find inverses: Indeed, we know that $a \cdot a^{\varphi(n)-1} = 1 \bmod n$ so that $a^{\varphi(n)-1}$ is the inverse of a modulo n. For instance if $ax \equiv b \bmod n$ has a solution that it must be

$$x = a^{\varphi(n)-1} b \bmod n.$$

You might wonder: Modulo p there was always a primitive root - i.e. an element of order $p - 1$ (the biggest possible order). Must there always be an element of order $\varphi(n)$ modulo n? The answer to this question is NO, but in some cases the answer is yes.

Example 12.5. *Consider the integers modulo* 4. *Let's denote the set of congruence classes by* $\mathbb{Z}/4\mathbb{Z} := \{0, 1, 2, 3\}$. *The units are the numbers relatively prime to* 4, *which are*

$$\mathbb{Z}/4\mathbb{Z}^\times = \{1, 3\}$$

This is a group of order two, and the element 3 *indeed has order* 2*: (*$3^1 \neq 1$, *but* $3^2 = 1$.*)*

On the other hand

$$\mathbb{Z}/8\mathbb{Z}^\times = \{1, 3, 5, 7\}$$

is a group of order 4. *However, the elements* $3, 5, 7$ *all have order* 2. *Thus there is no element of order* 4. *(However, they still all satisfy* $g^4 = 1$.*)*

Definition 12.6. *We say that* a *is a primitive root modulo* n *if* $(a, n) = 1$ *and the following* $\varphi(n)$ *numbers*

$$\{a, a^2, a^3, \ldots, a^{\varphi(n)}\}$$

are distinct modulo n. *In other words,* a *has (maximal) order equal to* $\varphi(n)$. *In this case we say that* n *has a primitive root.*

As the above example shows, 3 is a primitive root modulo 4 but that 8 has no primitive roots. The following theorem states exactly which integers n have primitive roots.

Theorem 12.7. *A positive integer* n *has a primitive root if and only if* n *is one of the following numbers*

$$2,\ 4,\ p^k,\ 2 \cdot p^k$$

where p *is an odd prime and* k *is a positive integer.*

We have seem already that 4 has a primitive root, and have also shown that p has a primitive root for all p. What remains is to show that that in the remaining cases p^k $(k \geqslant 1)$ and $2p^k$ there is indeed a primitive root, and that also all other numbers lack a primitive root. The proof of this theorem is involved, but depending on class interest, we may come back and prove this later. The statement of this theorem is important to know, though. As is the following:

Theorem 12.8. *If n has a primitive root then it has $\varphi(\varphi(n))$ primitive roots.*

Proof. The double φ is NOT a typo, and although this looks a bit intimidating, the proof is actually really calming. Indeed, it's just three baby steps:

First, the group of units modulo n is a group G with $|G| = \varphi(n)$.

Second, if n has a primitive root, then this means that G is cyclic.

Third, if G is a cyclic group of order m then G has $\varphi(m)$ many generators.

The result now follows. □

Lemma 12.9. *If G is a cyclic group of order m then G has $\varphi(m)$ many generators.*

We won't present the proof in class, but because it was on the homework, I'm including a different solution here:

Proof. This was essentially one of your homework problems, but here's the idea. If G is cyclic, then that means that $G =< g >$ is generated by some element g. Hence every element of the group is a power of g:

$$G = \{g, g^2, \ldots, g^m\}$$

Now we just have to figure out how many elements have order m. Well g^k has order m if and only if m is the smallest positive integer such that $(g^k)^m = 1$. (The word "smallest" is the important word here, we already know that $(g^k)^m = 1$, since EVERY element in a group satisfying $x^m = 1$.)

Now suppose that n is the smallest power such that $(g^k)^n = 1$. Then $(g^k)^n = g^{kn}$ and $g^{kn} = 1$ if and only if $kn \equiv 0 \bmod m$. Hence we have that kn is a multiple of m, and n is the smallest positive integer n with this property. Now if $(k, m) = d$ then

$$(x^k)^{m/d} = (x^m)^{k/d} = (x^m)^{\text{an integer}} = 1$$

so certainly $n < m/d$. Hence if $n = m$ then $d = 1$ and m and k are relatively prime. Conversely, it's easy to check that if k and m are relatively prime that kn is a multiple of m if and only if n is a multiple of m. □

This proof had a lot of steps, and the technicalities obscure the fact that this result is simple, nice, and really intuitive. If this proof doesn't feel like a part of your repertoire, then try a few examples:

(If G is cyclic of order 20, then $G = \{g, g^2, \ldots, g^{20}\}$. Work out the order of the elements g^10, g^2, g^3, g^5, g^7 and look for patterns. Try to write your own proof of the above, etc.)

13 Perfect Numbers and Mersenne Primes

13.1 Perfect numbers

A positive integer n is called *perfect* if it is the sum of its proper (i.e., excluding n itself) divisors. Thus $\sigma(n) = 2n$ for n perfect.

Theorem 13.1. *An even number n is perfect iff it of the form $n = 2^{p-1}(2^p - 1)$ for some prime p with the property that $2^p - 1$ is also prime.*

Proof. First suppose that $n = 2^{p-1}(2^p - 1)$ and $2^p - 1$ is prime. Then since σ is multiplicative, we have that

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

Check that $\sigma(2^k) = (2^{k+1} - 1)$ and $\sigma(q) = q + 1$ for q prime. Thus

$$\sigma(n) = (2^p - 1)(2^p) = 2n.$$

Conversely, suppose that n is an even perfect number. Then we can write $n = 2^k \cdot t$ where t is an odd number. Then perfection implies that

$$2^{k+1}t = 2n = \sigma(n) = (2^{k+1} - 1)\sigma(t) \tag{6}$$

This implies that 2^{k+1} divides $\sigma(t)$ (since the other factor on the right is odd). Thus we can write $\sigma(t) = 2^{k+1}s$. Our goal is to show that $k + 1 = p$ is prime and $s = 1$. Now canceling we see that

$$t = (2^{k+1} - 1)s$$

If $s > 1$ then t clearly has $1, s, t$ as factors. Thus

$$\sigma(t) \geqslant 1 + t + s = 1 + (2^{k+1}s - s) + s = 1 + (2^{k+1}s)$$

but this is a contradiction, because we assume that $\sigma(t) = 2^{k+1}s$.

Hence $s = 1$ and we have that $\sigma(t) = 2^{k+1}$, and Equation (6) says that

$$t = (2^{k+1} - 1).$$

This information about t and $\sigma(t)$ implies that t must in fact be prime, as required. □

Prime numbers of the form $2^p - 1$ are called *Mersenne primes*. (Unsolved problem: are there infinitely many such primes?) It is easy to show as you did on the first homework, that if $2^p - 1$ is prime, then p must be prime. The converse is not true. Later, when we discuss primality testing we will see that there is a relatively efficient algorithm to check whether a number of the form $2^n - 1$ is prime. (The Lucas-Lehmar test) A good source of information on Mersenne numbers is

`http://primes.utm.edu/mersenne/index.html`

It is an unsolved problem as to whether there are any odd perfect numbers. See e.g., `http://en.wikipedia.org/wiki/Perfect_number` for lots on this problem.

14 Möbius Inversion and Some Applications

No study of multiplicative functions would be complete without the classic formula of Möbius Inversion. The idea is simple: We want to invert the operation of going from $f \mapsto \widehat{f}$. Let's do a few examples - following the notation of Rosen, we will let $F(n) = \widehat{f}(n)$.

$$\begin{aligned} F(1) &= f(1) \\ F(2) &= f(1) + f(2) \\ F(3) &= f(1) + f(3) \\ F(4) &= f(1) + f(2) + f(4) \\ F(5) &= f(1) + f(5) \\ F(6) &= f(1) + f(2) + f(3) + f(6) \end{aligned}$$

If we solve these equations for $f(n)$ in terms of $F(n)$ we see

$$\begin{aligned} f(1) &= F(1) \\ f(2) &= F(2) - F(1) \\ f(3) &= F(3) - F(1) \\ f(4) &= F(4) - F(2) \\ f(5) &= F(5) - F(1) \\ f(6) &= F(6) - F(2) - F(3) + F(1). \end{aligned}$$

Experimentally it seems that $f(n) = \sum_{d|n} \mu F(d)$ where μ seems to be either $0, 1$ or -1. We will prove that there exists a multiplicative function $\mu(n)$ such that $f(n) = \sum_{d|n} \mu(n/d)F(d)$. This allows us to invert the process of passing to a summatory function $f \mapsto \widehat{f}$.

The *Möbius function* $\mu(n)$ is defined as

$$\mu(n) = \begin{cases} 0 \text{ if } p^2 \mid n \text{ for some prime } p; \\ (-1)^k \text{ if } n = p_1 p_2 \dots p_k \text{ for distinct primes } p_i. \end{cases}$$

In particular, $\mu(1) = 1$ and $\mu(p) = -1$ for a prime p. It is immediate from the definition that μ is multiplicative.

Proposition 14.1. *The summatory function of μ is the 1-detecting function $\Delta(n)$:*

$$\widehat{\mu}(n) = \Delta(n)$$

Proof. We need to check that $\sum_{k|n} \mu(k) = \Delta(n)$. Recall that $\Delta(n)$ is one if $n = 1$ and zero otherwise. Since μ is multiplicative, it suffices to compute $\sum_{k|n} \mu(k)$ when $n = p^e$ is a power of a prime. If $e > 0$ then

$$\sum_{k|p^e} \mu(k) = \mu(1) + \mu(p) + \cdots \mu(p^e) = 1 - 1 = 0.$$

If $e = 0$ then of course $\widehat{mu}(1) = mu(1) = 1$. Hence $\widehat{mu}(n) = 0$ unless $n = 1$, and we are done. □

Integers with $\mu(n) = \pm 1$ are called *squarefree*.

The Möbius function arises in many kinds of *inversion* formulae. The fundamental one is the following.

