M3P14 Number Theory

Lectured by Prof Toby Gee
Typeset by David Kurniadi Angdinata

Autumn 2018

Contents

Lecture 1
Friday
05/10/18

0 Introduction

Roughly speaking number theory is the study of the integers. More specifically, problems in number theory often have a lot to do with primes and divisibility, congruences, and include problems about the rational numbers. For example, solving equations in integers or in the rationals, such as $x^2 - 2y^2 = 1$, etc. We will be looking at problems that can be tackled by elementary means, but this does not mean easy. Also the statements of problems can be elementary without the solution being elementary, such as Fermat's last theorem, or even known, such as the twin prime conjecture. Sometimes we will state interesting things, like the prime number theorem, without proving them. Typically these will be things that we could prove if the course was much longer. We will start the course with a look at prime numbers and factorisation, a review of Euclid's algorithm and consequences, congruences, the structure of $(\mathbb{Z}/n\mathbb{Z})^\times$, RSA algorithm, and quadratic reciprocity. We will return to primes at the end, too. Typical questions here include the following.

- How do you tell if a number is prime?
- How many primes are there congruent to $a \mod b$ for given a, b?
- How many primes are there less than n?

A warning is that we will be using plenty of things from the compulsory first and second year algebra courses, about groups, rings, ideals, fields, Lagrange's theorem, the first isomorphism theorem, and so on. You may want to revise this material if you are not comfortable with it. The course is not based on any particular book, although some material, such as continued fractions, was drawn from the following.

1. A Baker, A concise introduction to the theory of numbers, 1984

Not everything we will do is in that book, though.

1 Euclid's algorithm and unique factorisation

1.1 Divisibility

Definition 1. If $a, b \in \mathbb{Z}$, we say that a **divides** b, $a \mid b$, if there exists $c \in \mathbb{Z}$ such that $b = ac$. If a does not divide b, write $a \nmid b$.

If $a \mid b$ and $a \mid c$ then $a \mid rb + sc$ for any $r, s \in \mathbb{Z}$.

Definition 2. The **greatest common divisor (gcd)** or **highest common factor (hcf)** of a, b is the largest positive integer dividing a and b. Write it as (a, b).

Example. $(-10, 15) = 5$.

Note that the ring $\mathbb{Z}$ is a principal ideal domain (PID). If $f_1, \ldots, f_n \in R$, write $(f_1, \ldots, f_n)$ for the ideal generated by the f_i. Then for $a, b \in \mathbb{Z}$, the ideal (a, b) is generated by the gcd (a, b), by Theorem 6 below.

Definition 3. $n \in \mathbb{Z}$ is **prime** if n has exactly two positive divisors, namely 1 and n.

Note that frequently when people talk about prime numbers they restrict to the positive case. If we write, let p be a prime number, then we will usually mean $p > 0$. Note that 1 is not prime.

1.2 Euclid's algorithm

Proposition 4. If $a, b \in \mathbb{Z}$, not both zero, then for any $n \in \mathbb{Z}$, $(a, b) = (a, b - na)$.

Proof. By definition, it is enough to show that if $r \mid a$ and $r \mid b$ then $r \mid a$ and $r \mid b - na$ and conversely. □

Theorem 5. Let $a, b \in \mathbb{Z}$ with $b > 0$. Then there exist unique $q, r \in \mathbb{Z}$ with $0 \leq r < b$ and $a = qb + r$.

Proof. Take $q = \lfloor a/b \rfloor$. By definition $0 \leq a/b - q < 1$, that is $0 \leq a - qb < b$, so take $r = a - qb$. Uniqueness is easy. □

Euclid's algorithm is as follows. Let $a, b \in \mathbb{Z}$ not both zero. Without loss of generality, $0 \leq b \leq a$.

Step 1. If $b = 0$, output a.

Step 2. Otherwise, replace (a, b) with (b, r) as in Theorem 5. Then go to step 1.

This algorithm terminates because $|a| + |b|$ decreases when we apply Step 2.

Example.

$$\begin{aligned}(120, 87) &= (87, 33) & 120 &= 87 + 33\\ &= (33, 21) & 87 &= 2\,(33) + 21\\ &= (21, 12) & 33 &= 21 + 12\\ &= (12, 9) & 21 &= 12 + 9\\ &= (9, 3) & 12 &= 9 + 3\\ &= (3, 0) & 9 &= 3\,(3) + 10.\end{aligned}$$

$$\begin{aligned}3 &= 12 - 9\\ &= 12 - (21 - 12)\\ &= 2\,(12) - 21\\ &= 2\,(33 - 21) - 21\\ &= 2\,(33) - 3\,(21)\\ &= 2\,(33) - 3\,(87 - 2\,(33))\\ &= 8\,(33) - 3\,(87)\\ &= 8\,(120 - 87) - 3\,(87)\\ &= 8\,(120) - 11\,(87)\,.\end{aligned}$$

Theorem 6. If $a, b \in \mathbb{Z}$, not both zero, then there exist $r, s \in \mathbb{Z}$ such that $(a, b) = ra + sb$.

Proof. Exercise: idea is to write (a_n, b_n) for the sequence of pairs in Euclid's algorithm, and use downwards induction on n. □

1.3 Unique factorisation

Proposition 7. Let $n, a, b \in \mathbb{Z}$ with $n \mid ab$ and $(n, a) = 1$. Then $n \mid b$.

Proof. Since $(n, a) = 1$, we can write $rn + sa = 1$, so $b = rnb + sab = n(rb) + (ab)s$, which is divisible by n, that is $n \mid b$. □

If $(n, a) = 1$, we say that n, a are **coprime**.

Lecture 2
Tuesday
09/10/18

Corollary 8. If p is prime and $p \mid ab$ then $p \mid a$ or $p \mid b$.

Proof. If $p \nmid a$ then $(p, a) = 1$, so Proposition 7 implies $p \mid b$. □

Proposition 9. If $(a, b) = 1$, and $a \mid n$ and $b \mid n$, then $ab \mid n$.

Proof. By 6, we can write $1 = ra + sb$ with $r, s \in \mathbb{Z}$. So $n = r(sa) + s(nb)$, which is divisible by ab, so $ab \mid n$. □

We say that $m_1, \ldots, m_n \in \mathbb{Z}$ are **pairwise coprime** if $(m_i, m_j) = 1$ for all $i \neq j$.

Corollary 10. If $m_1, \ldots, m_n$ are pairwise coprime and $m_i \mid N$ for all i then $m_1 \ldots m_n \mid N$.

Proof. Induction on n. $n = 2$ is Proposition 9. (Exercise) □

Proposition 11. Every $n \in \mathbb{Z}^\times$ can be written as $\pm p_1 \ldots p_r$ where p_i are prime, and r could be zero.

Proof. Use induction on $|n|$. The case $|n|$ is trivial, so suppose $|n| > 1$. Then either $|n|$ is prime, or $|n| = ab$ with $1 < a, b < |n|$, and by induction each of a, b is a product of primes. □

Theorem 12. Every $n \in \mathbb{Z}_{>0}$ can be written as $\pm p_1 \ldots p_r$ where p_i are prime and are uniquely determined up to ordering.

Proof. Existence is Proposition 11. For uniqueness, suppose that

$$n = p_1 \ldots p_r = q_1 \ldots q_s,$$

with p_i, q_i prime. Then without loss of generality suppose $r, s \geq 1$. Then $p_1 \mid p_1 \ldots p_r$, so $p_1 \mid q_1 \ldots q_s$. By Corollary 8, either $p_1 \mid q_1$ or $p_1 \mid q_2 \ldots q_s$. Proceeding inductively, eventually $p_1 \mid q_i$ for some i. Since q_i is prime this means $p_1 = q_i$. We then have

$$p_2 \ldots p_r = q_1 \ldots q_i \ldots q_s.$$

Since this product is smaller than n, by the inductive hypothesis we must have $r - 1 = s - 1$ and the p_i, except p_1, are a rearrangement of the q_i, except q_i. □

1.4 Linear diophantine equations

Let $a, b, c \in \mathbb{Z}^\times$. Want to solve $ax + by = c$ with $x, y \in \mathbb{Z}$.

Example. $2x + 6y = 3$ has no solutions.

In general, there are no solutions if $(a,b) \nmid c$. Suppose that $(a,b) \mid c$. Then

$$ax + by = c \qquad \iff \qquad \frac{a}{(a,b)}x + \frac{b}{(a,b)}y = \frac{c}{(a,b)}.$$

By Theorem 6, since

$$\left(\frac{a}{(a,b)}, \frac{b}{(a,b)}\right) = 1,$$

we can find $r, s \in \mathbb{Z}$ with

$$\frac{a}{(a,b)}r + \frac{b}{(a,b)}s = 1,$$

so

$$\frac{a}{(a,b)}\left(\frac{rc}{(a,b)}\right) + \frac{b}{(a,b)}\left(\frac{sc}{(a,b)}\right) = \frac{c}{(a,b)}.$$

So

$$x = \frac{rc}{(a,b)}, \qquad y = \frac{sc}{(a,b)}$$

is a solution. X, Y is another solution if and only if

$$\frac{a}{(a,b)}X + \frac{b}{(a,b)}Y = \frac{a}{(a,b)}x + \frac{b}{(a,b)}y,$$

if and only if

$$\frac{a}{(a,b)}(X - x) = \frac{b}{(a,b)}(y - Y).$$

For this to hold, we need

$$\frac{a}{(a,b)} \,\Big|\, y - Y, \qquad \frac{b}{(a,b)} \,\Big|\, X - x.$$

See that the solutions are exactly

$$X = x + \frac{nb}{(a,b)}, \qquad Y = y - \frac{na}{(a,b)}.$$

2 Congruences and modular arithmetic

2.1 Congruences

Definition 13. Let $n \in \mathbb{Z}^\times$, usually $n > 0$. Let $a, b \in \mathbb{Z}$. We say that a is **congruent to** $b \mod n$ if and only if $n \mid a - b$. Write $a \equiv b \mod n$.

$\equiv$ is an equivalence relation, and we write $\mathbb{Z}/n\mathbb{Z}$ for the equivalence classes, which is a ring.

Example. If $a \equiv b \mod n$, $c \equiv d \mod n$, then $a + c \equiv b + d \mod n$ and $ac \equiv bd \mod n$.

If $a \in \mathbb{Z}$, we sometimes write $\overline{a}$ for the image of a in $\mathbb{Z}/n\mathbb{Z}$.

Example. If $n = 12$, then $\overline{25} = \overline{1}$.

So every element of $\mathbb{Z}/n\mathbb{Z}$ is equal to $\overline{r}$ for some unique $r \in \{0, \dots, n-1\}$. We often write

$$\frac{\mathbb{Z}}{n\mathbb{Z}} = \{0, \dots, n-1\}.$$

Example. If $n = 6$, we could write $3 + 4 = 1$ and $3 \times 4 = 0$.

Let R be a commutative ring with unity. Then a **unit** of R is an element x such that there exists $y \in R$ with $xy = 1$. Write $R^\times$ for the set of units in R. This is a group under multiplication.

Example.

- $\mathbb{Z}^\times = \{\pm 1\}$.
- $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\} = \{x \in \mathbb{Q} \mid x \neq 0\}$.

We want to understand $(\mathbb{Z}/n\mathbb{Z})^\times$. Which elements of $\{0, \dots, n-1\}$ are in $(\mathbb{Z}/n\mathbb{Z})^\times$? If $r \in \mathbb{Z}$ and $r \in (\mathbb{Z}/n\mathbb{Z})^\times$ then there exists $s \in \mathbb{Z}$ such that $rs \equiv 1 \mod n$. This implies that $(r, n) = 1$. Conversely, if $(r, n) = 1$, then there exist $x, y \in \mathbb{Z}$ such that $rx + ny = 1$, that is $\overline{rx} = 1$, that is $\overline{r}$ is a unit. So

$$\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^\times = \{0 \le i < n \mid (i, n) = 1\}.$$

Example. If p is a prime, then

$$\left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times = \{1, \dots, p-1\}.$$

So $\mathbb{Z}/p\mathbb{Z}$ is a ring with the property that every non-zero element has a multiplicative inverse, so it is a field. Another equivalent way to see this is to check that $p\mathbb{Z}$ is a maximal ideal of $\mathbb{Z}$.

Thus every non-zero congruence class mod p is a unit,

Lecture 3
Wednesday
10/10/18

2.2 Linear congruence equations

Consider the question of solving $ax \equiv b \mod c$ for $a, b, c, x \in \mathbb{Z}$. This is equivalent to solving $ax + cy = b$ for $y \in \mathbb{Z}$. We saw yesterday that this has solutions if and only if $(a, c) \mid b$. Furthermore, there is a unique solution mod $c/(a, c)$, because all the solutions are obtained by adding multiples of $c/(a, c)$ to our given x, and subtracting the corresponding multiple of $a/(a, c)$ from y. This implies that there are (a, c) solutions to the original congruence mod c. If x_0 is one solution, the others are

$$x_0 + \frac{cj}{(a, c)}, \qquad 0 \le j < (a, c).$$

In particular, if $(a, c) = 1$ then there is a unique solution to $ax \equiv b \mod c$. Indeed $a \in (\mathbb{Z}/c\mathbb{Z})^\times$, so it has an inverse a^{-1}, and $x \equiv a^{-1}b \mod c$ is the unique solution.

Example.

- $2x \equiv 3 \mod 6$ has no solutions as $(2, 6) = 2 \nmid 3$.
- $2x \equiv 4 \mod 6$ if and only if $x \equiv 2 \mod 3$ has solutions $x \equiv 2 \mod 6$ and $x \equiv 5 \mod 6$.

2.3 Chinese remainder theorem

Theorem 14 (Chinese remainder theorem)**.** Let $m_1, \dots, m_n \in \mathbb{Z}_{>0}$ be pairwise coprime. Then the natural map

$$\frac{\mathbb{Z}}{m_1 \dots m_n \mathbb{Z}} \to \frac{\mathbb{Z}}{m_1\mathbb{Z}} \times \dots \times \frac{\mathbb{Z}}{m_n\mathbb{Z}}$$

is an isomorphism of rings. Consequently,

$$\left(\frac{\mathbb{Z}}{m_1 \dots m_n \mathbb{Z}}\right)^\times \to \left(\frac{\mathbb{Z}}{m_1\mathbb{Z}}\right)^\times \times \dots \times \left(\frac{\mathbb{Z}}{m_n\mathbb{Z}}\right)^\times$$

is an isomorphism of abelian groups.

Remark that this is false without the assumption that m_i pairwise coprime, such as $m_1 = m_2 = 2$.