Proposition 14.2 (Möbius inversion)**.** *Let $f(n)$ be a multiplicative function, and let $F(n) = \widehat{f}(n) = \sum_{d|n} f(d) \quad (n \in \mathbb{N})$. Then for all $n \in \mathbb{N}$ we have $f(n) = \sum_{d|n} \mu(n/d)F(d)$.*

Proof. We will simplify $\sum_{d|n} \mu(n/d)F(d) = \sum_{d|n} \mu(n/d) \sum_{k|d} f(k) \quad (n \in \mathbb{N})$ by interchanging the order of summation to make $\sum_{k|n}$ the outer sum. First note that we can swap d and n/d in the sum:

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

$$= \sum_{e|n} (\sum_{k|(n/e)} \mu(e) \cdot f(k)).$$

Notice that the pairs of integers (e, k) such that $e \mid n$ and $k \mid n/e$ is the same as the set of pairs is the same as those with $k \mid n$ and $e \mid (n/k)$. So the

$$\sum_{e|n} (\sum_{k|(n/e)} \mu(e) \cdot f(k)) = \sum_{k|n} (\sum_{e|(n/k)} \mu(e) f(k))$$

$$= \sum_{k|n} f(k) \left(\sum_{e|(n/k)} \mu(e) \right).$$

The inner bracket is $\widehat{\mu}(n/k)$ which is nonzero if and only if $(n/k) = 1$. In this case, $\widehat{\mu}(1) = 1$. Hence we have that the sum above reduces to

$$\sum_{k|n} f(k) \left(\sum_{e|(n/k)} \mu(e) \right) = f(n) \cdot 1 = f(n).$$

□

14.1 Some Examples of Using Möbius Inversion

There are two main uses of Möbius Inversion. The first is that we can just apply the formula to immediately obtain identities which might be difficult to obtain directly. For instance, recall that $\sigma(n) = \widehat{n}$ (Indeed, $\sigma(n) = \sum_{d|n} d.$) so we have that

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

Similarly, since $\tau(n) = \widehat{1}(n)$, (Indeed, $\tau(n) = \sum_{d|n} 1$) we have

$$1 = \sum_{d|n} \mu(n/d)\tau(d).$$

Finally, since $n = \widehat{\varphi}(n)$ we have

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

For example, $\varphi(12) = 4$ and

$$\sum_{d|12} \mu(12/d)d = \mu(12/1)1 + \mu(12/2)2 + \mu(12/3)3 + \mu(12/4)4 + \mu(12/6)6 + \mu(12/12)12$$

$$= 0 + 2 + 0 - 4 - 6 + 12 = 4.$$

The second main use of Möbius inversion is the following

Theorem 14.3. *If $\widehat{f}(n)$ is a multiplicative function, then so is $f(n)$.*

Proof. Let $F(n) = \widehat{f}(n)$. By Möbius Inversion, we have that

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

If a and b are coprime, then

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

We can write every factor of ab as a product of a factor of a and a factor of b

$$f(ab) = \sum_{\substack{d_1|a \\ d_2|b}} \mu(\frac{ab}{d_1 d_2})F(ab)$$

Since F and μ are multiplicative:

$$f(ab) = \sum_{\substack{d_1|a \\ d_2|b}} \mu(a/d_1)F(a)\mu(b/d_2)F(b)$$

$$f(ab) = \sum_{d_1|a} \mu(a/d_1)F(a) \sum_{d_2|b} \mu(b/d_2)F(b)$$

$$f(ab) = f(a) \cdot f(b)$$

again by Möbius Inversion. Thus f is multiplicative. □

Remark 14.4. *This gives another proof that φ is multiplicative, say from the fact that $\widehat{\varphi} = n$.*

15 Quadratic Residues

15.1 Quadratic residues and nonresidues

Take p an odd prime, and $r \in \mathbb{F}_p^\times$. If the equation $x^2 = r$ has a solution $x \in \mathbb{F}_p^\times$ then r is called a *quadratic residue* mod$p$. If there is no such solution x, then r is called a *quadratic nonresidue* mod$p$.

Proposition 15.1. *Take p an odd prime, and g a primitive root* mod$p$. *Then the quadratic residues* mod$p$ *are the even powers of g, while the quadratic nonresidues* mod$p$ *are the odd powers of g. (So there are $\frac{p-1}{2}$ of each.)*

In particular, $\left(\frac{g^k}{p}\right) = (-1)^k$.

Proof. Suppose $r \in \mathbb{F}_p^\times$, with $r = g^k$ say. If k is even then $r = (g^{k/2})^2$, so that r is a quadratic residue mod$p$. Conversely, if $x = g^\ell$, $x^2 = r$, then $g^{2\ell - k} = 1$, so that $2\ell - k$ is a multiple of $p - 1$, which is even. So k is even. □

15.2 The Legendre symbol

Let p be an odd prime, and $r \in \mathbb{F}_p^\times$. Then the *Legendre symbol* is defined as

$$\left(\frac{r}{p}\right) = \begin{cases} 1 \text{ if } r \text{ is a quadratic residue;} \\ -1 \text{ if } r \text{ is a quadratic nonresidue.} \end{cases}$$

Note that, on putting $r = g^k$ for a primitive root g we see that

$$\left(\frac{g^k}{p}\right) = (-1)^k = \begin{cases} 1 \text{ if } k \text{ is even;} \\ -1 \text{ if } k \text{ is odd.} \end{cases}$$

Next, recall Fermat's Theorem: that $r^{p-1} = 1$ for all $r \in \mathbb{F}_p^\times$. This is simply a consequence of $\mathbb{F}_p^\times$ being a group of size (order) $p-1$. (We know that $g^{\#G} = 1$ for each g in a finite group G.)

Proposition 15.2 (Euler's Criterion)**.** *For p an odd prime and $r \in \mathbb{F}_p^\times$ we have in $\mathbb{F}_p^\times$ that*

$$\left(\frac{r}{p}\right) = r^{\frac{p-1}{2}}. \tag{7}$$

Proof. If $r = g^k$ then for k even

$$r^{\frac{p-1}{2}} = g^{k\frac{p-1}{2}} = (g^{p-1})^{k/2} = 1^{k/2} = 1,$$

while if k is odd, $k\frac{p-1}{2}$ is not a multiple of $p-1$, so $r^{\frac{p-1}{2}} \neq 1$. However, $r^{p-1} = 1$ by Fermat, so $r^{\frac{p-1}{2}} = \pm 1$ and hence $r^{\frac{p-1}{2}} = -1$. So, by Proposition 15.1, we have (7), as required. □

Theorem 15.3. *In particular ($r = -1$), for p an odd prime, we have*

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} = \begin{cases} 1, & \text{if } p \equiv 1 \bmod 4 \\ -1, & \text{if } p \equiv -1 \bmod 4. \end{cases}$$

Lemma 15.4. *Let p be an odd prime, and a, b be integers not divisible by p. We have*

1. *$a \equiv b \bmod p$ implies that $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$;*
2. *$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$;*
3. *$\left(\frac{a^2}{p}\right) = 1$, $\left(\frac{a^2 b}{p}\right) = \left(\frac{b}{p}\right)$.*

Proof. Let g be a primitive root mod$p$. Then $\left(\frac{g^k}{p}\right) = (-1)^k$, from which the results follow easily. □

Example 15.5. *Determine whether or not 90 is a square modulo 11.*

Solution: We are asked to compute $\left(\frac{90}{11}\right)$ By the Lemma above we see that

$$\left(\frac{90}{11}\right) = \left(\frac{9}{11}\right)\left(\frac{10}{11}\right) = 1 \cdot \left(\frac{-1}{11}\right).$$

(Since 9 is clearly a square). Now by the Theorem, we know that -1 is a quadratic residue if and only if p is congruent to 1 modulo 4. Since $11 \equiv 3 \bmod 4$ we see that -1 is not a quadratic residue. Thus

$$\left(\frac{90}{11}\right) = -1$$

and 90 is not a quadratic residue.

To some extent, Euler's criterion allows us to determine whether or not a is a quadratic residue. However, since it involves taking a large power, it is not clear how effective this is. At the same time, the Lemma above shows that to compute Legendre symbol $\left(\frac{a}{p}\right)$, it suffices to consider only the prime factors q of a and compute $\left(\frac{q}{p}\right)$. Since we can reduce modulo p, we can assume that $q < p$. From this perspective, it is natural to ask whether or not there is some relationship between $\left(\frac{q}{p}\right)$ and $\left(\frac{p}{q}\right)$. The answer is a resounding Yes and is one of the most celebrated results of number theory. In the next section we present this theorem of Quadratic Reciprocity and a sketch of the proof.

16 Quadratic Reciprocity

16.1 Introduction

Recall that the Legendre symbol $\left(\frac{a}{p}\right)$ is defined for an odd prime p and integer a coprime to p as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 \text{ if } a \text{ is a quadratic residue } \bmod p; \\ -1 \text{ otherwise;} \end{cases}$$

Recall too that for a, b coprime to p

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

(easily proved by writing a, b as powers of a primitive root), and that, by Euler's Criterion,

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} = \begin{cases} 1, & \text{if } p \equiv 1 \bmod 4 \\ -1, & \text{if } p \equiv -1 \bmod 4. \end{cases}$$

Theorem 16.1 (Law of Quadratic Reciprocity (Legendre, Gauss))**.** *For distinct odd primes p and q we have*

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

(Thus $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ unless p and q are both $\equiv -1 \bmod 4$, in which case $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$.)
There are now 240 recorded proofs of this (not all different), including six by Gauss – see `http://www.rzuser.uni-heidelberg.de/~hb3/rchrono.html`.
We'll give one of Gauss's proofs, using

Lemma 16.2 (Gauss's Lemma)**.** *For an odd prime p, put $p' = \frac{p-1}{2}$, and let a be an integer coprime to p. Consider the sequence*

$$a, 2a, 3a, ..., p'a,$$

reduced $\bmod p$ *to lie in* $(-\frac{p}{2}, \frac{p}{2})$*. Then* $\left(\frac{a}{p}\right) = (-1)^\nu$*, where ν is the number of negative numbers in this sequence.*

Proof. Now all of $a, 2a, 3a, ..., p'a$ are $\equiv \bmod p$ to one of $\pm 1, \pm 2, \ldots, \pm p'$. Further,

- no two are equal, as $ia \equiv ja \bmod p \Rightarrow i \equiv j \bmod p$;
- none is minus another, as $ia \equiv -ja \bmod p \Rightarrow i + j \equiv 0 \bmod p$.

So they must be $\pm 1, \pm 2, \ldots, \pm p'$, with each of $1, 2, \ldots, p'$ occurring with a *definite sign*. Hence

$$a \cdot 2a \cdot 3a \cdot \ldots \cdot p'a \equiv (\pm 1) \cdot (\pm 2) \cdot \ldots \cdot (\pm p') \bmod p,$$

giving

$$a^{p'}(p')! \equiv (-1)^{\nu}(p')! \bmod p,$$

and so, as $(p')!$ is coprime to p, that

$$a^{p'} \equiv (-1)^{\nu} \bmod p.$$

Finally, using Euler's criterion (Prop. 2.8), we have

$$\left(\frac{a}{p}\right) \equiv a^{p'} \equiv (-1)^{\nu} \bmod p.$$

Hence $\left(\frac{a}{p}\right) = (-1)^{\nu}$. □

We can use Gauss's Lemma to evaluate $\left(\frac{2}{p}\right)$.