Proof. The map

$$\frac{\mathbb{Z}}{m_1 \dots m_n \mathbb{Z}} \to \frac{\mathbb{Z}}{m_1\mathbb{Z}} \times \dots \times \frac{\mathbb{Z}}{m_n\mathbb{Z}}$$

is a ring homomorphism between two rings of order, or cardinality, $m_1 \dots m_n$. So to show that it is an isomorphism, it is enough to show that it is an injection, so we only need to check that the kernel is zero. So we need to know that if $m_i \mid N$ for all i, then $m_1 \dots m_n \mid N$. This is Corollary 10. For the second part, just use that if R, S are rings, then

$$(R \times S)^\times \cong R^\times \times S^\times.$$

□

The first part says that given any $a_i \in \mathbb{Z}$, there is a unique $x \mod m_1 \dots m_n$ with $x \equiv a_i \mod m_i$. Write

$$M = m_1 \dots m_n, \qquad M_i = \frac{M}{m_i}.$$

Choose q_i such that $q_i M_i \equiv 1 \mod m_i$, using $(M_i, m_i) = 1$ because $(m_j, m_i) = 1$ for all $j \neq i$. Then take

$$x = a_1 q_1 M_1 + \dots + a_n q_n M_n.$$

Then

$$x \equiv a_i q_i M_i \equiv a_i \mod m_i.$$

3 The structure of $(\mathbb{Z}/n\mathbb{Z})^\times$

3.1 The Euler Φ function

Let $\Phi(n)$ be the order of $(\mathbb{Z}/n\mathbb{Z})^\times$, that is

$$\Phi(n) = \#\{1 \le i < n \mid (i,n) = 1\}.$$

Example. If p is prime, $\Phi(p) = p - 1$.

Φ is called **Euler's Φ function**.

Definition 15. Let f be a function on the positive integers. Say that f is **strongly multiplicative** if

$$f(mn) = f(m) f(n)$$

for all m, n. Say f is **multiplicative** if this holds whenever $(m,n) = 1$.

Φ is multiplicative by Theorem 14 because if $(m,n) = 1$ then

$$\frac{\mathbb{Z}}{m_1 \dots m_n \mathbb{Z}} \to \frac{\mathbb{Z}}{m_1\mathbb{Z}} \times \cdots \times \frac{\mathbb{Z}}{m_n\mathbb{Z}}.$$

Φ is not strongly multiplicative, since

$$\Phi(4) = 2 \ne 1 = \Phi(2)\Phi(2).$$

If p is prime then

$$\Phi(p^a) = \#\{1 \le i < p^a \mid (i,p^a) = 1\} = \#\{1 \le i < p^a \mid p \nmid i\} = p^a - p^{a-1} = p^a\left(1 - \frac{1}{p}\right).$$

If $n = \prod_i p_i^{a_i}$, then

$$\Phi(n) = \prod_i \Phi(p_i^{a_i}) = \prod_i p_i^{a_i}\left(1 - \frac{1}{p_i}\right) = n\prod_i\left(1 - \frac{1}{p_i}\right) = n\prod_{p \mid n}\left(1 - \frac{1}{p}\right).$$

3.2 Euler's theorem

Theorem 16 (Euler's theorem)**.** If $(a,n) = 1$, then

$$a^{\Phi(n)} \equiv 1 \mod n.$$

Proof. This is equivalent to showing that $\overline{a}^{\Phi(n)} = 1$ in $(\mathbb{Z}/n\mathbb{Z})^\times$. This is a group of order $\Phi(n)$, so this is immediate from Lagrange's theorem. □

Corollary 17 (Fermat's little theorem)**.** If p is prime and $p \nmid a$, then

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

Proof. Theorem 16 with $n = p$, so $\Phi(n) = p - 1$. □

Next is to understand the structure of $(\mathbb{Z}/n\mathbb{Z})^\times$. By Theorem 14, it is enough to study the case that n is a prime power. We will begin by considering the case that n is prime.

Example. Let $n = 5$.

$$\left(\frac{\mathbb{Z}}{5\mathbb{Z}}\right)^\times = \{1,2,3,4\}.$$

This has order four. So it is either cyclic of order four or a product of two cyclic groups of order two.

$$2^2 = 4, \qquad 2^3 = 3, \qquad 2^4 = 1$$

in $(\mathbb{Z}/5\mathbb{Z})^\times$. So $(\mathbb{Z}/5\mathbb{Z})^\times$ is cyclic of order four.

Next is $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic of order $p-1$ for any prime p.

Lecture 4
Friday
12/10/18

Definition 18. If G is a group and $g \in G$ is an element, the **order** of g is the least $a \geq 1$ such that $g^a = 1$. In particular, if $(g, n) = 1$, then we write $ord_n(g)$ for the order of g in $(\mathbb{Z}/n\mathbb{Z})^\times$, **the order of** $g \mod n$.

Proposition 19. If G is a group and g is an element of order a, then

$$g^n = 1 \qquad \Longleftrightarrow \qquad a \mid n.$$

Proof.

$\Longleftarrow$ If $n = ab$ then $g^n = (g^a)^b = 1^b = 1$.

$\Longrightarrow$ Write $n = ab + r$ with $0 \leq r < a$. Then $g^r = 1$ and since $r < a$ we have $r = 0$.

□

In particular, if $(g, n) = 1$, then $g^{\Phi(n)} = 1$, by Euler's theorem, so Proposition 19 gives $ord_n(g) \mid \Phi(n)$. We want to prove that if p is prime, then $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic. Equivalently, we need to show that there exists g such that $ord_p(g) = \Phi(p) = p-1$. We will do this by counting the number of elements of each order. Key point is that $\mathbb{Z}/p\mathbb{Z}$ is a field. For any $d \geq 1$, the elements of $(\mathbb{Z}/p\mathbb{Z})^\times$ of order dividing d are exactly the roots of the $X^d - 1$ in $\mathbb{Z}/p\mathbb{Z}$, by Proposition 19.

Example. The equation $X^2 = 1$ has exactly two solutions mod p for any prime p, namely ± 1, but it can have more mod n if n is composite. For example, if $n = 15$, then $4, 11$ are also solutions.

$$X^2 - 1 \equiv 0 \mod n \qquad \Longleftrightarrow \qquad n \mid (X+1)(X-1),$$

for example, $15 \mid (4+1)(4-1)$.

Definition 20. $g \in \mathbb{Z}$ with $(g, p) = 1$ is a **primitive root** if $ord_p(g) = p-1$, that is $(\mathbb{Z}/p\mathbb{Z})^\times = \langle g \rangle$.

Lemma 21. Let R be a commutative ring, and let $P(X) \in R[X]$. If $\alpha \in R$ has $P(\alpha) = 0$, then there exists $Q(X) \in R[X]$ such that $P(X) = (X - \alpha) Q(X)$.

Example. If $R = \mathbb{Z}/15\mathbb{Z}$,

$$X^2 - 1 = (X+1)(X-1) = (X+4)(X-4).$$

Proof. Induction on $\deg(P)$. $\deg(P) = 0$ is obvious. Let $\deg(P) = d$, assume the result holds for degree at most $d-1$. Let

$$P(X) = cX^d + \ldots, \qquad S(X) = P(X) - cX^{d-1}(X - \alpha).$$

Then $S(X)$ has degree at most $d-1$. Also $S(\alpha) = 0$. By induction, we can write $S(X) = (X - \alpha) R(X)$. Set $Q(X) = cX^{d-1} + R[X]$. Then

$$(X - \alpha) Q(X) = cX^{d-1}(X - \alpha) + S(X) = P(X).$$

□

Theorem 22. Let F be a field. Let $P(X)$ be a polynomial in $F[X]$. Then $P(X)$ has at most d distinct roots in F.

Proof. Induction on $d = \deg(P)$. $d = 1$ is obvious. If P has no roots, then we are done. Otherwise, let α be a root. By Lemma 21,

$$P(X) = (X - \alpha) Q(X),$$

$Q(X)$ has degree $d-1$, so we are done by induction. □

Corollary 23. Let d be any divisor of $p-1$. Then there are exactly d elements of $(\mathbb{Z}/p\mathbb{Z})^\times$ of order dividing d.

Proof. We have to show that $X^d - 1$ has exactly d roots in $\mathbb{Z}/p\mathbb{Z}$. $X^{p-1} - 1$ has exactly $p-1$ roots, by Fermat's little theorem. Since $d \mid p-1$, we can write

$$X^{p-1} - 1 = \left(X^d - 1\right)\left(\left(X^d\right)^{\frac{p-1}{d}-1} + \cdots + 1\right) = \left(X^d - 1\right) Q\left(X\right), \qquad \deg\left(Q\right) = p - 1 - d.$$

$X^{p-1} - 1$ has exactly $p-1$ roots, $X^d - 1$ has at most d roots, and $Q\left(X\right)$ has at most $p-1-d$ roots, by Theorem 22. So $X^d - 1$ has exactly d roots. □

Example. Let $p = 7$. There are

- one element of order one,
- two elements of order dividing two, so one element of order two,
- three elements of order dividing three, so two elements of order three, and
- six elements of order dividing six, so two elements of order six.

Lemma 24. For any $n \geq 1$, we have

$$\sum_{d \mid n} \Phi\left(d\right) = n.$$

Proof. For each $d \mid n$, the elements of $\{1, \ldots, n\}$ with $(i, n) = n/d$ are exactly those of the form $i = (n/d)\, j$ with $1 \leq j \leq d$ and $(j, d) = 1$. There are exactly $\Phi\left(d\right)$ such elements. Since the n/d run over all the divisors of n, we are done. □

Theorem 25. Let p be prime, and let $d \mid p-1$. Then there are exactly $\Phi\left(d\right)$ elements of $\left(\mathbb{Z}/p\mathbb{Z}\right)^\times$ of order d. In particular, there are $\Phi\left(p-1\right)$ primitive roots, and $\left(\mathbb{Z}/p\mathbb{Z}\right)^\times$ is cyclic.

Proof. Induction on d. $d = 1$ is obvious. Assume the result holds for all $d' \mid d$, $d' \neq d$. Then by Lemma 24,

$$\Phi\left(d\right) = d - \sum_{d' \mid d,\ d' \neq d} \Phi\left(d'\right).$$

Now use inductive hypothesis and Corollary 23. □

Lecture 5
Tuesday
16/10/18

Proposition 26. Let p be an odd prime and $n \geq 1$. Then $\left(\mathbb{Z}/p^n\mathbb{Z}\right)^\times$ is cyclic.

Proof. Consider three cases.

$n = 1$ Theorem 25.

$n = 2$ Let g be a primitive root mod p. Claim that either $g^{p-1} \not\equiv 1 \mod p^2$, and g is a generator for $\left(\mathbb{Z}/p^2\mathbb{Z}\right)^\times$, or $g^{p-1} \equiv 1 \mod p^2$, and $g + p$ is a generator for $\left(\mathbb{Z}/p^2\mathbb{Z}\right)^\times$. Either way, $\left(\mathbb{Z}/p^2\mathbb{Z}\right)^\times$ is cyclic. Suppose firstly that $g^{p-1} \not\equiv 1 \mod p^2$.

$$\#\left(\frac{\mathbb{Z}}{p^2\mathbb{Z}}\right)^\times = \Phi\left(p^2\right) = p\left(p-1\right).$$

So $ord_{p^2}\left(g\right) \mid p\left(p-1\right)$. On the other hand, $g^{ord_{p^2}(g)} \equiv 1 \mod p^2$ gives $g^{ord_{p^2}(g)} \equiv 1 \mod p$, so $p-1 \mid ord_{p^2}\left(g\right)$, because $ord_p\left(g\right) = p-1$ by assumption. But $ord_{p^2}\left(g\right) \neq p-1$, as $g^{p-1} \not\equiv 1 \mod p^2$. So $ord_{p^2}\left(g\right) = p\left(p-1\right)$ as required. Suppose now that $g^{p-1} \equiv 1 \mod p$. It suffices to show that $(g+p)^{p-1} \not\equiv 1 \mod p^2$, as we can then apply the analysis above with $g + p$ in place of g. By the binomial theorem,

$$(g+p)^{p-1} \equiv g^{p-1} + (p-1)\, g^{p-2} p \equiv 1 + (p-1)\, g^{p-2} p \mod p^2.$$

Since $p \nmid (p-1)\, g^{p-2}$, $(g+p)^{p-1} \not\equiv 1 \mod p^2$, as required.

$n \geq 2$ It suffices to show that if $ord_{p^2}(g) = p(p-1)$, then $ord_{p^n}(g) = p^{n-1}(p-1)$. We do this by induction on n. So assume that $ord_{p^n}(g) = p^{n-1}(p-1)$. Then

$$p^{n-1}(p-1) = ord_{p^n}(g) \mid ord_{p^{n+1}}(g) \mid \Phi\left(p^{n+1}\right) = p^n(p-1).$$

So either $ord_{p^{n+1}}(g) = p^n(p-1)$, or $ord_{p^{n+1}}(g) = p^{n-1}(p-1)$. So we need to show that

$$g^{p^{n-1}(p-1)} \not\equiv 1 \mod p^{n+1}.$$

To do this, consider $g^{p^{n-2}(p-1)} \mod p^{n-1}$ and $g^{p^{n-2}(p-1)} \mod p^n$. Since $\Phi\left(p^{n-1}\right) = p^{n-2}(p-1)$,

$$g^{p^{n-2}(p-1)} \equiv 1 \mod p^{n-1},$$

by Euler's theorem. Write

$$g^{p^{n-2}(p-1)} = 1 + p^{n-1}t.$$

Since $ord_{p^n}(g) = p^{n-1}(p-1)$ by assumption, $g^{p^{n-2}(p-1)} \not\equiv 1 \mod p^n$, that is $p \nmid t$. Then

$$g^{p^{n-1}(p-1)} = \left(g^{p^{n-2}(p-1)}\right)^p = \left(1 + p^{n-1}t\right)^p \equiv 1 + p^n t + \binom{p}{2} p^{2(n-1)}t^2 + \cdots + p^{p(n-1)}t^p \mod p^{n+1},$$

Since $r(n-1) \geq n+1$ if and only if $(r-1)n \geq r+1$ and $p > 2$,

$$p \left| \binom{p}{2} \right. \qquad \Longrightarrow \qquad p^{n+1} \left| p^{2n-1} = p^{2(n-1)+1} \right| \binom{p}{2} p^{2(n-1)}.$$

So $g^{p^{n-1}(p-1)} \equiv 1 + p^n t \not\equiv 1 \mod p^{n+1}$, because $p \nmid t$.

□

Example.

- $(\mathbb{Z}/2\mathbb{Z})^\times = \{1\}$.
- $(\mathbb{Z}/4\mathbb{Z})^\times = \{1, 3\}$ is cyclic of order two, with 3 as a generator.
- $(\mathbb{Z}/8\mathbb{Z})^\times = \{1, 3, 5, 7\}$ is not cyclic.

$$1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \mod 8,$$

so every element has order two.

Lemma 27. For $n \geq 0$ we have $5^{2^n} \equiv 1 + 2^{n+2} \mod 2^{n+3}$.

Proof. Induction on n. $n = 0$ is obvious. Assume that $5^{2^n} = 1 + 2^{n+2}t$ with t odd. Then

$$5^{2^{n+1}} = \left(1 + 2^{n+1}t\right)^2 = 1 + 2^{n+3}t + 2^{2(n+2)}t^2 = 1 + 2^{n+3}\left(t + 2^{n+1}t^2\right),$$

where $t + 2^{n+1}t^2$ is odd. □

Proposition 28. If $n \geq 2$ then there is an isomorphism

$$\left(\frac{\mathbb{Z}}{2^n\mathbb{Z}}\right)^\times \to \frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2^{n-2}\mathbb{Z}}.$$

In particular, if $n \geq 3$, then $(\mathbb{Z}/2^n\mathbb{Z})^\times$ is not cyclic.

Proof. Let $\langle g \rangle$ denote the group

$$\left\{1, \ldots, g^{ord(g)-1}\right\}$$

generated by g. Consider the natural map

$$\langle -1 \rangle \times \langle 5 \rangle \to \left(\frac{\mathbb{Z}}{2^n\mathbb{Z}}\right)^{\times}.$$

This is injective, because if $\pm 1\,(5)^s \equiv 1 \mod 2^n$ then in particular $\pm 1\,(5)^s \equiv 1 \mod 4$ so $\pm 1 \equiv 1 \mod 4$, so we must have $5^s \equiv 1 \mod 2^n$, that is $5^s = 1$ in $\langle 5 \rangle$. $\langle -1 \rangle$ has order 2 and $\langle 5 \rangle$ has order $ord_{2^n}(5) = 2^{n-2}$ by Lemma 27. So $\langle -1 \rangle \times \langle 5 \rangle$ has order

$$2\left(2^{n-2}\right) = 2^{n-1} = \Phi\left(2^n\right) = \#\left(\frac{\mathbb{Z}}{2^n\mathbb{Z}}\right)^{\times}.$$

So the map

$$\langle -1 \rangle \times \langle 5 \rangle \to \left(\frac{\mathbb{Z}}{2^n\mathbb{Z}}\right)^{\times}$$

is an injection of groups of the same order, so it is a bijection. □

Theorem 29. $(\mathbb{Z}/n\mathbb{Z})^{\times}$ is cyclic if and only if either

- $n = 1, 2, 4$,
- $n = p^r$ for $p > 2$ prime and $r \geq 1$, or
- $n = 2p^r$ for $p > 2$ prime and $r \geq 1$.

Lecture 6
Wednesday
17/10/18

Primitive roots are generators of $(\mathbb{Z}/n\mathbb{Z})^{\times}$. Find them in practice by guessing small values of g, and seeing if g is a generator. There are $\Phi(p-1)$ primitive roots, which means that you have a high probability of success. Could work out $1, \ldots, g^{p-2}$ and check these are distinct. This would be inefficient. Better is to check for some prime $q \mid p-1$ whether $g^{(p-1)/q} = 1$ or not. This works, because if $g^{(p-1)/q} = 1$ then g is not a primitive root, while if $g^{(p-1)/q} \neq 1$ then $ord_p(g) \mid p-1$ and $ord_p(g) \nmid (p-1)/q$. If this holds for all $q \mid p-1$, then $ord_p(g) = p-1$, because otherwise it would be a proper divisor, and so would divide $(p-1)/q$ for some prime $q \mid p-1$.

Example. Let $p = 31$, so $p - 1 = 30 = (2)(3)(5)$. g is a primitive root if and only if $g^{15} \neq 1$, $g^{10} \neq 1$, $g^6 \neq 1$.

- Is 2 a primitive root? $2^2 = 4$, $2^4 = 16$, $2^6 = 2$, but $2^{10} = 2^{15} = 1$ because $2^5 = 32 = 1$.
- How about 3? $3^2 = 9$, $3^4 = 19$, $3^6 = 16$, $3^8 = 20$, $3^{10} = 25$, $3^{15} = 30$. So 3 is a primitive root mod 31.

4 Primality testing and factorisation

Idea is that testing whether $n \in \mathbb{Z}$ is prime is easy. Factoring n is expected to be hard. Easy here means that there is an algorithm to check whether n is prime or not which runs in time polynomial in $\log n$. It is known that a deterministic algorithm exists to do this, the Agrawal-Kayal-Saxena (AKS) algorithm, in 2005. We will see an algorithm that runs faster than this in practice. On the other hand, for factoring there are algorithms which are better than exponential in $\log n$, but there is nothing close to polynomial time, and the general expectation is that no such algorithm should exist.

4.1 Factorisation

How do we factor three digit numbers, or small four digit numbers, say at most 400 if we wanted to factor with paper or calculator? If $n \leq 400$ and n is composite, then has a prime factor at most $\sqrt{400} = 20$, since if $d \mid n$ then $d(n/d) = n$, so either $d \leq \sqrt{n}$ or $n/d \leq \sqrt{n}$. So you only have to be able to check for divisibility by $2, 3, 5, 7, 11, 13, 17, 19$.

$2, 5$ Checking for divisibility is easy, by just looking at the last digit.

$3, 9, 11$ Use that $10 \equiv 1 \mod 3$ and $10 \equiv -1 \mod 3$. So

$$\sum_i a_i 10^i \equiv \sum_i a_i \mod 3, \qquad \sum_i a_i 10^i \equiv \sum_i a_i (-1)^i \mod 11.$$

So you can check divisibility by 3, or 9, by checking for the sum of the digits, and 11 by taking the alternating sum.

7 $10x + y \equiv 0 \mod 7$ if and only if $-2(10x + y) \equiv 0 \mod 7$, if and only if $x - 2y \equiv 0 \mod 7$.

$13, 17, 19$ There are no good tests. If $n \leq 400$ and n is not divisible by $2, 3, 5, 7, 11$, then the smallest prime factor of n is at least 13. Since $13^3 > 400$, it can have at most two prime factors. So if you want to factor numbers at most 400, you only have to remember a short list

$$13^2, \quad 13(17), \quad 13(19), \quad 13(23), \quad 13(29), \quad 17^2, \quad 17(19), \quad 17(23), \quad 19^2.$$

Example.

- $143 \equiv 1 - 4 + 3 \equiv 0 \mod 11$.
- $144 \equiv 1 + 4 + 4 \equiv 0 \mod 9$.
- $154 \equiv 15 - 2(4) = 7 \equiv 0 \mod 7$.

Lecture 7
Friday
19/10/18

Factor four digit numbers by an algorithm due to Fermat. Idea is to first check for small prime factors by hand, say $p = 2, \ldots 19$. If n is composite and does not have any small factors, then the prime factors of n should be close to $\sqrt{n}$. If $n = ab$ for a, b odd and $a \leq b$, then

$$n = ab = \left(\frac{a+b}{2}\right)^2 - \left(\frac{b-a}{2}\right)^2, \qquad \left(\frac{a+b}{2}\right)^2 - n = \left(\frac{b-a}{2}\right)^2.$$

If you know $(a+b)/2$ and $(b-a)/2$, you can recover a, b. So take m such that $m^2 \leq n < (m+1)^2$. If $n = m^2$, done. Otherwise check if $(m+i)^2 - n$ is a square for increasing i.

Example. Let $n = 6077$. $77^2 < 6077 < 78^2$, so

$$\begin{aligned} 78^2 - 6077 &= 7, \\ 79^2 - 6077 &= 164, \\ 80^2 - 6077 &= 323, \\ 81^2 - 6077 &= 484 = 22^2. \end{aligned}$$

Thus $6077 = 81^2 - 22^2 = (103)(59)$.

There exist algorithms for factoring n which run in better than exponential time in $\log n$, such as the quadratic sieve and the general number field sieve. The following is the **quadratic sieve**.

Example. Let $n = 1649$. $40^2 < 1649 < 41^2$, so

$$\begin{aligned} 41^2 - 1649 &= 32 = 2^5, \\ 42^2 - 1649 &= 115, \\ 43^2 - 1649 &= 200 = (2)^3 (5)^2 . \end{aligned}$$

$41^2 \equiv 2^5 \mod 1649$ and $43^2 \equiv (2)^3 (5)^2 \mod 1649$, so

$$80^2 \equiv (41)^2 (43)^2 = 1763^2 \equiv 114^2 \mod 1649.$$

Then

$$0 \equiv 114^2 - 80^2 = (194)(34) = (2)^2 (17)(97) \mod 1649.$$

In fact, $1649 = (17)(97)$. Better for this last step would be to have computed $(194, 1649) = 97$ and $(34, 1649) = 17$. Can do this quickly using Euclid's algorithm.

To make this into an efficient algorithm, need to have a way given $x_1, \ldots, x_r$ to find a subset whose product is a square. If we know the prime factorisation for the x_i, we can write

$$x_i = p_1^{a_{i1}} \ldots p_k^{a_{ik}}.$$

Want to choose $\epsilon_i \in \{0, 1\}$ such that $\prod_{i=1}^r x_i^{\epsilon_i}$ is a square. Equivalently, for each j, want the exponent of p_j to be even, that is

$$\sum_{i=1}^r \epsilon_i a_{ij} \equiv 0 \mod 2.$$

Example.

$$x_1 = 2^5, \qquad x_2 = (5)(23), \qquad x_3 = (2)^3 (5)^2, \qquad p_1 = 2, \qquad p_2 = 5, \qquad p_3 = 23.$$

Ignore all numbers with a large prime factor, so here ignore 23.

$$\begin{pmatrix} \epsilon_1 & \epsilon_2 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 3 & 2 \end{pmatrix} \equiv \begin{pmatrix} 0 & 0 \end{pmatrix} \mod 2 \qquad \Longleftrightarrow \qquad \begin{pmatrix} \epsilon_1 & \epsilon_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

in $\mathbb{Z}/2\mathbb{Z}$, a field, $\mathbb{F}_2$, that is $\epsilon_1 + \epsilon_2 = 0$, so $\epsilon_1 = \epsilon_2 = 1$.

This step, solving linear equations in $\mathbb{Z}/2\mathbb{Z}$, can be done efficiently. The remaining difficulty is to find a supply of $m \in \mathbb{Z}$ such that $m^2 - n$ has only small prime factors. Idea is that if we fix a list of small primes to start with, we get congruence conditions on m. It turns out that there is a straightforward algorithm for solving $m^2 \equiv n \mod p$. This gives two possible values for $m \mod p$. If you do this for lots of primes p, you get a supply of congruence conditions for m, so you can eliminate ever considering m such that $m^2 - n$ has large prime factors.

Example. $m^2 = 1649 \equiv 2 \mod 3$ has no solutions.

Lecture 8
Tuesday
23/10/18

4.2 Testing primality

Euler's theorem states that if $(a, n) = 1$ then $a^{\Phi(n)} \equiv 1 \mod n$. In particular if p is prime then $a^{p-1} \equiv 1 \mod p$ for all $1 \leq a \leq p - 1$. In particular, if $2^{n-1} \equiv 1 \mod n$, then n cannot be prime. Problem is that there exists n composite such that $a^{n-1} \equiv 1 \mod n$ for all $(a, n) = 1$, the **Carmichael numbers**. It is known that infinitely many of these exist. **Miller-Rabin test** is a test for whether odd $n \in \mathbb{Z}$ is prime or not. Today let $n \equiv 3 \mod 4$. Example sheet is $n \equiv 1 \mod 4$.

Lemma 30. Let $n > 1$ be congruent to 3 mod 4. Then n is prime if and only if for all $(a, n) = 1$,

$$a^{\frac{n-1}{2}} \equiv \pm 1 \mod n.$$

Proof.

- If n is prime, then $a^{n-1} \equiv 1 \mod n$ by Fermat's little theorem, so

  $$\left(a^{(n-1)/2}\right)^2 \equiv 1 \mod n,$$

  so $a^{(n-1)/2} \equiv \pm 1 \mod n$.

- Suppose firstly that $n = p^k$ with p prime, and $k \geq 2$. Try $a = 1 + p$. Then

  $$(1+p)^{\frac{n-1}{2}} \equiv 1 + \left(\frac{n-1}{2}\right) p \mod p^2,$$

  by the binomial theorem. If $(1+p)^{(n-1)/2} \equiv \pm 1 \mod p^k = n$, then $(1+p)^{(n-1)/2} \equiv \pm 1 \mod p, p^2$ gives

  $$\pm 1 \equiv (1+p)^{\frac{n-1}{2}} \equiv 1 + \left(\frac{n-1}{2}\right) p \equiv 1 \mod p \qquad \Longrightarrow \qquad 1 \equiv 1 + \left(\frac{n-1}{2}\right) p \mod p^2,$$

  then $p \mid (n-1)/2$, so $p \mid n-1$. But $p \mid n$, a contradiction.

- The remaining case is that n is composite but not a power of a prime. Write $n = rs$, for $r, s > 1$, and odd, and $(r, s) = 1$. By the Chinese remainder theorem,

  $$\frac{\mathbb{Z}}{n\mathbb{Z}} \cong \frac{\mathbb{Z}}{r\mathbb{Z}} \times \frac{\mathbb{Z}}{s\mathbb{Z}}.$$

  Choose a such that

  $$a \equiv -1 \mod r, \qquad a \equiv 1 \mod s.$$

  Then $(a, r) = (a, s) = 1$, so $(a, n) = 1$. Since $n \equiv 3 \mod 4$, $(n-1)/2$ is odd, so

  $$a^{(n-1)/2} \equiv -1 \mod r, \qquad a^{(n-1)/2} \equiv 1 \mod s.$$

  So $a^{(n-1)/2} \not\equiv \pm 1 \mod n$.

□

Lemma 31. Suppose that $n \equiv 3 \mod 4$ is composite. Then the set of $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ which satisfy

$$a^{\frac{n-1}{2}} \equiv \pm 1 \mod n$$

is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$.

Proof. Certainly $1^{(n-1)/2} \equiv 1 \mod n$. If $a^{(n-1)/2} \equiv \pm 1 \mod n$ and $b^{(n-1)/2} \equiv \pm 1 \mod n$,

$$(ab)^{\frac{n-1}{2}} \equiv a^{\frac{n-1}{2}} b^{\frac{n-1}{2}} \equiv (\pm 1)(\pm 1) \equiv \pm 1 \mod n, \qquad \left(a^{-1}\right)^{\frac{n-1}{2}} \equiv \left(a^{\frac{n-1}{2}}\right)^{-1} \equiv (\pm 1)^{-1} \equiv \pm 1 \mod n.$$

So this set is a subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$. It is a proper subgroup by Lemma 30. □

Corollary 32. At most half the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$ satisfy

$$a^{\frac{n-1}{2}} \equiv \pm 1 \mod n.$$

Proof. The set of such elements is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$ by Lemma 31, so it has index at least two. □

In fact, with a bit more work, you can improve this to show that at least $3/4$ of the numbers $1 \leq a \leq n-1$ satisfy

$$a^{(n-1)/2} \not\equiv \pm 1 \mod n.$$

So if you randomly choose numbers $1 \leq a \leq n-1$ x times, and n is composite, the probability that you find some a with

$$a^{(n-1)/2} \not\equiv \pm 1 \mod n$$

is at least $1 - (1/4)^x$. This gives a probabilistic algorithm to check if n is prime in polynomial time. If you assume generalised Riemann hypothesis (GRH) you can find some $1 \leq a \leq \left\lceil 2 (\log n)^2 \right\rceil$ with

$$a^{(n-1)/2} \not\equiv \pm 1 \mod n.$$

In practice it is even better.