Proposition 16.3. *For p an odd prime we have $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.*

(This is equal to 1 when $p \equiv \pm 1 \bmod 8$, and to -1 when $p \equiv \pm 3 \bmod 8$.)

Proof. There are four similar cases, depending on $p \bmod 8$. We give the details for $p \equiv 3 \bmod 8$, $p = 8\ell + 3$ say. Then $p' = 4\ell + 1$, and, taking $a = 2$ in Gauss's Lemma, we see that for the sequence

$$2, 4, 6, \ldots, 4\ell, 4\ell + 2, \ldots, 8\ell + 2$$

that this becomes

$$2, 4, 6, \ldots, 4\ell, -(4\ell + 1), -(4\ell - 1), \ldots, -3, -1$$

when reduced $\bmod p$ into the range $(-\frac{p}{2}, \frac{p}{2})$. This clearly has 2ℓ positive members, and hence $\nu = p' - 2\ell = 2\ell + 1$ negative members. Hence $\left(\frac{2}{p}\right) = (-1)^{2\ell+1} = -1$. □

Doing the other three cases would be a good exercise!

We now use Gauss's Lemma with $a = q$ to prove the Law of Quadratic Reciprocity.

Proof of Theorem 16.1. Take distinct odd primes p and q. For $k = 1, 2, \ldots, p'$ write (one step of the Euclidean algorithm)

$$kq = q_k p + r_k \tag{8}$$

say, where $1 \leq r_k \leq p - 1$ and

$$q_k = \left\lfloor \frac{kq}{p} \right\rfloor. \tag{9}$$

Now, working in $\mathbb{F}_p$ we have

$$\{q, 2q, \ldots, p'q\} = \{r_1, r_2, \ldots, r_{p'}\} = \{a_1, a_2, \ldots, a_t\} \cup \{-b_1, -b_2, \ldots, -b_\nu\},$$

as in Gauss's Lemma. So the a_i's are in $(0, \frac{p}{2})$ and the $-b_i$'s are in $(-\frac{p}{2}, 0)$. (In fact $t = p' - \nu$, but not needed.) Now put

$$a = \sum_{i=1}^{t} a_i, \qquad\qquad b = \sum_{i=1}^{\nu} b_i.$$

So, by the definition of the a_i's and $-b_i$'s we have

$$\sum_{k=1}^{p'} r_k = a - b + \nu p. \tag{10}$$

Now, in the proof of Gauss's Lemma we saw that

$$\{a_1, a_2, \ldots, a_t\} \cup \{b_1, b_2, \ldots, b_\nu\} = \{1, 2, \ldots, p'\},$$

so that

$$\frac{p^2 - 1}{8} = 1 + 2 + \cdots + p' = a + b. \tag{11}$$

and

$$\begin{aligned} \frac{p^2 - 1}{8} q &= \sum_{k=1}^{p'} kq \\ &= p \sum_{k=1}^{p'} q_k + \sum_{k=1}^{p'} r_k \qquad\qquad \text{(using (8))} \\ &= p \sum_{k=1}^{p'} q_k + a - b + \nu p, \qquad\qquad \text{(using (10).)} \end{aligned} \tag{12}$$

Next, on subtracting (12) from (11) we get

$$\frac{p^2 - 1}{8}(q - 1) = p \sum_{k=1}^{p'} q_k - 2b + \nu p.$$

Reducing this modulo 2 we have $0 \equiv \sum_{k=1}^{p'} q_k - \nu \bmod 2$, or $\nu \equiv \sum_{k=1}^{p'} q_k \bmod 2$. Thus Gauss's Lemma gives

$$\left(\frac{q}{p}\right) = (-1)^\nu = (-1)^{\sum_{k=1}^{p'} q_k} = (-1)^{\sum_{k=1}^{p'} \left\lfloor \frac{kq}{p} \right\rfloor},$$

using (9).

Now, reversing the rôles of p and q we immediately get

$$\left(\frac{p}{q}\right) = (-1)^{\sum_{\ell=1}^{q'} \left\lfloor \frac{\ell p}{q} \right\rfloor},$$

where of course $q' = (q-1)/2$, and we've replaced the dummy variable k by ℓ. So

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\left\{\sum_{k=1}^{p'} \left\lfloor \frac{kq}{p} \right\rfloor + \sum_{\ell=1}^{q'} \left\lfloor \frac{\ell p}{q} \right\rfloor\right\}},$$

which equals $(-1)^{p'q'}$, by the following proposition. □

Proposition 16.4. *Let p and q be two coprime odd positive integers. Then*

$$\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{kq}{p} \right\rfloor + \sum_{\ell=1}^{\frac{q-1}{2}} \left\lfloor \frac{\ell p}{q} \right\rfloor = \frac{p-1}{2} \cdot \frac{q-1}{2}.$$

We shall see later that this result will be used in the proof of the Law of Quadratic Reciprocity.

Proof. Consider the rectangle with corners $(0,0)$, $(p/2,0)$, $(0,q/2)$ and $(p/2,q/2)$. (Suggest you draw it, along with its diagonal from $(0,0)$ to $(p/2,q/2)$, and the horizontal axis the k-axis, the vertical axis the ℓ-axis. The diagonal is then the line with equation $\ell = kq/p$.) We count the number of integer lattice points (k,ℓ) strictly inside this rectangle in two different ways. First we note that these points form a rectangle with corners

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

so that there are $\frac{p-1}{2} \cdot \frac{q-1}{2}$ of them in all.

On the other hand, we count separately those below, above and on the diagonal. *Below* the diagonal we have, for $k = 1, \ldots \frac{p-1}{2}$ that $\left\lfloor \frac{kq}{p} \right\rfloor$ is the number of points (k,ℓ) with $1 \le \ell \le \frac{kq}{p}$, i.e., below the diagonal, in the kth column. So the total is $\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{kq}{p} \right\rfloor$.

To count the number of lattice points above the diagonal, we flip the diagram over, reversing the rôles of p and q, and of k and ℓ. Then we get that the number of points *above* the diagonal is $\sum_{\ell=1}^{\frac{q-1}{2}} \left\lfloor \frac{\ell p}{q} \right\rfloor$. It remains to check that there are no lattice points actually *on* the diagonal. For if the integer lattice point (k,ℓ) were on the diagonal $\ell = kq/p$ we would have $\ell p = kq$ so that, as p and q are coprime, $p \mid k$. But $k < p$, so this is impossible. □

16.2 Proofs of Infinitude of Primes

Euclid's proof of the infinitude of primes was remarkable for its simplicity. To review, he argued that if there are only finitely many primes $p_1, \ldots, p_k$ then the number $p_1 \cdots p_k + 1$ is not divisible by any prime, an absurdity. Thus there must be infinitely many.

You can actually get pretty far by just modifying this proof a little bit. For instance, consider the following

Claim 16.5. *If n is congruent to 3 modulo 4 then it has a prime factor congruent to 3 modulo 4.*

The proof is simple: n is odd, so its factors are all odd. Such factors are either 1 or 3 modulo 4. If they were all 1, then n itself would be 1 mod 4, which is isn't. So n has to have a factor congruent to 3 modulo 4.

This claim gives us an easy proof of the following

Theorem 16.6. *There are infinitely many primes of the form $4k + 3$.*

Proof. Suppose there were only finitely many such primes, call them $p_1, \ldots, p_k$. Then consider the number

$$N = 4p_1 \cdots p_k - 1.$$

Then N is clearly congruent to 3 modulo 4, and thus has a prime factor that is 3 mod 4 by the claim. However, at the same time, N is not divisible by any of the p_i. Hence they cannot have been a full list of primes that were 3 mod 4. □

We might hope that we could continue in this fashion to do other cases, but we soon run into difficulties. For instance if we tried to prove that there are infinitely many primes congruent to 1 modulo 4 then this approach would not work. The problem is that the Claim above is not true if we replace 3 with 1. Indeed, the issue is that pairs of factors that are congruent to 3 modulo 4 multiply to give 1 mod 4. For instance $3 \cdot 7 = 21$.

Using quadratic reciprocity, we can improve a this approve by a bit.

Theorem 16.7. *There are infinitely many primes of the form $4k + 1$.*

Proof. Suppose there were only finitely many such primes, call them $p_1, \ldots, p_k$. Then consider the number

$$N = 4(p_1 \cdots p_k)^2 + 1.$$

Then N is clearly congruent to 1 modulo 4. Now let q be a prime factor of N. Then modulo q we see that

$$0 \equiv 4(p_1 \cdots p_k)^2 + 1$$

and thus

$$-1 \equiv (2p_1 \cdots p_k)^2$$

so -1 is a square modulo q. But by QR, we know that -1 is a square if and only if q is congruent to 1 modulo 4. Hence all prime factors of N are congruent to 1 modulo 4. However, at the same time, N is not divisible by any of the p_i. Hence they cannot have been a full list of primes that were 1 mod 4. □

Problem 2 on Workshop 5 deals with showing there are infinitely many primes congruent to 7 modulo 8. For that problem you want to consider a number of the form $(4p_1 \cdots p_k)^2 - 2$. You can also use this approach to show that there are infinitely many primes of the form $8k+3$ and $8k+5$, where you would use numbers of the form $(p_1 \cdots p_k)^2+2$ and $(p_1 \cdots p_k)^2+4$ respectively. (The case $8k + 3$ was on the practice exam).

17 Some Diophantine Equations

17.1 Fermat's method of descent

Equations to be solved in integer variables are called *Diophantine* equations, in honour of Diophantus of Alexandria, who in the 3rd century AD is first recorded as working on them.

Around 1640, Fermat developed a method for showing that certain Diophantine equations had no (integer) solutions. In essence, the method is as follows: assume that the equation *does* have a solution. Pick the 'smallest' (suitably defined) one. Use the assumed solution to construct a smaller solution, contradicting the fact that the one you started with was the smallest. This contradiction proves that there is in fact no solution. The technique is called *Fermat's method of descent.* It is, in fact, a form of strong induction. (Why?)

We illustrate the method with three examples.

17.2 A 2-variable quadratic equation with no nonzero integer solution

Proposition 17.1. *The equation $x^2 = 2y^2$ has no solution in positive integers.*

First proof. Assume there is a solution in positive integers x, y. Find the highest power of 2 dividing x and the highest power of two dividing y. Then we see that x^2 is exactly divisible by an even power of 2, while $2y^2$ is exactly divisible by an odd power of 2, a contradiction. So there is no such solution x, y.

Second proof. We use Fermat Descent. Again assume that there is a solution x, y in positive integers. Define the size of the solution to be $x + y$. Choose a solution of smallest size (not necessarily unique – that doesn't matter). Then, because $2 \mid 2y^2$, $2 \mid x^2$, and so $2 \mid x$. Put $x = 2x_1$, so that $(2x_1)^2 = 2y^2$, or $y^2 = 2x_1^2$. Hence we have another solution y, x_1 of the original equation. But its size is $y+x_1 < y+x$, contradicting the assumption that we started with a solution of smallest size. Hence the assumption that there was a solution must be wrong. □

We next look at a more complicated example. The principle of proving that there is no solution is just the same, however.

17.3 A 4-variable quadratic equation with no nonzero integer solution

Theorem 17.2. *Let p and q be odd primes such that at least one of $\left(\frac{p}{q}\right), \left(\frac{q}{p}\right)$ is -1. Then the equation*

$$x^2 + pqy^2 = pz^2 + qw^2 \tag{13}$$

has no solution in positive integers x, y, z, w.

Proof. In fact we'll prove the slightly stronger assertion that (13) has no solution in nonnegative integers x, y, z, w not all 0.

Suppose that say $\left(\frac{p}{q}\right) = -1$ and there is such a solution (x, y, z, w). Clearly we can assume that x, y, z, w are all ≥ 0. Define the *size* of such a solution by $s(x, y, z, w) = x + y + z + w$. We know that $s(x, y, z, w)$ is a positive integer. Among all such solutions, we choose one that has size $s(x, y, z, w)$ as small as possible.

Considering (13) modulo q, we have that $x^2 \equiv pz^2 \bmod q$. If $z \not\equiv 0 \bmod q$, we would have $(xz^{-1})^2 \equiv p \bmod q$, contradicting $\left(\frac{p}{q}\right) = -1$. Hence $q \mid z$, and so also $q \mid x$. Thus we can write $x = qx_1$, $z = qz_1$, and so from (13) we have

$$(qx_1)^2 + pqy^2 = p(qz_1)^2 + qw^2.$$

Dividing by q and reordering the terms, we have

$$w^2 + pqz_1^2 = py^2 + qx_1^2,$$

which gives a new solution (w, z_1, y, x_1) of (13). Now x and z can't both be 0, as then (13) would give $py^2 = w^2$. This is impossible, as y and w aren't both 0, and so the LHS is exactly divisible by an odd power of p while the RHS is exactly divisible by an even power of p. Hence either $0 < z_1 < z$ or $0 < x_1 < x$ (or both!), so we have $s(w, z_1, y, x_1) = w + z_1 + y + x_1 < w + z + y + x = s(x, y, z, w)$. So we have found a solution of smaller size, contradicting the fact that we started with one of minimal size. Hence no solution can exist.

Of course if, instead, $\left(\frac{q}{p}\right) = -1$, then we simply swap the rôles of p and q in the above argument. □

Corollary 17.3. *If both p and q are primes $\equiv -1 \bmod 4$ then (13) has no solution in positive integers.*

Proof. In this case quadratic reciprocity tells us that $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$, so that one of these Legendre symbols is -1. Hence the condition that at least one of $\left(\frac{p}{q}\right), \left(\frac{q}{p}\right)$ is -1 in Theorem 17.2 applies. □

Notes.

1. If both $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right) = 1$, then (13) can have a nonzero solution. For instance, when $p = 5$, $q = 11$ we have $\left(\frac{5}{11}\right) = \left(\frac{11}{5}\right) = \left(\frac{1}{5}\right) = 1$, from Quadratic Reciprocity. And indeed the equation $x^2 + 55y^2 = 5z^2 + 11w^2$ has the nonzero solutions $(x, y, z, w) = (4, 0, 1, 1), (3, 1, 2, 2), (7, 1, 1, 3)$.

2. Theorem 17.2 can be strengthened to show that (13) has no solutions in nonnegative integers not all 0. To see this, you follow the proof as above, but the size of the new solution obtained, $w + z_1 + y + x_1$, is $< w + z + y + x$ only if x and z are not both 0. In this case the proof goes through as before.

   However, if $x = z = 0$ then (13) gives $py^2 = w^2$. But this has no nonzero solution, as is easily seen by replacing '2' by 'p' in Proposition 17.1 – the proof is just the same. Hence the case $x = z = 0$ cannot occur.

Our third example has a trickier proof, but again the underlying 'descent' method is the same.

17.4 Fermat's Last Theorem for exponent 4

Theorem 17.4. *The equation*

$$x^4 + y^4 = z^2 \tag{14}$$

has no solution in positive integers x, y, z.

Corollary 17.5 (Fermat's Last Theorem for exponent 4)**.** *The equation $x^4 + y^4 = z^4$ has no solution in positive integers x, y, z.*

This corollary is simply the special case of Theorem 17.4 where z is assumed to be a perfect square.

Proof of Theorem 17.4. (From H. Davenport, *The higher arithmetic. An introduction to the theory of numbers*, Longmans 1952, p.162). Suppose that (14) has such a solution. We can clearly assume that $z \neq 1$, i.e., that $z > 1$. We measure the size of a solution simply by z. Assume we have a solution with z minimal. If $d = \gcd(x, y) > 1$ we can divide by d^4, replacing x by x/d, y by y/d and z by z/d^2 in (14), obtaining a solution with z smaller. So we must have $\gcd(x, y) = 1$ for our minimal solution.

Now from Corollary 7.9 we know that

$$X^2 + Y^2 = Z^2$$

has general solution (with $\gcd(X, Y) = 1$), possibly after interchanging X and Y of

$$X = p^2 - q^2 \qquad Y = 2pq \qquad Z = p^2 + q^2,$$

where $p, q \in \mathbb{N}$ and $\gcd(p, q) = 1$, so

$$x^2 = p^2 - q^2 \qquad y^2 = 2pq \qquad z = p^2 + q^2.$$

As a square is $\equiv 0$ or $1 \bmod 4$, and x is odd (because $\gcd(x,y)=1$), we see that p is odd and q is even, say $q=2r$. So

$$x^2 = p^2 - (2r)^2 \qquad \left(\frac{y}{2}\right)^2 = pr.$$

Since $\gcd(p,r)=1$ and pr is a square, we have $p=v^2$ and $r=w^2$ say, so

$$x^2 + (2w^2)^2 = v^4.$$

Note that, as $\gcd(p,q)=1$, we have $\gcd(x,q)=1=\gcd(x,2w^2)$. Hence, on applying Corollary 7.9 again, we have

$$x = p_1^2 - q_1^2 \qquad 2w^2 = 2p_1q_1 \qquad v^2 = p_1^2 + q_1^2,$$

where $\gcd(p_1,q_1)=1$ and not both are odd. Say p_1 odd, q_1 even. Thus $w^2=p_1q_1$, giving $p_1=v_1^2$, $q_1=r_1^2$, say. Hence

$$v^2 (= p_1^2 + q_1^2) = v_1^4 + r_1^4,$$

which is another solution of (14)! But

$$v^2 = p = \sqrt{z-q^2} < \sqrt{z},$$

giving $v < z^{1/4}$, so certainly $v < z$ (as $z > 1$), contradicting the minimality of z. □

18 Representation of integers as sums of two squares

Which $n \in \mathbb{Z}$ can be represented as a sum $n = x^2 + y^2$ for $x, y \in \mathbb{Z}$? Obviously need $n \geq 0$. Can clearly assume that x and y are nonnegative. We have $0 = 0^2+0^2$, $1 = 1^2+0^2$, $2 = 1^2+1^2$, $4 = 2^2+0^2$, $5 = 2^2+1^2$, but no such representation for $n = 3$, 6 or 7.

Important note: $(2k)^2 \equiv 0 \bmod 4$, and $(2k+1)^2 = 8\binom{k+1}{2} + 1 \equiv 1 \bmod 8$ (and so certainly $\equiv 1 \bmod 4$).

18.1 The case $n = p$, prime

Which primes are the sum of two squares?

Theorem 18.1. *An odd prime p is a sum of two squares (of integers) iff $p \equiv 1 \bmod 4$.*

Proof. As $x^2, y^2 \equiv 0$ or $1 \bmod 4$, so $x^2+y^2 \equiv 0$ or 1 or $2 \bmod 4$. Assuming $p = x^2+y^2$, then as p is odd, we have $p \equiv 1 \bmod 4$.

Conversely, assume $p \equiv 1 \bmod 4$, and, knowing that then $\left(\frac{-1}{p}\right) = 1$, take $r \in \mathbb{N}$ with $r^2 \equiv -1 \bmod p$. Define $f(u,v) = u + rv$ and $K = \lfloor \sqrt{p} \rfloor$. Note that

$$K < \sqrt{p} < K+1, \tag{15}$$

as $\sqrt{p} \notin \mathbb{Z}$. Consider all pairs (u, v) with $0 \leq u \leq K$ and $0 \leq v \leq K$. There are $(K+1)^2 > p$ such pairs, and so the multiset of all $f(u, v)$ for such u, v has, by the Pigeonhole Principle, two such pairs $(u_1, v_1) \neq (u_2, v_2)$ for which $f(u_1, v_1) \equiv f(u_2, v_2) \bmod p$. Hence

$$\begin{aligned} u_1 + rv_1 &\equiv u_2 + rv_2 \bmod p \\ u_1 - u_2 &\equiv -r(v_1 - v_2) \bmod p \\ a &\equiv -rb \bmod p, \end{aligned}$$

say, where $a = u_1 - v_1$ and $b = v_1 - v_2$ are not both 0. Hence $a^2 \equiv -b^2 \bmod p$ as $r^2 \equiv -1 \bmod p$, so that $p \mid (a^2 + b^2)$. But $|a| \leq K$, $|b| \leq K$, giving

$$0 < a^2 + b^2 \leq 2K^2 < 2p.$$

So $a^2 + b^2 = p$. □

18.2 The general case

We now look at what happens if a prime $\equiv -1 \bmod 4$ *divides* a sum of two squares.

Proposition 18.2. *Let $q \equiv 3 \bmod 4$ be prime, and $q \mid (x^2 + y^2)$. Then $q \mid x$ and $q \mid y$, so that $q^2 \mid (x^2 + y^2)$.*

Proof. Assume that it is not the case that both x and y are divisible by q, say $q \nmid x$. Then from $x^2 + y^2 \equiv 0 \bmod q$ we get $(yx^{-1})^2 \equiv -1 \bmod q$, contradicting $\left(\frac{-1}{q}\right) = -1$. □

Proposition 18.3. *If n is a sum of two squares and m is a sum of two squares then so is nm.*

Proof. If $n = a^2 + b^2$ and $m = c^2 + d^2$ then

$$nm = (a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

□

(The latter identity comes from complex numbers:

$$(a + ib)(c + id) = ac - bd + i(ad + bc)$$

gives

$$|a + ib|^2 \cdot |c + id|^2 = |ac - bd + i(ad + bc)|^2$$

and hence the identity.)