Example. If $n < 341550071728321$, then one of $a = 2, 3, 5, 7, 11, 13, 17$ will work.

5 Public-key cryptography

Public-key cryptography is private communication and identity verification.

5.1 Messages as sequences of classes mod n

How do we turn messages into numbers in $\mathbb{Z}/n\mathbb{Z}$? Idea is to choose n very large. Say $n > 2^{8k}$. Write down your message. Break it up into strings of at most k characters. Encode each character as an 8 bit binary number. String these integers together to get an $8k$ bit binary number. Regard that as an integer mod n.

5.2 The Rivest-Shamir-Adleman (RSA) algorithm

Now apply some function $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, and then tell whoever you are trying to communicate with the result of this computation. Then they should apply some other function $g : \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, to get back the number you started with. So want f to be injective. Want to be able to make f public without making g public. Idea is to choose two large prime numbers p, q and set $n = pq$. Choose $(e, \Phi(n)) = 1$. Find d such that

$$de = 1 \mod \Phi(n) = (p-1)(q-1) = n - (p+q) + 1.$$

Publish n and e, you keep $p, q, \Phi(n), d$ secret. $f(x) = x^e \mod n$ and $g(x) = x^d \mod n$.

$$x^{de} \equiv x^1 \equiv x \mod n,$$

because $de \equiv 1 \mod \Phi(n)$ and $x^{\Phi(n)} \equiv 1 \mod n$. So if someone wants to send you a message $c \in \mathbb{Z}/n\mathbb{Z}$, they compute $c^e \in \mathbb{Z}/n\mathbb{Z}$, and send it to you. To decode it, you compute

$$(c^e)^d = c^{de} \equiv c \mod n.$$

This assumes that $(c, n) = 1$, but the probability of this is extremely high. The prevailing assumption is that with only the information n and e, it is hopeless to discover d, or to find any other way of recovering c from c^e.

Lecture 9 Wednesday 24/10/18

Lecture 9 is a problem class.

5.3 Signing with RSA

If you have functions $f, g : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ with $f \circ g = g \circ f = id$, then you can also verify your identity, that is sign messages. Again, make f public, and any time you publish a message m, you also publish $g(m)$. Then anyone can apply f to $g(m)$ to recover $m = f(g(m))$, but without g, no one can forge your signature.

Lecture 10 Friday 26/10/18

5.4 Discrete logarithms

Suppose that n is prime, or more generally that $(\mathbb{Z}/n\mathbb{Z})^\times$ is cyclic. Let g be a generator for this group, that is a primitive root. For any $a \in (\mathbb{Z}/n\mathbb{Z})^\times$, we can write $a = g^m$ for some unique $0 \leq m < \Phi(n)$. We call m the **discrete logarithm** of a to base g, and write $m = \log_g(a)$.

Example. If you want to solve $x^r \equiv a \mod n$, write $x = g^y$, and the congruence becomes equivalent to $yr \equiv \log_g(a) \mod \Phi(n)$.

Unfortunately, or fortunately for cryptography, computing $\log_g$ is believed to be a hard problem. In particular, no known polynomial time algorithm.

Example. Imagine that you have a system where you need to store passwords for different users, but you do not want to store the actual passwords. One way to do this is to choose a large prime p and a primitive root g, and if someone inputs x as their password, you store $g^x \mod p$. If they later input y, you compute g^y, and check it matches what you stored. If it does then $y \equiv x \mod p - 1$.

6 Quadratic reciprocity

6.1 Quadratic residues

Let p be a prime number.

Definition 33. If $(a,p)=1$, then a is a **quadratic residue** (QR) if and only if there is a solution to $x^2 \equiv a \mod p$. If $(a,p)=1$ and is not a QR, it is called a **quadratic non-residue** (QNR).

Example.

- If $p=2$, 1 is a QR.
- If $p=3$, 1 is a QR, -1 is a QNR, since $1^2 \equiv (-1)^2 \equiv 1 \mod 3$.
- If $p=5$, $1,4$ are QRs, $2,3$ are QNRs, since $1^2 \equiv (-1)^2 \equiv 1 \mod 5$ and $2^2 \equiv 3^2 \equiv 4 \mod 5$.

Lemma 34. If $p>2$ then there are exactly $(p-1)/2$ QRs, and $(p-1)/2$ QNRs mod p.

Proof. The map

$$\begin{aligned} \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times &\to \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times \\ x &\mapsto x^2 \end{aligned}$$

is a group homomorphism with kernel $\{\pm 1\}$. So the image has order $(p-1)/2$, and the image is exactly the QRs. □

Proposition 35. Suppose that $(a,p)=(b,p)=1$. Then

- if a,b are both QRs, then ab is a QR,
- if one of a,b is a QR and one is a QNR, then ab is a QNR, and
- if a,b are both QNRs, then ab is a QR.

Proof. Let H be the image of

$$\begin{aligned} \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times &\to \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times \\ x &\mapsto x^2 \end{aligned}$$

that is H is the QRs. Then $(\mathbb{Z}/p\mathbb{Z})^\times / H$ is a group of order two by Lemma 34, so it is cyclic of order two. This statement is a restatement of Proposition 35, since

$$\left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times = H \cup 1 + H.$$

□

Definition 36. Let $a \in \mathbb{Z}$ and p a prime. Then the **Legendre symbol** is

$$\left(\tfrac{a}{p}\right) = \begin{cases} 1 & a \text{ is a QR mod } p \\ 0 & p \mid a \\ -1 & a \text{ is a QNR mod } p \end{cases}.$$

Proposition 35 can be restated as saying that

$$\left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times \to \{\pm 1\}$$
$$a \mapsto \left(\tfrac{a}{p}\right)$$

is a group homomorphism, that is

$$\left(\tfrac{ab}{p}\right) = \left(\tfrac{a}{p}\right)\left(\tfrac{a}{p}\right).$$

Even holds if we do not assume that $(a,p) = (b,p) = 1$.

Lecture 11
Tuesday
30/10/18

Theorem 37 (Euler's criterion)**.** If p is an odd prime, and $p \nmid a$, then

$$\left(\tfrac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p.$$

Proof. Let g be a primitive root mod p, and write $a \equiv g^r \mod p$ for $0 \le r < p-1$. Now

$$\left(g^{\frac{p-1}{2}}\right)^2 = g^{p-1} \equiv 1 \mod p.$$

So $g^{(p-1)/2} \equiv \pm 1 \mod p$. Since g is a primitive root, $g^{(p-1)/2} \not\equiv 1 \mod p$, so $g^{(p-1)/2} \equiv -1 \mod p$. So

$$a^{\frac{p-1}{2}} \equiv (g^r)^{\frac{p-1}{2}} \equiv \left(g^{\frac{p-1}{2}}\right)^r = (-1)^r \mod p.$$

But

$$\begin{aligned}
\left(\tfrac{a}{p}\right) = 1 \quad &\iff \quad \exists s \in \mathbb{Z},\ (g^s)^2 \equiv a \mod p \\
&\iff \quad 2s \equiv r \mod p-1 \\
&\iff \quad r \in 2\mathbb{Z} \\
&\iff \quad (-1)^r \equiv 1 \mod p.
\end{aligned}$$

□

6.2 Computing Legendre symbols

Proposition 38. -1 is a square mod p if and only if $p = 2$ or $p \equiv 1 \mod 4$.

Proof. $p = 2$ is trivial. If $p > 2$, then by Euler's criterion,

$$\left(\tfrac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \mod p,$$

so in fact $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$. Then

$$(-1)^{\frac{p-1}{2}} = \begin{cases} 1 & p \equiv 1 \mod 4 \\ -1 & p \equiv 3 \mod 4 \end{cases}.$$

□

Proposition 39.

$$\left(\tfrac{2}{p}\right) = \begin{cases} 1 & p \equiv \pm 1 \mod 8 \\ -1 & p \equiv \pm 3 \mod 8 \end{cases},$$

that is $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$.

Example.

- $\left(\frac{2}{7}\right) = 1$, since $2 \equiv 3^3 \mod 7$.
- $\left(\frac{2}{11}\right) = -1$, since squares mod 11 are $1, 4, 9, 5, 3$.
- $\left(\frac{-1}{11}\right) = -1$, so $\left(\frac{-2}{11}\right) = \left(\frac{2}{11}\right)\left(\frac{-1}{11}\right) = (-1)^2 = 1$ and $-2 \equiv 3^2 \mod 11$.

Proof. $\left(\frac{2}{p}\right) \equiv 2^{(p-1)/2} \mod p$ by Euler's criterion. Let $q = (p-1)/2$, and let

$$Q = (2)(4)\ldots(p-3)(p-1) = (2(1))\ldots(2(q)) = 2^q q! = 2^{\frac{p-1}{2}} q!.$$

Subtracting p from every term which is bigger than q,

$$Q \equiv (2)(4)\ldots(-3)(-1) \equiv (-1)^r q! \mod p,$$

where r is the number of odd integers in $1, \ldots, q$. Since $p \nmid q!$, we have

$$2^{\frac{p-1}{2}} \equiv (-1)^r \mod p.$$

(Exercise: now check the following)

$$(-1)^r = \begin{cases} 1 & p \equiv \pm 1 \mod 8 \\ -1 & p \equiv \pm 3 \mod 8 \end{cases}.$$

□

Example. If $p \equiv 1 \mod 8$, say $p = 1 + 8n$, so $q = 4n$. Odd integers in $1, \ldots, 4n$ are

$$1, 3, \ldots, 4n-3, 4n-1,$$

so $r = 2n$.

Theorem 40 (Law of quadratic reciprocity)**.** If p, q are odd primes, then

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

that is $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ unless $p \equiv q \equiv 3 \mod 4$, when $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$.

Example.

- $\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right)$ for $p \neq 5$. QRs mod 5 are $1, 4$. So

$$\left(\frac{5}{p}\right) = \begin{cases} 1 & p \equiv \pm 1 \mod 5 \\ -1 & p \equiv \pm 2 \mod 5 \end{cases}.$$

- What is $\left(\frac{3}{p}\right)$ for $p \neq 3$?
  - If $p \equiv 1 \mod 4$, then
  $$\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right) = \begin{cases} 1 & p \equiv 1 \mod 3 \\ -1 & p \equiv -1 \mod 3 \end{cases}.$$
  - If $p \equiv -1 \mod 4$, then
  $$\left(\frac{3}{p}\right) = -\left(\frac{p}{3}\right) = \begin{cases} 1 & p \equiv -1 \mod 3 \\ -1 & p \equiv 1 \mod 3 \end{cases}.$$

So

$$\left(\tfrac{3}{p}\right) = \begin{cases} 1 & p \equiv \pm 1 \mod 12 \\ -1 & p \equiv \pm 5 \mod 12 \end{cases}.$$

For example,

- if $p = 7$, QRs are $1, 2, 4$, so $\left(\tfrac{3}{p}\right) = -1$, and
- if $p = 11$, $5^2 \equiv 3 \mod 11$, so $\left(\tfrac{3}{p}\right) = 1$.

Example.

- $\left(\tfrac{6}{19}\right) = \left(\tfrac{2}{19}\right)\left(\tfrac{3}{19}\right) = (-1)(-1) = 1$.
- $\left(\tfrac{2}{19}\right) = -1$, because $19 \equiv 3 \mod 8$.
- $\left(\tfrac{3}{19}\right) \equiv -1 \mod 12$, by the above.

In general to compute $\left(\tfrac{a}{p}\right)$, we could do the following. Use that if $a \equiv b \mod p$ then $\left(\tfrac{a}{p}\right) = \left(\tfrac{b}{p}\right)$. So without loss of generality $|a| < p$. Then write $a = \pm \prod_i q_i^{s_i}$ for q_i prime. Then

$$\left(\tfrac{a}{p}\right) = \left(\tfrac{\pm 1}{p}\right) \prod_i \left(\tfrac{q_i}{p}\right)^{s_i}.$$

If s_i is even, then $\left(\tfrac{q_i}{p}\right)^{s_i} = 1$. If s_i is odd, then $\left(\tfrac{q_i}{p}\right)^{s_i} = \left(\tfrac{q_i}{p}\right)$. We have formulas for $\left(\tfrac{-1}{p}\right)$ and $\left(\tfrac{2}{p}\right)$. If q is an odd prime, $q < p$, then use quadratic reciprocity to relate $\left(\tfrac{q}{p}\right)$ and $\left(\tfrac{p}{q}\right)$. Then repeat mod q.

Lecture 12
Wednesday
31/10/18

6.3 Proof of quadratic reciprocity

Proof of this is due to Rousseau, in 1991. Resembles the proof we gave that $\left(\tfrac{2}{p}\right) = (-1)^{(p^2-1)/8}$.

$$\left(\frac{\mathbb{Z}}{pq\mathbb{Z}}\right)^\times \cong \left(\frac{\mathbb{Z}}{p\mathbb{Z}}\right)^\times \times \left(\frac{\mathbb{Z}}{q\mathbb{Z}}\right)^\times.$$

We will write down several choices of coset representatives for $\{\pm 1\}$, and compare them, that is we will write down choices of x or $-x$ for each $x \in (\mathbb{Z}/pq\mathbb{Z})^\times$.

Theorem 41 (Wilson's theorem)**.** If p is prime, then $(p-1)! \equiv -1 \mod p$.

Write elements of $(\mathbb{Z}/pq\mathbb{Z})^\times$ as pairs $(\alpha, \beta) \in (\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times$.

- For our first set of coset representatives, take

$$\left\{ (x, y) \;\middle|\; 1 \le x \le \frac{p-1}{2},\ 1 \le y \le q-1 \right\}.$$

Let A be the product of these coset representatives. This is by definition

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

- The second set of representatives is

$$\left\{ (x, y) \;\middle|\; 1 \le x \le p-1,\ 1 \le y \le \frac{q-1}{2} \right\}.$$

Let B be the product of these representatives. Then by symmetry,

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

- For the third set of representatives, select the pairs (x, y) which correspond via the Chinese remainder theorem to the set
$$\left\{ 1 \leq i \leq \frac{pq-1}{2} \;\middle|\; (i, pq) = 1 \right\}.$$
Let C be the product of these coset representatives. What is the x-coordinate of C? It is
$$\prod_{i=1,\ (i,pq)=1}^{\frac{pq-1}{2}} i.$$
So
$$\prod_{i=1,\ (i,pq)=1}^{\frac{pq-1}{2}} i = \left(\prod_{i=1,\ (i,p)=1}^{\frac{pq-1}{2}} i \right) \Big/ \left(\prod_{i=1,\ (i,p)=1,\ q|i}^{\frac{pq-1}{2}} i \right), \tag{1}$$
$$\prod_{i=1,\ (i,p)=1}^{\frac{pq-1}{2}} i = \left(\prod_{i=1,\ (i,p)=1}^{p\left(\frac{q-1}{2}\right)} i \right) \left(\prod_{i=p\left(\frac{q-1}{2}\right)+1,\ (i,p)=1}^{p\left(\frac{q-1}{2}\right)+\frac{p-1}{2}} i \right), \tag{2}$$
$$\prod_{i=1,\ (i,p)=1,\ q|i}^{\frac{pq-1}{2}} i = \prod_{j=1,\ (j,p)=1}^{\frac{p-1}{2}} qj = q^{\frac{p-1}{2}} \left(\frac{p-1}{2} \right)!. \tag{3}$$
Combining (1), (2), (3), get that the x-coordinate of the product is
$$\prod_{i=1,\ (i,pq)=1}^{\frac{pq-1}{2}} i = \frac{(p-1)!^{\frac{q-1}{2}} \left(\frac{p-1}{2}\right)!}{q^{\frac{p-1}{2}} \left(\frac{p-1}{2}\right)!} = \frac{(-1)^{\frac{q-1}{2}}}{q^{\frac{p-1}{2}}}.$$
So C, the product of these representatives, is
$$C = \left((-1)^{\frac{q-1}{2}} \left(\frac{q}{p}\right), (-1)^{\frac{p-1}{2}} \left(\frac{p}{q}\right) \right).$$

A, B, C all agree up to sign, that is up to multiplication by ± 1, that is up to multiplication by $(-1, -1) \in (\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times$. Looking at y-coordinates, $C = \left(\frac{p}{q}\right) A$. Similarly $C = \left(\frac{q}{p}\right) B$. So
$$B = \left(\frac{q}{p}\right) \left(\frac{p}{q}\right) A.$$
How did we define A and B? A is the product of But we can compare A and B more directly. To swap between A and B, just change the signs of everything with $1 \leq x \leq (p-1)/2$ and $(q+1)/2 \leq y \leq q-1$. So
$$B = (-1)^{\left(\frac{p-1}{2}\right)\left(\frac{q-1}{2}\right)} A.$$
So $\left(\frac{q}{p}\right) \left(\frac{p}{q}\right) = (-1)^{((p-1)/2)((q-1)/2)}$, that is
$$\left(\frac{q}{p}\right) = (-1)^{\left(\frac{p-1}{2}\right)\left(\frac{q-1}{2}\right)} \left(\frac{p}{q}\right).$$

6.4 Jacobi symbols

These are an extension of Legendre symbols which are useful for making computations.

Definition 42. Write $b = \prod_i p_i^{r_i}$ for p_i distinct primes. Then the **Jacobi symbol** is

$$\left(\tfrac{a}{b}\right) = \prod_i \left(\tfrac{a}{p_i}\right)^{r_i}.$$

Warning that $\left(\frac{a}{b}\right) = 1$ does not imply that a is a square mod b. On the other hand, $\left(\frac{a}{b}\right) = -1$ implies that a is not a square mod b.

Lecture 13
Friday
02/11/18

Lemma 43.

1. $\left(\frac{a_1 a_2}{b}\right) = \left(\frac{a_1}{b}\right)\left(\frac{a_2}{b}\right)$ and $\left(\frac{a}{b_1 b_2}\right) = \left(\frac{a}{b_1}\right)\left(\frac{a}{b_2}\right)$.
2. $\left(\frac{a}{b}\right)$ depends only on $a \mod b$.
3. $\left(\frac{a^2}{b}\right) = 1$.
4. $\left(\frac{-1}{b}\right) = (-1)^{(b-1)/2}$.
5. $\left(\frac{2}{b}\right) = (-1)^{\left(b^2-1\right)/8}$.
6. If $a, b > 0$ are both odd $\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = (-1)^{((a-1)/2)((b-1)/2)}$.

Proof. All of these statements are true for Legendre symbols, that is for b prime, and a prime in 6. 1 to 3 follow immediately. 4 to 6 also follows from 1 and the corresponding statements for Legendre symbols. For 5, it is enough to show that if it holds for b_1, b_2, then it holds for $b_1 b_2$. Since $\left(\frac{2}{b_1 b_2}\right) = \left(\frac{2}{b_1}\right)\left(\frac{2}{b_2}\right)$, we need to show that

$$(-1)^{\frac{b_1^2-1}{8}} (-1)^{\frac{b_2^2-1}{8}} = (-1)^{\left(\frac{(b_1 b_2)^2-1}{8}\right)},$$

that is need

$$\left(b_1^2 - 1\right) + \left(b_2^2 - 1\right) \equiv \left(b_1 b_2\right)^2 - 1 \mod 16,$$

that is

$$\left(b_1^2 - 1\right)\left(b_2^2 - 1\right) \equiv 0 \mod 4.$$

True because $b_1^2 \equiv b_2^2 \equiv 1 \mod 4$. □

Example.

$$\begin{aligned}\left(\tfrac{7411}{9283}\right) &= -\left(\tfrac{9283}{7411}\right) = -\left(\tfrac{1872}{7411}\right) = -\left(\tfrac{16}{7411}\right)\left(\tfrac{117}{7411}\right) = -\left(\tfrac{117}{7411}\right) = -\left(\tfrac{7411}{117}\right) = -\left(\tfrac{40}{117}\right) \\ &= -\left(\tfrac{8}{117}\right)\left(\tfrac{5}{117}\right) = -\left(\tfrac{2}{117}\right)\left(\tfrac{5}{117}\right) = \left(\tfrac{5}{117}\right) = \left(\tfrac{117}{5}\right) = \left(\tfrac{2}{5}\right) = -1.\end{aligned}$$

So 7411 is not a square mod 9283.

7 Sums of squares

Which integers are the sum of two squares? Which integers are the sum of four squares?

7.1 Sums of two squares

Definition 44. We say that $n \in \mathbb{Z}$ is a **sum of two squares** if $n = x^2 + y^2$ for $x, y \in Z$.

Example. If $n = x^2 + y^2$, then since $x^2, y^2 \equiv 0, 1 \mod 4$, we cannot have $n \equiv 3 \mod 4$.

Example. $21 \equiv 1 \mod 4$, but 21 is not a sum of two squares. On the other hand, we will see that all primes which are $1 \mod 4$ are sums of two squares.

Definition 45. The **Gaussian integers** $\mathbb{Z}[i]$ are the subring of $\mathbb{C}$ consisting of $a + bi$ for $a, b \in \mathbb{Z}$. The **norm** is defined by

$$\begin{aligned} N : \mathbb{Z}[i] &\to \mathbb{Z}_{>0} \\ a + bi &\mapsto a^2 + b^2, \end{aligned}$$

that is $N(z) = z\bar{z}$.

$$N(zw) = (zw)(z\bar{w}) = (z\bar{z})(w\bar{w}) = N(z)N(w).$$

If $z = a + bi$, $w = c + di$, then $zw = (ac - bd) + (ad + bc)i$, so

$$\left(a^2 + b^2\right)\left(c^2 + d^2\right) = (ac - bd)^2 + (ad + bc)^2.$$

Lemma 46. If m, n are each a sum of two squares, then so is mn.

Theorem 47 (Fermat's two square theorem)**.** If $p \equiv 1 \mod 4$ is prime, then p is a sum of two squares.

Lemma 46 and Theorem 47 together allow you to give a complete classification of the integers which are sums of two squares, in terms of their prime factorisations.

Definition 48. A ring R is a **Euclidean domain** if it is an integral domain, that is $ab = 0$ gives $a = 0$ or $b = 0$, and there exists a function $N : R \to \mathbb{Z}_{\geq 0}$ such that for all $a, b \in R$ with $b \neq 0$, there exist $q, r \in R$ such that $a = qb + r$, and $r = 0$ or $N(r) < N(b)$.

If R is a Euclidean domain, then you can carry out Euclid's algorithm. In particular, irreducible elements are the same as prime elements, and every element can be factored as a product of primes, uniquely up to reordering and multiplication by units. $\mathbb{Z}[i]$ together with N is a Euclidean domain. By definition, $n \in \mathbb{Z}$ is a sum of two squares if and only if there exists $z \in \mathbb{Z}[i]$ with $N(z) = n$. Since $N(zw) = N(z)N(w)$, all we have to do is to figure out what the primes in $\mathbb{Z}[i]$ are, and what their norms are. (Exercise: show that the units in $\mathbb{Z}[i]$ are ± 1, $\pm i$) Two elements of $\mathbb{Z}[i]$ are **associates** if their ratio is a unit, that is z, w are associates if $z = uw$, for $u \in \{\pm 1, \pm i\}$.

Lecture 14
Tuesday
06/11/18

Lemma 49. Let p be a prime in $\mathbb{Z}[i]$. Then there is a prime q of $\mathbb{Z}$ such that either $N(p) = q$ or $N(p) = q^2$. In the latter case, p is an associate of q. Given q a prime in $\mathbb{Z}$, there exists p such that $N(p) = q$ if and only if q is a sum of two squares.

Proof. Write $n = N(p)$, and let $n = q_1^{s_1} \dots q_r^{s_r}$ be the prime factorisation of n in $\mathbb{Z}$. By definition $n = p\bar{p}$, so $p \mid n$ in $\mathbb{Z}[i]$, and so since p is prime, $p \mid q_i$ for some i. Write $q = q_i$. Then $p \mid q$ gives $q = pv$ for some v, so

$$N(p)N(v) = N(pv) = N(q) = q^2.$$

If $N(p) = 1$, then p is a unit, a contradiction. So $N(p) \mid q^2$ gives $N(p) = q$ or $N(p) = q^2$, as claimed. If $N(p) = q^2$, then $N(v) = 1$, so v is a unit, and since $q = pv$, p is an associate of q, by definition. If $N(p) = q$, then writing $p = a + bi$, we have $q = a^2 + b^2$. Conversely, if

$$q = a^2 + b^2 = (a + bi)(a - bi),$$

then since $p \mid q$, we have either $p \mid a + bi$ or $p \mid a - bi$, so $N(p) \mid N(a + bi) = q$ or $N(p) \mid N(a - bi) = q$, and either way $N(p) = q$. □

Corollary 50. The primes in $\mathbb{Z}[i]$ are either of the form $a+bi$ with a^2+b^2 a prime in $\mathbb{Z}$, or are primes of $\mathbb{Z}$ which are not sums of two squares.

Theorem 51. If $p=2$ or $p\equiv 1 \mod 4$, then p is a sum of two squares.

Proof. By Corollary 50, we just have to show that p is not a prime in $\mathbb{Z}[i]$. There exists n such that $n^2\equiv -1 \mod p$. If $p=2$ obvious, and if $p\equiv 1 \mod 4$,

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

by Euler's criterion. That is,

$$p \mid n^2+1 = (n+i)(n-i).$$

If p were prime, then $p \mid n+i$ or $p \mid n-i$, that is there exist $c,d\in\mathbb{Z}$ such that $n\pm i = p(c+di)$, so $1=pd$, a contradiction. □

Remark that if $p\equiv 3 \mod 4$ then p is not a sum of two squares, even mod 4. Remark that in practice, to go from $n^2+1\equiv 0 \mod p$ to finding a,b with $a^2+b^2=p$, you just compute $(n+i,p)=a+bi$. You can do this computation with Euclid's algorithm in $\mathbb{Z}[i]$.

Theorem 52. $n\in\mathbb{Z}$ is a sum of two squares if and only if its prime factorisation only contains primes congruent to $3 \mod 4$ to even powers, that is

$$n = 2^a \prod_{p_i\equiv 1 \mod 4} p_i^{r_i} \prod_{q_i\equiv 3 \mod 4} q_i^{2s_i}.$$

Proof. Suppose n is of this form. Then 2, each p_i, and each q_i^2 are all sums of two squares, so n is a sum of two squares by Lemma 46. Conversely suppose that $n=a^2+b^2$, and write $a+bi$ as a product of primes in $\mathbb{Z}[i]$. Then $n=N(a+bi)$ is the product of the norms of these primes, and we already saw that the norms of primes in $\mathbb{Z}[i]$ are either 2, a prime which is $1 \mod 4$, or the square of a prime which is $3 \mod 4$. □

7.2 Sums of four squares - the ring of quaternions

Lagrange's theorem states that every positive integer is a sum of four squares.

Definition 53. $\mathbb{H}$, the **ring of quaternions**, is the ring of sums $a+bi+cj+dk$ for $a,b,c,d\in\mathbb{R}$, such that

- addition is

$$(a+bi+cj+dk)+(A+Bi+Cj+Dk) = (a+A)+(b+B)i+(c+C)j+(d+D)k,$$

- multiplication is

$$ij=-ji=k, \qquad jk=-kj=i, \qquad ki=-ik=j.$$

If $z=a+bi+cj+dk$, we write $z^*=a-bi-cj-dk$, so $(zw)^*=w^*z^*$.

Define $N(z)=zz^*=a^2+b^2+c^2+d^2$. Then

$$N(zw)=zw(zw)^*=zww^*z^*=zN(w)z^*=zz^*N(w)=N(z)N(w),$$

because $N(w)\in\mathbb{R}$. So

$$\begin{aligned}\left(a^2+b^2+c^2+d^2\right)\left(x^2+y^2+z^2+w^2\right) &= N(a+bi+cj+dk)N(x+yi+zj+wk)\\ &= N((a+bi+cj+dk)(x+yi+zj+wk))\\ &= (ax-by-cz-dw)^2+(ay+bx+cw-dz)^2\\ &+ (az-bw+cx+dy)^2+(aw+bz-cy+dx)^2.\end{aligned}$$

In particular, if m,n are sums of four squares, then mn is a sum of four squares. So to prove Lagrange's theorem, it suffices to show that all primes are sums of four squares.

7.3 Proof of Lagrange's theorem

We already saw that 2, and any prime congruent to 1 mod 4, is a sum of two squares. It remains to show that any prime congruent to 3 mod 4 is a sum of four squares.

Lecture 15
Wednesday
07/11/18

Lemma 54. If $p \equiv 3 \mod 4$ is prime, then there exist x, y such that $x^2 + y^2 + 1 \equiv 0 \mod p$.