Corollary 18.4. *If $n = A^2 \prod_i n_i$ where $A, n_i \in \mathbb{Z}$ and each n_i is a sum of two squares, then so is n.*

Proof. Use induction on i to get $n/A^2 = \prod_i n_i = a^2 + b^2$ say. Then $n = (Aa)^2 + (Ab)^2$. □

We can now state and prove our main result.

Theorem 18.5 (Fermat). *Write n in factorised form as*

$$n = 2^{f_2} \prod_{p \equiv 1 \bmod 4} p^{f_p} \prod_{q \equiv -1 \bmod 4} q^{g_q},$$

where (of course) all the p's and q's are prime. Then n can be written as the sum of two squares of integers iff all the g_q's are even.

Proof. If all the g_q are even then $n = A^2 \times$(product of some p's) and also $\times 2$ if f_2 is odd. So we have $n = A^2 \times \prod_i (a_i^2 + b_i^2)$ by Theorem 18.1 (using also $2 = 1^2 + 1^2$ if f_2 odd). Hence, by Corollary 18.4, n is the sum of two squares.

Conversely, suppose $q \mid n = a^2 + b^2$, where $q \equiv -1 \bmod 4$ is prime. Let q^k be the highest power of q dividing both a and b, so say $a = q^k a_1$, $b = q^k b_1$. Then

$$\frac{n}{q^{2k}} = a_1^2 + b_1^2.$$

Now $q \nmid \frac{n}{q^{2k}}$, as otherwise q would divide both a_1 and b_1, by Prop. 18.2. Hence q^{2k} is the highest power of q dividing n, i.e., $g_q = 2k$ is even. Hence all the g_q's are even. □

18.3 Related results

Proposition 18.6. *If an integer n is the sum of two squares of rationals then it's the sum of two squares of integers.*

Proof. Suppose that

$$n = \left(\frac{a}{b}\right)^2 + \left(\frac{c}{d}\right)^2$$

for some rational numbers a/b and c/d. Then

$$n(bd)^2 = (da)^2 + (bc)^2.$$

Hence, by Thm 18.5, for every prime $q \equiv -1 \bmod 4$ with $q^i || n(bd)^2$, i must be even. But then if $q^\ell || bd$ then $q^{i-2\ell} || n$, with $i - 2\ell$ even. Hence, by Thm 18.5 (in the other direction), n is the sum of two squares of integers. □

Corollary 18.7. *A rational number n/m is the sum of two squares of rationals iff nm is the sum of two squares of integers.*

Proof. If $nm = a^2 + b^2$ for $a, b \in \mathbb{Z}$ then

$$\frac{n}{m} = \left(\frac{a}{m}\right)^2 + \left(\frac{b}{m}\right)^2.$$

Conversely, if

$$\frac{n}{m} = \left(\frac{a}{b}\right)^2 + \left(\frac{c}{d}\right)^2$$

then

$$nm = \left(\frac{am}{b}\right)^2 + \left(\frac{cm}{d}\right)^2.$$

Hence, by Prop. 18.6, nm is the sum of two squares of integers. □

18.4 Finding all ways of expressing a rational as a sum of two rational squares

Now let h be a rational number that can be written as the sum of two squares of rationals. We can then describe *all* such ways of writing h.

Theorem 18.8. *Suppose that $h \in \mathbb{Q}$ is the sum of two rational squares: $h = s^2 + t^2$, where $s, t \in \mathbb{Q}$. Then the general solution of $h = x^2 + y^2$ in rationals x, y is*

$$x = \frac{s(u^2 - v^2) - 2uvt}{u^2 + v^2} \qquad y = -\left(\frac{t(u^2 - v^2) + 2uvs}{u^2 + v^2}\right), \tag{16}$$

where $u, v \in \mathbb{Z}$ and are not both zero.

Proof. We are looking for all points $(x, y) \in \mathbb{Q}^2$ on the circle $x^2 + y^2 = h$. If (x, y) is such a point, then for $x \neq s$ the chord through (s, t) and (x, y) has rational slope $(t - y)/(s - x)$.

Conversely, take a chord through (s, t) of rational slope r, which has equation $y = r(x - s) + t$. Then for the intersection point (x, y) of the chord and the circle we have

$$x^2 + (r(x - s) + t)^2 = h,$$

which simplifies to

$$x^2(1 + r^2) + 2rx(t - rs) + (r^2 - 1)s^2 - 2rst = 0,$$

using the fact that $t^2 - h = -s^2$. This factorises as

$$(x - s)((1 + r^2)x + 2rt + s(1 - r^2)) = 0.$$

For $x \neq s$ we have

$$x = \frac{s(r^2 - 1) - 2rt}{1 + r^2}$$

and

$$\begin{aligned} y &= t + r(x - s) \\ &= -\left(\frac{t(r^2 - 1) + 2sr}{1 + r^2}\right), \end{aligned}$$

on simplification. Finally, substituting $r = u/v$ gives (16). Note that $v = 0$ in (16) (i.e., $r = \infty$) gives the point $(r, -s)$. □

Corollary 18.9. *The general integer solution x, y, z of the equation $x^2 + y^2 = nz^2$ is*

$$(x, y, z) = (a(u^2 - v^2) - 2uvb, b(u^2 - v^2) + 2uva, u^2 + v^2),$$

where $n = a^2 + b^2$, with $a, b, u, v \in \mathbb{Z}$, and u, v arbitrary.

(If n is not the sum of two squares, then the equation has no nonzero solution, by Prop. 18.6.)

In particular, for $n = 1 = 1^2 + 0^2$, we see that the general integer solution to Pythagoras' equation $x^2 + y^2 = z^2$ is

$$(x, y, z) = (u^2 - v^2, 2uv, u^2 + v^2).$$

For a socalled *primitive* solution — one with $\gcd(x, y) = 1$ — choose u, v with $\gcd(u, v) = 1$ and not both odd.

The same method works for $Ax^2 + By^2 + Cz^2 = 0$.

18.5 Sums of three squares, sums of four squares

Proposition 18.10. *No number of the form $4^a(8k + 7)$, where a is a nonnegative integer, is the sum of three squares (of integers).*

Proof. Use induction on a. For $a = 0$: Now $n^2 \equiv 0, 1$ or $4 \bmod 8$, so a sum of three squares is $\equiv 1$ or 1 or 2 or 3 or 4 or 5 or $6 \bmod 8$, but $\not\equiv 7 \bmod 8$.

Assume result true for some integer $a \geq 0$. If $4^{a+1}(8k + 7) = n_1^2 + n_2^2 + n_3^2$ then all the n_i must be even, and so $= 4(n_1'^2 + n_2'^2 + n_3'^2)$ say. But then $4^a(8k + 7) = n_1'^2 + n_2'^2 + n_3'^2$, contrary to the induction hypothesis. □

In fact (won't prove)

Theorem 18.11 (Legendre 1798, Gauss)**.** *All positive integers except those of the form $4^a(8k + 7)$ are the sum of three squares.*

Assuming this result, we can show

Corollary 18.12 (Lagrange 1770)**.** *Every positive integer is the sum of four squares.*

Proof. The only case we need to consider is $n = 4^a(8k+7)$. But then $n-(2^a)^2 = 4^a(8k+6) = 2^{2k+1}(4k + 3)$, which (being exactly divisible by an odd power of 2) is not of the form $4^{a'}(8k' + 7)$, so is the sum of three squares. □

19 Primality testing

19.1 Introduction

Factorisation is concerned with the problem of developing efficient algorithms to express a given positive integer $n > 1$ as a product of powers of distinct primes. With primality testing, however, the goal is more modest: given n, decide whether or not it is prime. If n does turn out to be prime, then of course you've (trivially) factorised it, but if you show that it is not prime (i.e., *composite*), then in general you have learnt nothing about its factorisation (apart from the fact that it's not a prime!).

One way of testing a number n for primality is the following: suppose a certain theorem, Theorem X say, whose statement depends on a number n, is true when n is prime. Then if Theorem X is false for a particular n, then n cannot be prime. For instance, we know (Fermat) that $a^{n-1} \equiv 1 \bmod n$ when n is prime and $n \nmid a$. So if for such an a we have $a^{n-1} \not\equiv 1 \pmod{n}$, then n is not prime. This test is called the *Pseudoprime Test to base a*. Moreover, a composite number n that passes this test is called a *Pseudoprime to base a*.

(It would be good if we could find a Theorem Y that was true *iff* n was prime, and was moreover easy to test. Then we would know that if the theorem was true for n then n was prime. A result of this type is the following (also on a problem sheet): n is prime iff $a^{n-1} \equiv 1 \bmod n$ for $a = 1, 2, \ldots, n-1$. This is, however, not easy to test; it is certainly no easier than testing whether n is divisible by a for $a = 1, \ldots, n$.)

19.2 Proving primality of n when $n-1$ can be factored

In general, primality tests can only tell you that a number n either 'is composite', or 'can't tell'. They cannot confirm that n is prime. However, under the special circumstance that we can factor $n-1$, primality can be proved:

Theorem 19.1 (Lucas Test, as strengthened by Kraitchik and Lehmer)**.** *Let $n > 1$ have the property that for every prime factor q of $n-1$ there is an integer a such that $a^{n-1} \equiv 1 \bmod n$ but $a^{(n-1)/q} \not\equiv 1 \pmod{n}$. Then n is prime.*

Proof. Define the subgroup G of $(\mathbb{Z}/n\mathbb{Z})^\times$ to be the subgroup generated by all such a's. Clearly the exponent of G is a divisor of $n-1$. But it can't be a proper divisor of $n-1$, for then it would divide some $(n-1)/q$ say, which is impossible as $a^{(n-1)/q} \not\equiv 1 \pmod{n}$ for the a corresponding to that q. Hence G has exponent $n-1$. But then $n-1 \leq \#G \leq \#(\mathbb{Z}/n\mathbb{Z})^\times = \varphi(n)$. Hence $\varphi(n) = n-1$, which immediately implies that n is prime. □

Corollary 19.2 (Pepin's Test, 1877)**.** *Let $F_k = 2^{2^k} + 1$, the kth Fermat number, where $k \geq 1$. Then F_k is prime iff $3^{\frac{F_k-1}{2}} \equiv -1 \bmod F_k$.*

Proof. First suppose that $3^{\frac{F_k-1}{2}} \equiv -1 \bmod F_k$. We apply the theorem with $n = F_k$. So $n - 1 = 2^{2^k}$ and $q = 2$ only, with $a = 3$. Then $3^{\frac{F_k-1}{2}} \not\equiv 1 \pmod{F_k}$ and (on squaring) $3^{F_k-1} \equiv 1 \bmod F_k$, so all the conditions of the Theorem are satisfied.

Conversely, suppose that F_k is prime. Then, by Euler's criterion and quadratic reciprocity (see Chapter 5) we have

$$3^{\frac{F_k-1}{2}} \equiv \left(\frac{3}{F_k}\right) = \left(\frac{F_k}{3}\right) = \left(\frac{2}{3}\right) = -1,$$

as 2 is not a square (mod 3).

□

We can use this to show that $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$ are all prime. It is known that F_k is composite for $5 \leq k \leq 32$, although complete factorisations of F_k are known only for $0 \leq k \leq 11$, and there are no known factors of F_k for $k = 20$ or 24. Heuristics suggest that there may be no more k's for which F_k is prime.

19.3 Carmichael numbers

A *Carmichael number* is a (composite) number n that is a pseudoprime to every base a with $1 \leq a \leq n$ and $\gcd(a, n) = 1$. Since it it immediate that $a^{n-1} \not\equiv 1 \pmod{n}$ when $\gcd(a, n) > 1$, we see that Carmichael numbers are pseudoprimes to as many possible bases as any composite number could be. They are named after the US mathematician Robert Carmichael (1879 – 1967).

[But even *finding* an a with $\gcd(a, n) > 1$ gives you a factor of n. (Imagine that n is around 10^{300} and is a product of three 100-digit primes – such a's are going to be few and far between!)]

Example 1. The number $n = 561 = 3.11.17$ is a Carmichael number. To see this tak $a : \gcd(a, 561) = 1$, so that a is coprime to each of 3, 11 and 17. So, by Fermat, we have $a^2 \equiv 1 \bmod 3$, $a^{10} \equiv 1 \bmod 11$ and $a^{16} \equiv 1 \bmod 17$. Now $\text{lcm}(2, 10, 16) = 80$ so that, taking appropriate powers, we have that $a^{80} \equiv 1 \bmod 3.11.17$. Finally $a^{560} = (a^{80})^7 \equiv 1^7 \equiv 1 \bmod 560$, so that indeed $n = 561$ is Carmichael.

For more examples of Carmichael numbers, see Workshop 4.

19.3.1 Properties of Carmichael numbers

Theorem 19.3 (See Qs 14 and 15, Workshop 4)**.** *An integer $n > 1$ is a Carmichael number iff n is squarefree and $p - 1 \mid n - 1$ for each prime p dividing n.*

Proposition 19.4 (See Q 18, Workshop 4)**.** *Every Carmichael number has at least 3 distinct prime factors.*

A curious result is the following.

Theorem 19.5 (See Q 17, Workshop 4)**.** *An integer $n > 1$ has the property that*

$$(a + b)^n \equiv a^n + b^n \bmod n \qquad \text{for all } a, b \in \mathbb{Z}$$

iff either n is a prime number or n is a Carmichael number.

19.4 Strong pseudoprimes

Given $n > 1$ odd and an a such that $a^{n-1} \equiv 1 \bmod n$, factorise $n - 1$ as $n - 1 = 2^f q$, where q is odd, $f \geq 1$ and consider the sequence

$$\mathcal{S} = [a^q, a^{2q}, a^{4q}, \ldots, a^{2^f q} \equiv 1],$$

taken $\pmod{n}$. If n is prime then, working left to right, either $a^q \equiv 1 \bmod n$, in which case $\mathcal{S}$ consists entirely of 1's, or the number before the first 1 must be -1. This is because the number following any x in the sequence is x^2, so if $x^2 \equiv 1 \bmod n$ for n prime, then

$x \equiv \pm 1 \bmod n$. (Why?) A composite number n that has this property, (i.e., is a pseudoprime to base a and for which either $\mathcal{S}$ consists entirely of 1's or the number before the first 1 in $\mathcal{S}$ is -1) is called a *strong pseudoprime to base a*.

Clearly, if n is a prime or pseudoprime but not a strong pseudoprime, then this stronger test proves that n isn't prime. This is called the *Miller-Rabin Strong Pseudoprime Test.*

Example 2. Take $n = 31621$. It is a pseudoprime to base $a = 2$, as $2^{n-1} \equiv 1 \bmod n$ but $5^{n-1} \equiv 12876 \bmod n$ (so n not prime). We have $n - 1 = 2^2 \cdot 7905$, $2^{7905} \equiv 31313 \bmod n$ and $2^{15810} \equiv 2^{31620} \equiv 1 \bmod n$, so n is not a strong pseudoprime to base 2.

19.5 Strong pseudoprimes to the smallest prime bases

It is known that

- 2047 is the smallest strong pseudoprime to base 2;
- 1373653 is the smallest strong pseudoprime to both bases 2, 3;
- 25326001 is the smallest strong pseudoprime to all bases 2, 3, 5;
- 3215031751 is the smallest strong pseudoprime to all bases 2, 3, 5, 7;
- 2152302898747 is the smallest strong pseudoprime to all bases 2, 3, 5, 7, 11;
- 3474749660383 is the smallest strong pseudoprime to all bases 2, 3, 5, 7, 11, 13;
- 341550071728321 is the smallest strong pseudoprime to all bases 2, 3, 5, 7, 11, 13, 17.

(In fact 341550071728321 is also a strong pseudoprime to base 19.)

Hence any odd $n < 341550071728321$ that passes the strong pseudoprime test for all bases 2, 3, 5, 7, 11, 13, 17 must be prime. So this provides a cast-iron primality test for all such n.

19.6 Factorising weak pseudoprimes

Let us call a pseudoprime to base a that is not a strong pseudoprime to base a a *weak pseudoprime to base a*.

Theorem 19.6. *An odd weak pseudoprime n to base a can be factored into $n = n_1 n_2$, where $n_1, n_2 > 1$.*

Proof. When the strong pseudoprime test detects n as being composite, what happens is that some $x \in \mathcal{S}$ is a solution to $x^2 \equiv 1 \bmod n$ with $x \not\equiv \pm 1 \pmod n$ because $x \equiv 1 \bmod n_1$ and $x \equiv -1 \bmod n_2$ for some coprime n_1, n_2 with $n_1 n_2 = n$. And then both $g_- := \gcd(x-1, n)$ (divisible by n_1) and $g_+ := \gcd(x+1, n)$ (divisible by n_2) are nontrivial factors of n. Further, $2 = (x+1) - (x-1) = k_+ g_+ - k_- g_-$ say, for some integers k_+, k_-. So, because n is (assumed) odd, g_+ and g_- are coprime. As they are also factors of n, they must actually *equal* n_1 and n_2 respectively. □

Example 2 revisited. Take $n = 31621$.Then $x = 31313$ and $\gcd(n, 31312) = 103$ and $\gcd(n, 31314) = 307$, giving the factorisation $n = 103 \cdot 307$.

Note that if $n = n_1 n_2$ where n_1 and n_2 are coprime integers, then by the Chinese Remainder Theorem we can solve each of the four sets of equations

$$x \equiv \pm 1 \bmod n_1 \qquad\qquad x \equiv \pm 1 \bmod n_2$$

to get four distinct solutions of $x^2 \equiv 1 \bmod n$. For instance, for $n = 35$ get $x = \pm 1$ or ± 6. For the example $n = 31621$ above, we have $31313 \equiv 1 \bmod 103$ and $31313 \equiv -1 \bmod 307$, so that four distinct solutions of $x^2 \equiv 1 \bmod 31621$ are ± 1 and ± 31313.

19.7 Primality testing in 'polynomial time'

In 2002 the Indian mathematicians Agrawal, Kayal and Saxena invented an algorithm, based on the study of the polynomial ring $(\mathbb{Z}/n\mathbb{Z})[x]$, that was able to decide whether a given n was prime in time $O((\log n)^{6+\varepsilon})$. (Here the constant implied by the 'O' depends on ε and so could go to infinity as $\varepsilon \to 0$.) (Search for 'AKS algorithm' on web.)

19.8 The Lucas-Lehmer primality test for Mersenne numbers

Given an odd prime p, let $M_p = 2^p - 1$, a *Mersenne number* (and a Mersenne prime iff it is prime). [It is an easy exercise to prove that if p is composite, then so is M_p.]

Define a sequence $S_1, S_2, \ldots, S_n, \ldots$ by $S_1 = 4$ and $S_{n+1} = S_n^2 - 2$ for $n = 1, 2, \ldots$. so we have

$$S_1 = 4, S_2 = 14, S_3 = 194, S_4 = 37634, S_5 = 1416317954, \ldots.$$

There is a very fast test for determining whether or not M_p is prime.

Theorem 19.7 (Lucas-Lehmer Test)**.** *For an odd prime p, the Mersenne number M_p is prime iff M_p divides S_{p-1}.*

So $M_3 = 7$ is prime as $7 \mid S_2$, $M_5 = 31$ is prime as $31 \mid S_4$,.... In this way get M_p prime for $p = 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, \ldots$ (47th) 43112609. There may be others between the 41st and 47th. [as at October 2012.]

For the proof, we need two lemmas.

Lemma 19.8. *Put $\omega = 2 + \sqrt{3}$ and $\omega_1 = 2 - \sqrt{3}$. Then $\omega\omega_1 = 1$ (immediate) and*

$$S_n = \omega^{2^{n-1}} + \omega_1^{2^{n-1}}$$

for $n = 1, 2, \ldots$.

The proof is a very easy induction exercise.

Lemma 19.9. *Let r be a prime $\equiv 1 \bmod 3$ and $\equiv -1 \bmod 8$ (i.e., $\equiv 7 \bmod 24$). Then*

$$\omega^{\frac{r+1}{2}} \equiv -1 \bmod r.$$

(So it's equal to $a + b\sqrt{3}$ where $a \equiv -1 \bmod r$ and $b \equiv 0 \bmod r$.)