Proof. Firstly, claim there exists a such that $\left(\frac{a}{p}\right) = 1$ and $\left(\frac{a+1}{p}\right) = -1$. If not, since $\left(\frac{1}{p}\right) = 1$, we must have

$$\left(\frac{2}{p}\right) = \cdots = \left(\frac{p-1}{p}\right) = 1.$$

But we know that there are $(p-1)/2$ values of b with $1 \leq b \leq p-1$ and $\left(\frac{b}{p}\right) = -1$, a contradiction. Since $p \equiv 3 \mod 4$, $\left(\frac{-1}{p}\right) = -1$ by Euler's criterion. So

$$\left(\frac{-(a+1)}{p}\right) = \left(\frac{a+1}{p}\right)\left(\frac{-1}{p}\right) = 1.$$

Choose x such that $x^2 \equiv a \mod p$ and y such that $y^2 \equiv -(a+1) \mod p$. Then $x^2 + y^2 \equiv -1 \mod p$. □

By Lemma 54, there exist $x, y \in \mathbb{Z}$ such that $x^2 + y^2 + 1 = pr$ for some r. Since the congruence $x^2 + y^2 + 1 \equiv 0 \mod p$ only depends on $x, y \mod p$, we can find x, y with $-p/2 < x, y < p/2$. Then

$$\frac{x^2 + y^2 + 1}{p} = r < p.$$

Proposition 55. Suppose that

$$x^2 + y^2 + z^2 + w^2 = pr, \qquad 1 \leq r < p.$$

If $r > 1$, there exist x', y', z', w', r', for

$$x'^2 + y'^2 + z'^2 + w'^2 = pr', \qquad 1 \leq r' < r.$$

Proposition 55 gives p is a sum of four squares, starting with x, y, r as above, $z = 1$, and $w = 0$.

Proof.

- Suppose firstly that r is even. Then either x, y, z, w are all even, all odd, or two are even and two are odd. So without loss of generality $x \equiv y \mod 2$ and $z \equiv w \mod 2$. Then take

$$x' = \frac{x+y}{2}, \qquad y' = \frac{x-y}{2}, \qquad z' = \frac{z+w}{2}, \qquad w' = \frac{z-w}{2}, \qquad r' = \frac{r}{2}.$$

- Suppose now that r is odd, and choose $a, b, c, d \in (-r/2, r/2)$ such that

$$x \equiv a \mod r, \qquad y \equiv b \mod r, \qquad z \equiv c \mod r, \qquad w \equiv d \mod r.$$

  Then

$$a^2 + b^2 + c^2 + d^2 \equiv x^2 + y^2 + z^2 + w^2 = pr \equiv 0 \mod r.$$

  Write $a^2 + b^2 + c^2 + d^2 = rr'$. Then $rr' < 4\left(r/2\right)^2 = r^2$, so $0 \leq r' < r$. If $r' = 0$ then $a = b = c = d = 0$, so r' divides each of x, y, z, w. Since $x^2 + y^2 + z^2 + w^2 = pr$, we get $r^2 \mid pr$ so $r \mid p$, and since $r < p$, we get $r = 1$, and we are done. Otherwise $1 \leq r' < r$.

$$\begin{aligned}(rr')(rp) &= \left(a^2 + b^2 + c^2 + d^2\right)\left(x^2 + y^2 + z^2 + w^2\right) \\ &= (ax + by + cz + dw)^2 + (-ay + bx + cw - dz)^2 \\ &\quad + (-az - bw + cx + dy)^2 + (-aw + bz - cy + dx)^2.\end{aligned}$$

Then

$$\begin{aligned} ax + by + cz + dw &\equiv x^2 + y^2 + z^2 + w^2 \equiv 0 \mod r, \\ -ay + bx + cw - dz &\equiv -xy + yx + zw - wz \equiv 0 \mod r, \\ -az - bw + cx + dy &\equiv -xz - yw + zx + wy \equiv 0 \mod r, \\ -aw + bz - cy + dx &\equiv -xw + yz - zy + wx \equiv 0 \mod r. \end{aligned}$$

So take

$$x' = \frac{ax + by + cz + dw}{r}, \qquad y' = \frac{-ay + bx + cw - dz}{r},$$
$$z' = \frac{-az - bw + cx + dy}{r}, \qquad w' = \frac{-aw + bz - cy + dx}{r}.$$

□

Remark 56. This can be interpreted as a version of Euclid's algorithm in the ring

$$\left\{ \frac{a + bi + cj + dk}{2} \;\middle|\; a \equiv b \equiv c \equiv d \mod 2 \right\}.$$

Note that this ring is non-commutative, and also, for example,

$$5 = (1 - 2i)(1 - 2i) = (1 + 2j)(1 - 2j),$$

so you have to be careful with unique factorisation, etc.

7.4 Sums of three squares

7 is the smallest positive integer which is not a sum of three squares. In fact no integer congruent to 7 mod 8 can be a sum of three squares, because the squares mod 8 are $0, 1, 4$.

Theorem 57. A positive integer is not a sum of three squares if and only if it is of the form $4^a(8k + 7)$.

Proving that numbers are not of this form is beyond this course. Serre's a course in arithmetic is a good place to look.

8 Pell's equation

8.1 Pell's equation

Let $d \in \mathbb{Z}_{>1}$ be squarefree. **Pell's equation** is $x^2 - dy^2 = 1$.

Example. Let $d = 2$. $(x, y) = (3, 2)$ is a solution. In fact, there are infinitely many solutions, and this is true for any d.

We will find it useful to write $x^2 - dy^2 = \left(x + \sqrt{d}y\right)\left(x - \sqrt{d}y\right)$. This suggests that we should look at a ring like

$$\mathbb{Z}\left[\sqrt{d}\right] = \left\{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\right\}.$$

Definition 58. If $\alpha \in \mathbb{C}$, then $\mathbb{Z}[\alpha]$ is the smallest subring of $\mathbb{C}$ containing α.

Example.

- If $\alpha = 1$, then $\mathbb{Z}[\alpha] = \mathbb{Z}$.
- If $\alpha = i$, $\mathbb{Z}[i]$ is what we wrote before.
- On the other hand $\mathbb{Z}[\pi]$ is the ring of

  $$a_0 + \cdots + a_n\pi^n, \qquad a_i \in \mathbb{Z},$$

  for n arbitrary.
- Also $\mathbb{Z}\left[\sqrt[3]{2}\right]$ is not just the set

  $$\left\{a + b\sqrt[3]{2} \mid a, b \in \mathbb{Z}\right\},$$

  because this set does not contain $\left(\sqrt[3]{2}\right)^2 = \sqrt[3]{4}$.
- Also $\mathbb{Z}[1/p]$ contains $1/p^n$ for all n, so in fact

  $$\mathbb{Z}[1/p] = \{a/p^n \mid a \in \mathbb{Z},\ n \geq 0\}.$$

An alternative definition is that $\mathbb{Z}[\alpha]$ is the intersection of all subrings of $\mathbb{C}$ containing α.
Lecture 16 is a problem class.

Lecture 16
Friday
09/11/18

8.2 Quadratic subrings of $\mathbb{C}$

Lecture 17
Tuesday
13/11/18

Definition 59. Say that $\alpha \in \mathbb{C}$ is an **algebraic integer of degree two** if it is a root of a polynomial $x^2 + ax + b$, for $a, b \in \mathbb{Z}$, and $\alpha \notin \mathbb{Z}$.

Example.

- $\alpha = i$ root of $X^2 + 1$.
- $\alpha = \sqrt{d}$ root of $X^2 - d$ for $d > 1$ squarefree.

Proposition 60. If α is an algebraic integer of degree two, then $\mathbb{Z}[\alpha] = \{x + y\alpha \mid x, y \in \mathbb{Z}\}$.

Proof. Since $\alpha \notin \mathbb{Z}$, we have $\alpha \notin \mathbb{Q}$, since if $\alpha = r/s$ for $(r, s) = 1$ then $r^2 + ars + bs^2 = 0$, so $s \mid r^2$, so $s \mid 1$, so $\alpha \in \mathbb{Z}$. So if $x, y \in \mathbb{Z}$ and $x + y\alpha = 0$, then $x = y = 0$. Certainly every $x + y\alpha \in \mathbb{Z}[\alpha]$. The set $\{x + y\alpha\}$ is closed under addition and subtraction, so we only have to check that is closed under multiplication. But

$$\begin{aligned}(x + y\alpha)(X + Y\alpha) &= xX + (xY + yX)\alpha + yY\alpha^2 \\ &= xX + (xY + yX)\alpha + yY(a\alpha + b) \\ &= (xX + byY) + (xY + yX + ayY)\alpha.\end{aligned}$$

□

If α is an algebraic integer of degree two, say that $\mathbb{Z}[\alpha]$ is a **real quadratic subring** of $\mathbb{C}$ if $\alpha \in \mathbb{R}$, and an **imaginary quadratic subring** of $\mathbb{C}$ if $\alpha \notin \mathbb{R}$. Let α^* be the other root of $X^2 + aX + b = 0$.

Example.

- $i^* = -i = \bar{i}$.
- $\sqrt{d}^* = -\sqrt{d}$.

If $z = x + y\alpha \in \mathbb{Z}[\alpha]$, write $z^* = x + y\alpha^*$. If $\mathbb{Z}[\alpha]$ is imaginary quadratic, then $\alpha^* = \bar{\alpha}$, and $z^* = \bar{z}$. This is not true if $\mathbb{Z}[\alpha]$ is real quadratic. Define $N(z) = zz^*$. Since α, α^* are the roots of $X^2 + aX + b$, we have

$$\alpha + \alpha^* = -a, \qquad \alpha\alpha* = b.$$

If $z = x + y\alpha$, then

$$N(z) = (x + y\alpha)(x + y\alpha^*) = x^2 + xy(\alpha + \alpha^*) + y^2\alpha\alpha^* = x^2 - axy + by^2 \in \mathbb{Z}.$$

We have $(zw)^* = z^*w^*$, so

$$N(z)N(w) = zz^*ww^* = (zw)(zw)^* = N(zw).$$

So $N : \mathbb{Z}[\alpha] \to \mathbb{Z}$ is multiplicative. If $\mathbb{Z}[\alpha]$ is imaginary quadratic then $z^* = \bar{z}$, and $N(z) \geq 0$. If $\mathbb{Z}[\alpha]$ is real quadratic, we can have $N(z) < 0$.

Example. $N\left(\sqrt{d}\right) = \left(\sqrt{d}\right)\left(-\sqrt{d}\right) = -d < 0$.

(Exercise: $N(x + y\alpha) = 0$ if and only if $x = y = 0$)

Example. If $\alpha = \sqrt{d}$,

$$N\left(x + y\sqrt{d}\right) = \left(x + y\sqrt{d}\right)\left(x + y\sqrt{d}\right)^* = \left(x + y\sqrt{d}\right)\left(x - y\sqrt{d}\right) = x^2 - dy^2.$$

So solutions to Pell's equation are the same thing as elements of $\mathbb{Z}[\alpha]$ of norm one.

8.3 Factorisation in quadratic rings

Definition 61. The units of $\mathbb{Z}[\alpha]$ are by definition the elements with multiplicative inverses, and they form a group $\mathbb{Z}[\alpha]^\times$ under multiplication. We say that $z, w \in \mathbb{Z}[\alpha]$ are associates if $z = uw$ for $u \in \mathbb{Z}[\alpha]^\times$.

If $u \in \mathbb{Z}[\alpha]^\times$, then write $1 = uv$. Then $1 = N(1) = N(u)N(v)$, so $N(u) = \pm 1$. Conversely if $N(u) = \pm 1$, then $\pm 1 = N(u) = u(u^*)$, so $u(\pm u^*) = 1$, so $u \in \mathbb{Z}[\alpha]^\times$. So

$$\mathbb{Z}[\alpha]^\times = \{z \in \mathbb{Z}[\alpha] \mid N(z) = \pm 1\}.$$

Write

$$\mathbb{Z}[\alpha]^{\times,1} = \{z \in \mathbb{Z}[\alpha] \mid N(z) = 1\}.$$

Then $\mathbb{Z}[\alpha]^{\times,1}$ is a multiplicative subgroup of $\mathbb{Z}[\alpha]^\times$.

8.4 Back to Pell's equation

Example. If $\alpha = \sqrt{d}$, for $d > 1$ squarefree, then

$$\mathbb{Z}\left[\sqrt{d}\right]^{\times,1} = \left\{x + y\sqrt{d} \mid x^2 - dy^2 = 1\right\}.$$

(Exercise: if $\mathbb{Z}[\alpha]$ is imaginary quadratic, show that $\mathbb{Z}[\alpha]^\times = \mathbb{Z}[\alpha]^{\times,1}$ is finite, so what are the possibilities for this group?) What is $\mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$? Certainly contains ± 1. Anything else will be of the form $x + y\sqrt{d}$ with $x, y \neq 0$.

Lemma 62. Let $x + y\sqrt{d}$ be an element of $\mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$. Then

$$\begin{aligned} x > 0, \quad y > 0 \quad &\iff \quad x + y\sqrt{d} > 1, \\ x > 0, \quad y < 0 \quad &\iff \quad 0 < x + y\sqrt{d} < 1, \\ x < 0, \quad y > 0 \quad &\iff \quad -1 < x + y\sqrt{d} < 0, \\ x < 0, \quad y < 0 \quad &\iff \quad x + y\sqrt{d} < -1. \end{aligned}$$

Proof. If $x, y > 0$ then $x + y\sqrt{d} > y\sqrt{d} \geq \sqrt{d} > 1$. Then

$$x - y\sqrt{d} = \frac{1}{x + y\sqrt{d}} \in (0, 1).$$

So replacing y by $-y$, we get $x > 0$ and $y < 0$, so $0 < x + y\sqrt{d} < 1$. Replacing (x, y) with $(-x, -y)$ gives the forward in the third and fourth lines. Since the four possibilities for the right hand side are exhaustive for $x, y \neq 0$, we are done. □

Lemma 63. Let $z = x + y\sqrt{d}$ and $z' = x' + y'\sqrt{d}$ be two elements of $\mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$ with $z, z' > 1$, that is $x, y, x', y' > 0$. Then

$$z > z' \qquad \iff \qquad y > y'.$$

Proof.

$$z - \frac{1}{z} = x + y\sqrt{d} - \left(x - y\sqrt{d}\right) = 2y\sqrt{d}.$$

Just need to check that

$$z > z' \qquad \iff \qquad z - \frac{1}{z} > z' - \frac{1}{z'}.$$

But $z - 1/z$ is increasing, since its derivative is $1 + 1/z^2 > 0$. □

Lecture 18
Wednesday
14/11/18

Suppose that there exists $z \in \mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$, so $z \neq \pm 1$. By replacing z by $\pm z^{\pm 1}$, we can assume that $z > 1$. So by Lemma 62, if $z = x + y\sqrt{d}$, then $x, y > 0$. Let $\epsilon = x + y\sqrt{d} \in \mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$ with $x, y > 0$ and y as small as possible. Call ϵ the **fundamental 1-unit** of $\mathbb{Z}\left[\sqrt{d}\right]$.