Proof. Put

$$\tau = \frac{1+\sqrt{3}}{\sqrt{2}} \qquad \text{and} \qquad \tau_1 = \frac{1-\sqrt{3}}{\sqrt{2}}.$$

Then we immediately get $\tau\tau_1 = -1$, $\tau^2 = \omega$ and $\tau_1^2 = \omega_1$. Next, from $\tau\sqrt{2} = 1+\sqrt{3}$ we have $(\tau\sqrt{2})^r = (1+\sqrt{3})^r$, so that

$$\begin{aligned} \tau^r 2^{\frac{r-1}{2}}\sqrt{2} &= 1 + \sum_{j=1}^{r-1} \binom{r}{j} (\sqrt{3})^j + 3^{\frac{r-1}{2}}\sqrt{3} \\ &\equiv 1 + 3^{\frac{r-1}{2}}\sqrt{3} \bmod r, \end{aligned} \tag{17}$$

as $r \mid \binom{r}{j}$. Since $r \equiv -1 \bmod 8$ we have

$$2^{\frac{r-1}{2}} \equiv \left(\frac{2}{r}\right) = (-1)^{\frac{r^2-1}{8}} \equiv 1 \bmod r,$$

using Euler's Criterion, and Prop. 5.3. Further, since $r \equiv 1 \bmod 3$ and $r \equiv -1 \bmod 4$ we have

$$3^{\frac{r-1}{2}} \equiv \left(\frac{3}{r}\right) = \left(\frac{r}{3}\right)(-1)^{\frac{r-1}{2}\cdot\frac{3-1}{2}} = \left(\frac{1}{3}\right)\cdot(-1) \equiv -1 \bmod r,$$

using Euler's Criterion again, and also Quadratic Reciprocity (Th. 5.1). So, from (17), we have successively

$$\begin{aligned} \tau^r\sqrt{2} &\equiv 1 - \sqrt{3} \bmod r \\ \tau^r &\equiv \tau_1 \bmod r \\ \tau^{r+1} &\equiv \tau\tau_1 = -1 \pmod r \\ \omega^{\frac{r+1}{2}} &\equiv -1 \bmod r, \end{aligned}$$

the last step using $\tau^2 = \omega$. □

Proof of Theorem 19.7. **$\mathbf{M_p}$ prime $\Rightarrow$ $\mathbf{M_p \mid S_{p-1}}$.** Assume M_p prime. Apply Lemma 19.9 with $r = M_p$, which is allowed as $M_p \equiv -1 \bmod 8$ and $M_p \equiv (-1)^p - 1 \equiv 1 \bmod 3$. So

$$\omega^{\frac{M_p+1}{2}} = \omega^{2^{p-1}} \equiv -1 \bmod M_p \tag{18}$$

and, using Lemma 19.8, including $\omega_1^{-1} = \omega$, we have

$$S_{p-1} = \omega^{2^{p-2}} + \omega_1^{2^{p-2}} = \omega_1^{2^{p-2}}\left(\left(\omega_1^{-1}\right)^{2^{p-2}}\omega^{2^{p-2}} + 1\right) = \omega_1^{2^{p-2}}\left(\omega^{2^{p-1}} + 1\right) \equiv 0 \bmod M_p, \tag{19}$$

the last step using (18).

$M_p \mid S_{p-1} \Rightarrow M_p$ prime. Assume $M_p \mid S_{p-1}$ but M_p composite. We aim for a contradiction. Then M_p will have a prime divisor q (say) with $q^2 \leq M_p$.

Now consider the multiplicative group $G = \left(\frac{\mathbb{Z}[\sqrt{3}]}{(q)}\right)^\times$ of units of the ring $\frac{\mathbb{Z}[\sqrt{3}]}{(q)}$. Then G has coset representatives consisting of numbers $a+b\sqrt{3}$ with $a, b \in \{0, 1, 2, \ldots, q-1\}$ that are also invertible (mod q). So G is a group of size (order) at most $q^2 - 1$, with multiplication defined modulo q. From $\omega(\omega_1 + q\sqrt{3}) \equiv 1 \bmod q$ we see that $\omega = 2 + \sqrt{3}$ is invertible, and so $\omega \in G$. [Strictly speaking, the coset $\omega \pmod q \in G$.]

Now, using $M_p \mid S_{p-1}$ we see that (19) holds even when M_p is composite, so we have successively that $\omega^{2^{p-1}} + 1 \equiv 0 \bmod M_p$, $\omega^{2^{p-1}} \equiv -1 \bmod q$ and $\omega^{2^p} \equiv 1 \bmod q$. Hence the order of ω in G is 2^p. Then $2^p \mid \#G \leq q^2 - 1 \leq M_p - 1 = 2^p - 2$, a contradiction. Hence M_p must be prime.

□

In practice, to test M_p for primality using Theorem 19.7, one doesn't need to compute $S_j (j = 1, 2, \ldots, p-1)$, but only the much smaller (though still large!) numbers $S_j \pmod{M_p} (j = 1, 2, \ldots, p-1)$.

A good source of information on Mersenne numbers is
`http://primes.utm.edu/mersenne/index.html`

20 Integer Factorisation

In this chapter we review the historic techniques of Trial Division, the Sieve of Eratosthenes, and Fermat's factorisation method. We then study two simply-programmable integer factorisation algorithms, both due to Pollard.

20.1 Trial Division

Given $n > 1$, try dividing n successively by the primes $2, 3, \ldots$, up to the largest prime $\leq \sqrt{n}$. If any such prime divides n, then of course you have found a factor, and you can continue the process by applying the same procedure to n/p. On the other hand, if none of these primes divides n, then n itself is prime. Why?

Lemma 20.1. *If $n > 1$ is composite then it is divisible by a prime $\leq \sqrt{n}$.*

Proof. Say $n = n_1 n_2$, where $n_1, n_2 > 1$. If both were $> \sqrt{n}$ then $n = n_1 n_2$ would be $> \sqrt{n}^2 = n$, a contradiction. Hence one of n_1 or n_2, say n_1, is $\leq \sqrt{n}$. Then any prime factor p of n_1 certainly divides n, and so $p \leq n_1 \leq \sqrt{n}$, as required. □

Trial division requires knowledge of all primes $\leq \sqrt{n}$. How to find them?

20.2 The Sieve of Eratosthenes

To find all primes up to N (e.g., for $N = \lfloor\sqrt{n}\rfloor$), write down $2, 3, 4, 5, 6, \ldots, N$ and

- cross off all multiples of 2, except 2 itself. Then the first uncrossedout number (3) is prime.
- cross off all multiples of 3, except 3. Then the first uncrossedout number (5) is prime.

Proceed in this way until you have crossed out all multiples of p, except p itself, for all primes $\leq \sqrt{N}$. Then the uncrossedout numbers consist of all the primes $\leq N$. This is because, by Lemma 20.1, all composite numbers $\leq N$ are divisible by a prime $\leq \sqrt{N}$, and so have been crossed out.

Thus to apply trial division on n you would need to apply the Sieve of Eratosthenes with $N \approx n^{1/4}$ in order to find all primes up to $n^{1/2}$.

20.3 Fermat's factorisation method

Take $n > 1$ and odd. Fermat's idea is to try to write n as $n = x^2 - y^2$, as then $n = (x+y)(x-y)$. So if $x > y+1$ we get a nontrivial factorisation of n.

We successively try $x = \lceil\sqrt{n}\rceil$, $\lceil\sqrt{n}\rceil + 1$, $\lceil\sqrt{n}\rceil + 2, \ldots$, until $x^2 - n$ is a square, $= y^2$ say. Then $x^2 - n = y^2$, or $n = (x+y)(x-y)$. (This process will eventually terminate, as for $x = (n+1)/2$ we have $x^2 - n = ((n-1)/2)^2$. But this only give the trivial factorisation $n = n \cdot 1$.)

Example 1. $n = 2479$, $\lceil\sqrt{n}\rceil = 50$, $50^2 - n = 21$, $51^2 - n = 122$, $52^2 - n = 225 = 15^2$, giving $n = 52^2 - 15^2 = (52+15)(52-15) = 67 \cdot 37$.

Example 2. $n = 3953$, $\lceil\sqrt{n}\rceil = 63$, $63^2 - n = 16 = 4^2$, giving $n = 63^2 - 4^2 = (63+4)(63-4) = 67 \cdot 59$.

This method works well if n has two factors close together (so that y is small), but is otherwise slow. However, the idea of trying to write n as a difference of two squares is a factorisation idea used in several other factorisation algorithms, for instance in the Quadratic Sieve algorithm.

20.4 Pollard's $p-1$ method

Take $n > 1$ and odd, and suppose that n has a prime factor p. Then, if $p - 1 \mid k!$ for some k, say $k! = (p-1)q$, then

$$2^{k!} = (2^q)^{p-1} \equiv 1 \bmod p,$$

by Fermat's Little Theorem, so that $p \mid 2^{k!} - 1$. Hence $p \mid \gcd(2^{k!} - 1, n)$. So long as this gcd isn't n, we obtain a nontrivial (i.e., not 1 or n) factor of n.

So algorithm is:

Compute modulo n $2, 2^{2!} = 2^2, 2^{3!} = 2^{2!3}, 2^{4!} = 2^{3!4}, \ldots, 2^{k!} = 2^{(k-1)!k}$ until $n > \gcd(n, 2^{k!} - 1) > 1$. Then $\gcd(n, 2^{k!} - 1)$ is a nontrivial factor of n.

Maple code for Pollard $p - 1$:

```
r:=2;g:=1;
for k to n while g=1 or g=n do
r:=r^k mod n; g:=gcd(r-1,n);
end do;
print(g,k);
```

At worst k could be near $(n-1)/2$, but is sometimes much smaller. It is generally large when all prime factors p of n are such that $p - 1$ has a large prime factor. It is small when n has a prime factor p for which all prime factors of $p - 1$ are small.

Example 1 again. $n = 2479$. Here $k = 6$ is enough, as $37 - 1 = 36 \mid 6!$, showing that $p = 37$ is a factor.

Example 2 again. $n = 3953$. Here $k = 11$ is enough, as $67 - 1 = 66 \mid 11!$, showing that $p = 67$ is a factor.

20.5 Pollard rho

The idea: for some function $f : \mathbb{N} \to \mathbb{N}$ define an integer sequence, starting with a 'seed' x_0, and defining $x_{k+1} \equiv f(x_k) \bmod n$ for $k \geq 0$. if these numbers are fairly random (mod n) then we'd expect to need about $\sqrt{n}$ of them before two will be equal (mod n). [Compare the 'Birthday Paradox' in Probability Theory, where 23 people chosen at random have, under standard assumptions, a 50% probability of containing a pair that share a birthday.] However, if p is the smallest prime factor of n, and p is much smaller than n, we'd expect that roughly $\sqrt{p}$ of the x_i are needed before two are equal (mod p). Then if indeed $x_i \equiv x_j \bmod p$ we have $p \mid \gcd(x_i - x_j, n)$. Provided that $x_i \not\equiv x_i \pmod{n}$, this will yield a proper factor of n.

The name 'Pollard rho' comes from the ρ-shaped diagram you can draw, consisting of a path from x_0 to x_1, x_1 to x_2, and so on, until the path curls around to intersect itself with $x_j \equiv x_i \bmod p$.

In practice we can take $x_0 = 2$ and $f(x) = x^2 + 1$. If $x_i \equiv x_j \bmod p$ with $0 < i < j$, then

$$x_{i+1} \equiv f(x_i) \equiv f(x_j) \equiv x_{j+1} \bmod p.$$

Proceeding in this way, we have

$$x_{i+s} \equiv x_{j+s} \bmod p \qquad \text{for } s = 1, 2, 3, \ldots. \tag{20}$$

Also

$$x_i \equiv x_j \equiv f^{j-i}(x_i) \equiv f^{j-i}(x_j) \equiv x_{j+(j-i)} \equiv x_{2j-i},$$

where f^{j-i} is the $(j-i)$-fold iterate of f. Hence we can add $j-i$ to the index repeatedly, to obtain

$$x_i \equiv x_j \equiv x_{j+(j-i)} \equiv x_{j+2(j-i)} \equiv \ldots \bmod p.$$

Thus, by if necessary replacing j by $j \quad +$ (a multiple of $j-i$), we can make j as large as we like. In particular, we can assume that $j \geq 2i$.

Now take $s = j - 2i$ in (20), giving $x_{j-i} \equiv x_{2(j-i)} \bmod p$. So in fact we just need to find some k such that

$$n > \gcd(x_{2k} - x_k, n) > 1.$$

(So we do not need to compare x_j with all previous x_i's for $i < j$.)

Maple code for Pollard rho:

```
g:=1;x[0]:=2;
for k to 100 while g=1 or g=n do
x[k]:=x[k-1]^2+1 mod n;
if k mod 2 = 0 then g:=gcd(x[k]-x[k/2],n); end if;
end do;
k:=k-1;
print(k,g);
```

(The choice of 100 as the maximum value for k is somewhat *ad hoc*, and can of course be increased.)

Example 1 yet again. $n = 2479$. Here $k = 6$, and $g = 37$ is a factor.

Example 2 yet again. $n = 3953$. Here $k = 12$, and $g = 59$ is a factor.

Example 3. $n = 1009^2$. Here $k = 98$ and $g = 1009$ is a (prime) factor.

This last example shows that the algorithm does not work so well (i.e., k is large) if the prime factors of n are large.

20.6 Final remarks.

- To specify a factoring algorithm, it's enough to have a general method that, for a given composite n, factors n as $n = n_1 n_2$, where both $n_1, n_2 > 1$. For then you can test n_1 and n_2 for primality and, if either is composite, recursively apply your algorithm to them. In this way you will eventually be able to write n as a product of powers of distinct primes. So your algorithm does not need to explicitly specify how to do this.

- In order to factor n, it's enough to find k: $1 < \gcd(k, n) < n$, as then $\gcd(k, n)$ is a nontrivial factor of n, with $n = n_1 n_2$, where $n_1 = \gcd(k, n)$ and $n_2 = n/\gcd(k, n)$.

  But if say $n = pq$ where p, q are primes $\approx 10^{300}$, then $\varphi(n) = (p-1)(q-1) = n-p-q+1 \approx 10^{600}$, and $n - \varphi(n) = p+q-1 \approx 2 \cdot 10^{300}$. So a random $k \in \{1, 2, \ldots, n\}$ has a probability of $\approx 2 \cdot 10^{-300}$ of having $\gcd(k, n) > 1$ – vanishingly small!

- If we can find a solution x to the equation $x^2 \equiv 1 \bmod n$ that's not $x = \pm 1$ then we can factor n. This is because such a solution will produce $n = n_1 n_2$ where $n_1 = \gcd(x-1, n)$ and $n_2 = \gcd(x+1, n)$. For more details see also the end of Chapter 6, where this method is applied to factorise a 'weak pseudoprime'.

  Conversely, any nontrivial factorisation $n = n_1 n_2$ with $\gcd(n_1, n_2) = 1$ gives rise to four solutions of $x^2 \equiv 1 \bmod n$. This is because we can use the Chinese Remainder Theorem to solve the equations $x \equiv -1 \bmod n_1$, $x \equiv 1 \bmod n_2$. Then x and $-x$ are both solutions of $x^2 \equiv 1 \bmod n$, and neither is either of ± 1.

- Other factorisation methods:

  - The Quadratic Sieve – the best general algorithm for numbers up to 10^{100};
  - The General Number Field Sieve – best for larger n (not of a special form).

  For more factorisation methods see Wikipedia "integer_factorization".

ADDITIONAL TOPICS
Notes by Prof. Chris Smyth

21 Dirichlet series

For an arithmetic function f, define its *Dirichlet series* $D_f(s)$ by

$$D_f(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

Here $s \in \mathbb{C}$ is a parameter. Typically, such series converge for $\Re s > 1$, and can be meromorphically continued to the whole complex plane. However, we will not be concerned with analytic properties of Dirichlet series here, but will regard them only as generating functions for arithmetic functions, and will manipulate them formally, without regard to convergence.

The most important example is for $f(n) = 1 \quad (n \in \mathbb{N})$, which gives the Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$. Also, taking $f(n) = n \quad (n \in \mathbb{N})$ gives $\zeta(s-1)$. (Check!).

Proposition 21.1. *If f is multiplicative then*

$$D_f(s) = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots + \frac{f(p^k)}{p^{ks}} + \ldots\right) = \prod_p D_{f,p}(s), \tag{21}$$

say.

Proof. Expanding the RHS of (21), a typical term is

$$\frac{f(p_1^{e_1})f(p_2^{e_2})\ldots f(p_r^{e_r})}{p_1^{e_1}p_2^{e_2}\ldots p_r^{e_r}} = \frac{f(n)}{n^s}$$

for $n = \prod_{i=1}^{r} p_i^{e_i}$, using the fact that f is multiplicative. □

Such a product formula $D_f(s) = \prod_p D_{f,p}(s)$ over all primes p is called an *Euler product* for $D_f(s)$.

For example

$$\zeta(s) = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots + \frac{1}{p^{ks}} + \ldots\right) = \prod_p \left(\frac{1}{1-p^{-s}}\right),$$

on summing the Geometric Progression (GP). Hence also

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - p^{-s}\right) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = D_\mu(s),$$

on expanding out the product.

Proposition 21.2. *We have*

$$\left(\sum_k \frac{a_k}{k^s}\right) \cdot \left(\sum_\ell \frac{b_\ell}{\ell^s}\right) = \left(\sum_n \frac{c_n}{n^s}\right),$$

where $c_n = \sum_{k|n} a_k b_{n/k}$.

Proof. On multiplying out the LHS, a typical term is

$$\frac{a_k}{k^s} \cdot \frac{b_\ell}{\ell^s} = \frac{a_k b_{n/k}}{n^s},$$

where $k\ell = n$. So all pairs k, ℓ with $k\ell = n$ contribute to the numerator of the term with denominator n^s. □

Corollary 21.3. *We have* $D_F(s) = D_f(s)\zeta(s)$.

Proof. Apply the Proposition with $a_k = f(k)$ and $b_\ell = 1$. □

Corollary 21.4 (Möbius inversion again)**.** *We have* $f(n) = \sum_{d|n} \mu(n/d)F(d)$ *for all* $n \in \mathbb{N}$.

Proof. From Corollary 21.3 we have

$$D_f(s) = D_F(s) \cdot \frac{1}{\zeta(s)} = \left(\sum_k \frac{F(k)}{k^s}\right) \cdot \left(\sum_\ell \frac{\mu(\ell)}{\ell^s}\right) = \left(\sum_n \frac{c_n}{n^s}\right),$$

where $c_n = \sum_{k|n} F(k)\mu(n/k)$. But $D_f(s) = \sum_{n=1}^\infty \frac{f(n)}{n^s}$, so, on comparing coefficients, $f(n) = \sum_{k|n} F(k)\mu(n/k)$. □

We now compute the Dirichlet series for a few standard functions. [Part (a) is already proved above.]

Proposition 21.5. *We have*

(a) $D_\mu(s) = \frac{1}{\zeta(s)}$;

(b) $D_\varphi(s) = \frac{\zeta(s-1)}{\zeta(s)}$;

(c) $D_\tau(s) = \zeta(s)^2$;

(d) $D_\sigma(s) = \zeta(s-1)\zeta(s)$.

Proof. (b) Now

$$\begin{aligned} D_\varphi(s) &= \prod_p \left(1 + \frac{\varphi(p)}{p^s} + \frac{\varphi(p^2)}{p^{2s}} + \cdots + \frac{\varphi(p^k)}{p^{ks}} + \dots\right) \\ &= \prod_p \left(1 + \frac{p-1}{p^s} + \frac{p^2-p}{p^{2s}} + \cdots + \frac{p^k - p^{k-1}}{p^{ks}} + \dots\right) \\ &= \prod_p \left(1 + \frac{p-1}{p^s} \cdot \frac{1}{1-p^{1-s}}\right), \qquad \text{on summing the GP} \\ &= \prod_p \left(\frac{1-p^{-s}}{1-p^{-(s-1)}}\right), \qquad \text{on simplification} \\ &= \frac{\zeta(s-1)}{\zeta(s)}. \end{aligned}$$

(c) Now

$$\begin{aligned} D_\tau(s) &= \prod_p \left(1 + \frac{\tau(p)}{p^s} + \frac{\tau(p^2)}{p^{2s}} + \cdots + \frac{\tau(p^k)}{p^{ks}} + \dots\right) \\ &= \prod_p \left(1 + \frac{2}{p^s} + \frac{3}{p^{2s}} + \cdots + \frac{k+1}{p^{ks}} + \dots\right) \\ &= \prod_p \frac{1}{(1-p^{-s})^2} \qquad \text{using } (1-x)^{-2} = \sum_{k=0}^{\infty} (k+1)x^k \\ &= \zeta(s)^2 \end{aligned}$$

(d) This can be done by the same method as (b) or (c) – a good exercise! But, given that we know the answer, we can work backwards more quickly:

$$\zeta(s-1)\zeta(s) = \left(\sum_k \frac{k}{k^s}\right) \cdot \left(\sum_\ell \frac{1}{\ell^s}\right) = \sum_n \frac{\sum_{k|n} k \cdot 1}{n^s} = D_\sigma(s),$$

using Prop. 21.2

□

22 Some Analytic Results about primes and the divisor function

22.1 The Prime Number Theorem

How frequent are the primes? At the end of the eighteenth century, Gauss and Legendre suggested giving up looking for a formula for the nth prime, and proposed instead estimating