Proposition 64. Suppose that $\mathbb{Z}\left[\sqrt{d}\right]^{\times,1} \neq \{\pm 1\}$, and let ϵ be the fundamental 1-unit. Then every element of $\mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$ is of the form $\pm\epsilon^n$ for some $n \in \mathbb{Z}$.

Conversely, $N(\pm\epsilon^n) = N(\pm 1)\, N(\epsilon)^n = 1$.

Proof. Let $z \in \mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$, so $z \neq \pm 1$. After replacing z by $\pm z^{\pm 1}$, we may assume that $z > 1$. Choose $n \geq 0$ such that $\epsilon^n \leq z < \epsilon^{n+1}$. Then $1 \leq z\epsilon^{-n} < \epsilon$, so

$$N\left(z\epsilon^{-n}\right) = N(z)\, N(\epsilon)^{-n} = 1.$$

So $z\epsilon^{-n} \in \mathbb{Z}\left[\sqrt{d}\right]^{\times,1}$. So by the choice of ϵ, and Lemma 63, we have $z\epsilon^{-n} = 1$, that is $z = \epsilon^n$. □

Example. Let $d = 2$ and $x^2 - 2y^2 = 1$. $y = 2$ and $x = 3$ is a solution. So $\epsilon = 3 + 2\sqrt{2}$.

$$\epsilon^2 = \left(3 + 2\sqrt{2}\right)^2 = 17 + 12\sqrt{2},$$

and $17^2 - 2(12)^2 = 1$.

8.5 Constructing the fundamental 1-unit

Idea is that if $x^2 - dy^2 = 1$, for $x, y > 0$, then $x/y \approx \sqrt{d}$.

$$\left|x - y\sqrt{d}\right| = \frac{1}{\left|x + y\sqrt{d}\right|},$$

which is small. So one way to to try to find 1-units is to find rational numbers which are good approximations to $\sqrt{d}$. Want to make $\left|x/y - \sqrt{d}\right|$ as small as possible for y of a given size. More generally, if $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, we might want to find $x, y > 0$ such that $|x/y - \alpha| < C/y^n$, where C, n are fixed.

$n = 0$ Trivial.

$n = 1$, $C = 1$ Trivial, by just choosing any y and x/y as close to α as you can.

$n = 2$, $C = 1$ Not obvious. In fact there always exist infinitely many x, y with $|x/y - \alpha| < 1/y^2$, as we now show.

Theorem 65 (Dirichlet)**.** Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and let $Q \in \mathbb{Z}_{>1}$. Then there exist $p, q \in \mathbb{Z}$, such that $1 \leq q < Q$, and $|p - q\alpha| < 1/Q$.

Proof. For $1 \leq k \leq Q - 1$, let $a_k = \lfloor k\alpha \rfloor$. Then $0 < k\alpha - a_k < 1$. Consider the Q intervals

$$\left[0, \frac{1}{Q}\right], \ldots, \left[\frac{Q-1}{Q}, 1\right].$$

The set

$$\{0, \alpha - a_1, \ldots, (Q-1)\,\alpha - a_{Q-1}, 1\},$$

contains $Q + 1$ elements, so some pair of them must be in the same interval. The difference of these two elements is of the form $p - q\alpha$, for $1 \leq q < Q$. □

Corollary 66. For any $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, there exist infinitely many pairs $p, q \in \mathbb{Z}$ such that $|\alpha - p/q| < 1/q^2$.

Proof. Certainly there exists p for $q = 1$. It is then enough to prove that if $|\alpha - p/q| < 1/q^2$, there exist p', q' such that $|\alpha - p'/q'| < 1/(q')^2$ and $|\alpha - p'/q'| < |\alpha - p/q|$. Choose Q such that $1/Q < |\alpha - p/q|$. By Theorem 65, there exist p', q' with $1 \leq q' < Q$, and $|\alpha - p'/q'| < 1/Qq' < 1/(q')^2$. Also

$$\left|\alpha - \frac{p'}{q'}\right| < \frac{1}{Qq'} \leq \frac{1}{Q} < \left|\alpha - \frac{p}{q}\right|,$$

as required. □

We can now show the following.

Theorem 67. If $d > 1$ is squarefree, then there exist x, y such that $y \neq 0$ and $x^2 - dy^2 = 1$.

Proof. By Corollary 66, there exist infinitely many (p_i, q_i) for $p_i, q_i > 0$ such that $\left|p_i/q_i - \sqrt{d}\right| < 1/q_i^2$, that is $\left|p_i - q_i\sqrt{d}\right| < 1/q_i$. Then

$$\left|p_i + q_i\sqrt{d}\right| \leq \left|p_i - q_i\sqrt{d}\right| + 2q_i\sqrt{d} < \frac{1}{q_i} + 2q_i\sqrt{d} < 3q_i\sqrt{d}.$$

So

$$\left|N\left(p_i + q_i\sqrt{d}\right)\right| = \left|p_i + q_i\sqrt{d}\right|\left|p_i - q_i\sqrt{d}\right| < \left(3q_i\sqrt{d}\right)\frac{1}{q_i} = 3\sqrt{d}.$$

So there exists $M \in \left(-3\sqrt{d}, 3\sqrt{d}\right)$ such that $N\left(p_i + q_i\sqrt{d}\right) = M$ for infinitely many i. Then there exists (p_0, q_0) such that

$$p_i \equiv p_0 \mod M, \qquad q_i \equiv q_0 \mod M,$$

for infinitely many i. Now consider $(p_i, q_i) \neq (p_j, q_j)$ of this form, that is

$$N\left(p_i + q_i\sqrt{d}\right) = N\left(p_j + q_j\sqrt{d}\right) = M, \qquad p_i \equiv p_j \mod M, \qquad q_i \equiv q_j \mod M.$$

Then

$$\frac{p_i - q_i\sqrt{d}}{p_j - q_j\sqrt{d}} = \frac{\left(p_i - q_i\sqrt{d}\right)\left(p_j + q_j\sqrt{d}\right)}{M} = \frac{(p_ip_j - dq_iq_j) + (p_iq_j - p_jq_i)\sqrt{d}}{M},$$

$$p_iq_j \equiv p_jq_i \mod M, \qquad p_ip_j - dq_iq_j \equiv p_i^2 - dq_i^2 = M \equiv 0 \mod M.$$

So

$$N\left(\frac{p_i - q_i\sqrt{d}}{p_j - q_j\sqrt{d}}\right) = \frac{M}{M} = 1,$$

and

$$\frac{p_i - q_i\sqrt{d}}{p_j - q_j\sqrt{d}} \in \mathbb{Z}\left[\sqrt{d}\right],$$

as required. □

Lecture 19
Friday
16/11/18

8.6 The equation $x^2 - dy^2 = -1$

$x^2 - dy^2 = -1$ has solution if and only if there exists $u \in \mathbb{Z}\left[\sqrt{d}\right]^\times$ such that $N(u) = -1$. Given such a u, all solutions to the equation are given by $\pm u\epsilon^n$, for $n \in \mathbb{Z}$, since

$$N(v) = -1 \qquad \iff \qquad N(v) = N(u) \qquad \iff \qquad N(v/u) = 1.$$

Example. If $d = 3$, no solutions, as $X^2 \equiv -1 \mod 3$ has no solutions.

9 Continued fractions

9.1 Rational continued fractions

Let $p/q \in \mathbb{Q}$. Write

$$\frac{p}{q} = a_0 + r_0, \qquad a_0 = \left\lfloor \frac{p}{q} \right\rfloor, \qquad 0 \le r_0 < 1.$$

If $r_i \neq 0$, write

$$\frac{1}{r_i} = a_{i+1} + r_{i+1}, \qquad a_{i+1} = \left\lfloor \frac{1}{r_i} \right\rfloor \ge 1, \qquad 0 \le r_{i+1} < 1.$$

Eventually get some $r_n = 0$. Write

$$\frac{p}{q} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}.$$

Example.

$$\frac{40}{19} = 2 + \frac{2}{19}, \qquad \frac{19}{2} = 9 + \frac{1}{2} \qquad \Longrightarrow \qquad \frac{40}{19} = 2 + \cfrac{1}{9 + \cfrac{1}{2+0}}.$$

9.2 Infinite continued fractions

Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. As above, set $a_0 = \lfloor \alpha \rfloor$, write

$$\alpha = a_0 + r_0, \qquad a_0 = \lfloor \alpha \rfloor, \qquad 0 \le r_0 < 1.$$

Define sequences a_i, r_i by

$$\frac{1}{r_i} = a_{i+1} + r_{i+1}, \qquad a_{i+1} = \left\lfloor \frac{1}{r_i} \right\rfloor \in \mathbb{Z}, \qquad 0 \le r_{i+1} < 1.$$

By definition, $a_i \ge 1$ if $i > 0$. Write

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots}}.$$

Example. Let $\alpha = \sqrt{3}$.

$$a_0 = 1, \qquad r_0 = \sqrt{3} - 1, \qquad \frac{1}{r_0} = \frac{1}{\sqrt{3}-1} = \frac{\sqrt{3}+1}{2} = 1 + \frac{\sqrt{3}-1}{2}.$$

$$a_1 = 1, \qquad r_1 = \frac{\sqrt{3}-1}{2}, \qquad \frac{1}{r_1} = \frac{2}{\sqrt{3}-1} = \sqrt{3}+1 = 2 + \left(\sqrt{3}-1\right).$$

$$a_2 = 2, \qquad r_2 = \sqrt{3} - 1 = r_0, \qquad \frac{1}{r_2} = \frac{1}{\sqrt{3}-1} = \frac{1}{r_0},$$

so

$$a_i = \begin{cases} 1 & i > 0 \text{ odd} \\ 0 & i > 0 \text{ even} \end{cases}.$$

If $a_0, \ldots, a_n \in \mathbb{R}$, then

$$[a_0; a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}.$$

Lemma 68. If $a_0, \ldots, a_n \in \mathbb{R}$, define p_i, q_i for $0 \leq i \leq n$ by

$$p_0 = a_0, \qquad q_0 = 1, \qquad p_1 = a_0 a_1 + 1, \qquad q_1 = a_1, \qquad p_i = a_i p_{i-1} + p_{i-2}, \qquad q_i = a_i q_{i-1} + q_{i-2}.$$

Assuming that no $q_i = 0$, we have

$$[a_0; a_1, \ldots, a_n] = \frac{p_n}{q_n}.$$

Proof. Induction on n.

$n = 0$ $a_0 = a_0/1$ is trivial.

$n = 1$ $a_0 + 1/a_1 = (a_0 a_1 + 1)/a_1$ is trivial.

$n > 1$ Define sequences p_i', q_i' for $0 \leq i \leq n-1$ by applying the definition to the sequence

$$a_0, \ldots, a_{n-2}, a_{n-1} + \frac{1}{a_n}.$$

By definition, $p_i' = p_i$ and $q_i' = q_i$ if $i \leq n-2$. By induction,

$$\left[a_0; a_1, \ldots, a_{n-2}, a_{n-1} + \frac{1}{a_n}\right] = \frac{p_{n-1}'}{q_{n-1}'}.$$

By definition,

$$[a_0; a_1, \ldots, a_n] = \left[a_0; a_1, \ldots, a_{n-2}, a_{n-1} + \frac{1}{a_n}\right].$$

So we only need to show that $p_{n-1}'/q_{n-1}' = p_n/q_n$.

$$\begin{aligned}\frac{p_{n-1}'}{q_{n-1}'} &= \frac{\left(a_{n-1} + \frac{1}{a_n}\right) p_{n-2}' + p_{n-3}'}{\left(a_{n-1} + \frac{1}{a_n}\right) q_{n-2}' + q_{n-3}'} = \frac{\left(a_{n-1} + \frac{1}{a_n}\right) p_{n-2} + p_{n-3}}{\left(a_{n-1} + \frac{1}{a_n}\right) q_{n-2} + q_{n-3}} \\ &= \frac{(a_n a_{n-1} + 1) p_{n-2} + a_n p_{n-3}}{(a_n a_{n-1} + 1) q_{n-2} + a_n q_{n-3}} = \frac{a_n (a_{n-1} p_{n-2} + p_{n-3}) + p_{n-2}}{a_n (a_{n-1} q_{n-2} + q_{n-3}) + q_{n-2}} = \frac{a_n p_{n-1} + p_{n-2}}{a_n q_{n-1} + q_{n-2}} = \frac{p_n}{q_n}.\end{aligned}$$

□

Suppose now that $a_i \geq 1$ if $i \geq 1$. Then $q_i = a_i q_{i-1} + q_{i-2} \geq q_{i-1} + q_{i-2}$. So the q_i form an increasing sequence, in fact with $q_i \geq q_{i-1} + q_{i-2} \geq 2q_{i-2}$, so it even increases exponentially. If $a_0, a_1, \cdots \in \mathbb{R}$ is an infinite sequence with $a_i \geq 1$ for all i, say that p_i/q_i is the ***i*-th convergent** to

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots}}.$$

Lemma 69. For all n,

$$p_n q_{n-1} - q_n p_{n-1} = (-1)^{n-1}.$$

Proof. Obvious for $n = 1$. For inductive step,

$$\begin{aligned}p_n q_{n-1} - q_n p_{n-1} &= (a_n p_{n-1} + p_{n-2}) q_{n-1} - (a_n q_{n-1} + q_{n-2}) p_{n-1} \\ &= p_{n-2} q_{n-1} - q_{n-2} p_{n-1} \\ &= -(p_{n-1} q_{n-2} - q_{n-1} p_{n-2}).\end{aligned}$$

□

Note that if $a_i \in \mathbb{Z}$, then $p_i, q_i \in \mathbb{Z}$, and Lemma 69 gives $(p_n, q_n) = 1$.
In general, Lemma 69 gives

$$\left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{1}{q_n q_{n-1}}.$$

If $a_i \geq 1$ for all $i \geq 1$, then the sequence q_i increases exponentially. So

$$\sum_{i=1}^{n} \frac{1}{q_i q_{i-1}}$$

converges, so that (p_n/q_n) is a Cauchy sequence, so it converges.

Lecture 20
Tuesday
20/11/18

Lemma 70. Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and let $[a_0; a_1, a_2, \ldots]$ be the corresponding continued fraction. Then $p_n/q_n < \alpha$ if n is even, and $p_n/q_n > \alpha$ if n is odd.

Proof. Induction on n. $n = 0$ is $a_0 = \lfloor \alpha \rfloor < \alpha$ and $p_0/q_0 = a_0/1 = a_0$. If n is odd, then by induction, we have $[a_1; a_2, \ldots, a_n] < 1/(\alpha - a_0)$, since $\alpha = a_0 + 1/\ldots$. That is,

$$\alpha - a_0 < \frac{1}{[a_1; a_2, \ldots, a_n]} \qquad \iff \qquad \alpha < a_0 + \frac{1}{[a_1; a_2, \ldots, a_n]} = [a_0; a_1, \ldots, a_n] = \frac{p_n}{q_n}.$$

If n is even, same argument with $>$. □

Corollary 71. Assume $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $a_0, a_1, \cdots \in \mathbb{Z}$ be coming from its continued fraction. Let $p_n/q_n = [a_0; a_1, \ldots, a_n]$ be the n-th convergent. Then $|\alpha - p_n/q_n| < 1/q_n q_{n+1}$. In particular, $p_n/q_n \to \alpha$ as $n \to \infty$.

Proof. Either $p_n/q_n < \alpha < p_{n+1}/q_{n+1}$ or $p_n/q_n > \alpha > p_{n+1}/q_{n+1}$, by Lemma 70. Either way, $|p_n/q_n - \alpha| < |p_n/q_n - p_{n+1}/q_{n+1}| \leq 1/q_n q_{n+1}$, by Lemma 69. □

Note that $1/q_n q_{n+1} < 1/q_n^2$, so the sequence (p_n/q_n) satisfies the requirements of Dirichlet's theorem.

9.3 Best approximations

Fix $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Define a_i, r_i by

$$\alpha = a_0 + r_0, \qquad a_0 = \lfloor \alpha \rfloor \in \mathbb{Z}, \qquad 0 < r_0 < 1,$$

If $i \geq 1$,

$$\frac{1}{r_i} = a_{i+1} + r_{i+1}, \qquad a_{i+1} = \left\lfloor \frac{1}{r_i} \right\rfloor \in \mathbb{Z}_{\geq 1}, \qquad 0 < r_{i+1} < 1.$$

Lemma 72. For all n,

$$\alpha = \frac{p_n + p_{n-1} r_n}{q_n + q_{n-1} r_n}.$$

Proof. $\alpha = [a_0; a_1, \ldots, a_n, 1/r_n]$. Set $p_{n+1} = p_n/r_n + p_{n-1}$, $q_{n+1} = q_n/r_n + q_{n-1}$. Then by Lemma 68, $\alpha = p_{n+1}/q_{n+1}$. □

Corollary 73. For all n, $|\alpha q_n - p_n| < |\alpha q_{n-1} - p_{n-1}|$, and $|\alpha - p_n/q_n| < |\alpha - p_{n-1}/q_{n-1}|$.

Proof. By Lemma 72, $\alpha(q_n + q_{n-1} r_n) = p_n + p_{n-1} r_n$, so $\alpha q_n - p_n = r_n(p_{n-1} - \alpha q_{n-1})$. So $|\alpha q_n - p_n| = r_n |\alpha q_{n-1} - p_{n-1}| < |\alpha q_{n-1} - p_{n-1}|$.

$$\left|\alpha - \frac{p_n}{q_n}\right| = \frac{1}{q_n} |\alpha q_n - p_n| < \frac{1}{q_n} |\alpha q_{n-1} - p_{n-1}| < \frac{1}{q_{n-1}} |\alpha q_{n-1} - p_{n-1}| = \left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right|.$$

□

Theorem 74. Let $h, k \in \mathbb{Z}$ and $0 < |k| < q_{n+1}$. Then $|k\alpha - h| \geq |\alpha q_n - p_n|$, with equality only if $|k| = q_n$. If $|k| \leq q_n$, then $|h/k - \alpha| \geq |p_n/q_n - \alpha|$, with equality if and only if $h/k = p_n/q_n$.

Proof. By Lemma 69 there exist $u, v \in \mathbb{Z}$ such that $h = up_n + vp_{n+1}$ and $k = uq_n + vq_{n+1}$, since

$$\begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \qquad \iff \qquad \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{pmatrix}^{-1} \begin{pmatrix} h \\ k \end{pmatrix} = \frac{1}{(-1)^n} \begin{pmatrix} q_{n+1} & -p_{n+1} \\ -q_n & p_n \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}.$$

By assumption, $0 < |k| < q_{n+1}$. So $u \neq 0$, else $k = vq_{n+1}$, so $|v| < 1$ is a contradiction. If $v \neq 0$, then u, v have opposite signs, else $|k| = |uq_n| + |vq_{n+1}| \geq q_n + q_{n+1} > q_{n+1}$. If $v = 0$, then $h = up_n$, $k = uq_n$, and everything is easy. If $v \neq 0$, then write $k\alpha - h = u(\alpha q_n - p_n) + v(\alpha q_{n+1} - p_{n+1})$. u, v have opposite signs. By Lemma 70, $\alpha q_n - p_n, \alpha q_{n+1} - p_{n+1}$ also have opposite signs. So $u(\alpha q_n - p_n)$ and $v(\alpha q_{n+1} - p_{n+1})$ have the same sign. So $|k\alpha - h| = |u(\alpha q_n - p_n)| + |v(\alpha q_{n+1} - p_{n+1})| > |\alpha q_n - p_n|$ if $u, v \neq 0$. For the last part, if $|k| \leq q_n$ then $1/|k| \geq 1/q_n$. So

$$\frac{1}{|k|} |k\alpha - h| \geq \frac{1}{q_n} |q_n\alpha - p_n|,$$

that is $|\alpha - h/k| \geq |\alpha - p_n/q_n|$. □

Corollary 75. If $h, k \in \mathbb{Z}$ with $|\alpha - h/k| < 1/2k^2$, then $h/k = p_n/q_n$ for some n.

Proof. Without loss of generality $k \geq 1$, and $q_n \leq k < q_{n+1}$ for some n. Then

$$\begin{aligned} \left|\frac{p_n}{q_n} - \frac{h}{k}\right| &\leq \left|\frac{p_n}{q_n} - \alpha\right| + \left|\alpha - \frac{h}{k}\right| = \frac{1}{q_n} |\alpha q_n - p_n| + \frac{1}{k} |\alpha k - h| \\ &\leq \left(\frac{1}{q_n} + \frac{1}{k}\right) |\alpha k - h| = k\left(\frac{1}{q_n} + \frac{1}{k}\right) \left|\alpha - \frac{k}{h}\right| < \frac{1}{2k} \left|\frac{1}{q_n} + \frac{1}{k}\right| \leq \frac{1}{kq_n}, \end{aligned}$$

by Theorem 74. So $|p_n/q_n - h/k| < 1/kq_n$. So $p_n/q_n - h/k = 0$, as required. □

Lecture 21
Wednesday
21/11/18

9.4 Returning to Pell's equation

Pell's equation is $X^2 - dY^2 = 1$. If (x, y) is a solution, then $\left|\sqrt{d} - x/y\right|$ is small.

Proposition 76. Let $d > 1$ be squarefree, and let p_n/q_n be the sequence of convergents for the continued fraction for $\sqrt{d}$. If $x, y > 0$ with $x^2 - dy^2 = \pm 1$, then $x = p_n$ and $y = q_n$ for some n.

Proof.

- Firstly suppose $x^2 - dy^2 = 1$. It is enough to show that $x/y = p_n/q_n$ for some n. Since $(p_n, q_n) = 1$, this implies that $x = rp_n$ and $y = rq_n$ for some r, and then $1 = x^2 - dy^2 = r^2\left(p_n^2 - dq_n^2\right)$, so $r = 1$. By Corollary 75, it suffices to prove that $\left|\sqrt{d} - x/y\right| < 1/2y^2$. $x - y\sqrt{d} = 1/\left(x + y\sqrt{d}\right) > 0$. So $x > y\sqrt{d}$, and $x/y > \sqrt{d}$. So

$$\left|\frac{x}{y} - \sqrt{d}\right| = \frac{x}{y} - \sqrt{d} = \frac{1}{y}\left(x - y\sqrt{d}\right) = \frac{1}{y}\left(\frac{1}{x + y\sqrt{d}}\right) < \frac{1}{y}\left(\frac{1}{y\sqrt{d} + y\sqrt{d}}\right) = \frac{1}{\left(2\sqrt{d}\right)y^2} < \frac{1}{2y^2}.$$

- Now assume $x^2 - dy^2 = -1$. Again enough to show that $x/y = p_n/q_n$. Trick is to rewrite as $y^2 - x^2/d = 1/d$. Then $y - x/\sqrt{d} = (1/d)/\left(y + x/\sqrt{d}\right) > 0$. So $y > x/\sqrt{d}$.

$$\left|\frac{y}{x} - \frac{1}{\sqrt{d}}\right| = \frac{y}{x} - \frac{1}{\sqrt{d}} = \frac{1}{x}\left(y - \frac{x}{\sqrt{d}}\right) = \frac{1}{x}\left(\frac{1/d}{y + x/\sqrt{d}}\right) < \frac{1}{x}\left(\frac{1/d}{x/\sqrt{d} + x/\sqrt{d}}\right) = \frac{1/\sqrt{d}}{2x^2} < \frac{1}{2x^2}.$$

  So Corollary 75 gives that y/x is a convergent for the continued fraction of $1/\sqrt{d}$. $\left\lfloor 1/\sqrt{d} \right\rfloor = 0$, so the continued fraction for $1/\sqrt{d}$ is of the form $[0; a_0, a_1, \ldots]$. Next step is $1/\left(1/\sqrt{d}\right) = \sqrt{d}$. So if

$\sqrt{d} = [a_0; a_1, a_2, \ldots]$ then $1/\sqrt{d} = [0; a_0, a_1, \ldots]$, since

$$\sqrt{d} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots}}, \qquad \frac{1}{\sqrt{d}} = 0 + \cfrac{1}{a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots}}}.$$

So the convergents for $1/\sqrt{d}$ are the q_n/p_n. So $y/x = q_n/p_n$ for some n, and $x/y = p_n/q_n$.

□

Example.

- $\sqrt{3} = [1; 1, 2, 1, 2, \ldots] = \left[1; \overline{1, 2}\right]$.
- $\sqrt{2} = 1 + \left(\sqrt{2} - 1\right)$, $1/\left(\sqrt{2} - 1\right) = \sqrt{2} + 1 = 2 + \left(\sqrt{2} - 1\right)$, so $\sqrt{2} = \left[1; \overline{2}\right]$.
- $\sqrt{5} = 2 + \left(\sqrt{5} - 2\right)$, $1/\left(\sqrt{5} - 2\right) = \sqrt{5} + 2 = 4 + \left(\sqrt{5} - 2\right)$, so $\sqrt{5} = \left[2; \overline{4}\right]$.
- $\sqrt{7} = \left[2; \overline{1, 1, 1, 4}\right]$.
- $\sqrt{13} = \left[3; \overline{1, 1, 1, 1, 6}\right]$.
- $\sqrt{43} = \left[6; \overline{1, 1, 3, 1, 5, 1, 3, 1, 1, 12}\right]$.
- $\sqrt{n^2 + 1} = \left[n; \overline{2n}\right]$.

Definition 77. We say that $[a_0; a_1, a_2, \ldots]$ is **eventually periodic** if there exist $N, d > 0$ such that $a_{n+d} = a_n$ for all $n \geq N$. We say that it is **periodic** if we can take $N = 0$.

Remark 78. The following are facts.

- The continued fraction of $\sqrt{d}$ is eventually periodic.
- In fact, it is of the form $\left[a_0; \overline{a_1, \ldots, a_{m-1}, 2a_0}\right]$.
- $a_1, \ldots, a_{m-1}$ is symmetric, that is $a_i = a_{m-i}$ for $1 \leq i \leq m - 1$.
- The n for which $p_n^2 - dq_n^2 = \pm 1$ are exactly the n for which $n \equiv -1 \mod m$. If $n = lm - 1$, then $p_n^2 - dq_n^2 = (-1)^{lm}$.
- The fundamental 1-unit is $p_{m-1} + q_{m-1}\sqrt{d}$ if m is even and $p_{2m-1} + q_{2m-1}\sqrt{d}$ if m is odd.
- There is a solution to $x^2 - dy^2 = -1$ if and only if m is odd, in which case the solutions are $(x, y) = (p_n, q_n)$ with $n \equiv m - 1 \mod 2m$.

Example.

- Let $x^2 - 43y^2 = \pm 1$. $m = 10$ is even, so no solutions to $x^2 - 43y^2 = -1$. Smallest solution for $x^2 - 43y^2 = 1$ is p_9, q_9.

i	0	1	2	3	4	5	6	7	8	9
a	6	1	1	3	1	5	1	3	1	1
p	6	7	13	46	59	341	400	1541	1941	3482
q	1	1	2	7	9	52	61	235	296	531

$p_9 = 3482$, so $3482^2 - 43\,(531)^2 = 1$ is the smallest solution.

- For 13, $m = 5$ so p_4, q_4 is the smallest solution for $x^2 - 13y^2 = -1$ and p_9, q_9 is the smallest solution for $x^2 - 13y^2 = 1$.

i	0	1	2	3	4
a	3	1	1	1	1
p	3	4	7	11	18
q	1	1	2	3	5

$18^2 - 13\,(5)^2 = -1$ is the smallest solution. $N\left(18 + 5\sqrt{13}\right) = -1$, so $N\left(\left(18 + 5\sqrt{13}^2\right)\right) = 1$. In fact, it follows from our facts that this is the fundamental 1-unit, that is $p_9 + q_9\sqrt{13}$. $\left(18 + 5\sqrt{13}\right)^2 = 649 + 180\sqrt{13}$.

Lecture 22
Friday
23/11/18

9.5 Periodic continued fractions

Definition 79. $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is a **quadratic irrational** if it is a root of some $ax^2 + bx + c = 0$ for $a, b, c \in \mathbb{Q}$ not all zero.

Proposition 80. If α has an eventually periodic continued fraction, then α is a quadratic irrational.

Proof. Suppose firstly that the continued fraction of α is periodic. Suppose $a_{n+d} = a_n$ for all n, for some $d \geq 1$. Then

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_{d-1} + \cfrac{1}{\alpha}}}}.$$

This gives an equation of the form $\alpha = (x\alpha + y)\,/\,(z\alpha + w)$ for $w, x, y, z \in \mathbb{Z}$, by applying Lemma 72 to $[a_0; a_1, \ldots, a_{d-1}, \alpha]$. Then $(z\alpha + w)\,\alpha - (x\alpha + y) = 0$, that is $z\alpha^2 + (w - x)\,\alpha - y = 0$. Since $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, we conclude that α is a quadratic irrational. Suppose now that α is only eventually periodic. Then

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_N + \cfrac{1}{\beta}}}},$$

where β has periodic continued fraction. So β is a quadratic irrational. To complete the proof, we need to show that if γ is a quadratic irrational, then $1/\gamma, \gamma + n$ are quadratic rationals for any $n \in \mathbb{Z}$. If γ is a root of $aX^2 + bX + c = 0$, then $1/\gamma$ is a root of $cX^2 + bX + a = 0$ and $\gamma + n$ is a root of $a\,(X - n)^2 + b\,(X - n) + c = 0$. □

In fact, the converse is also true. All quadratic irrationals have eventually periodic continued fractions.

10 Diophantine approximation

10.1 Liouville's theorem

Definition 81. Let $d \in \mathbb{Z}_{\geq 1}$. Then $\alpha \in \mathbb{C}$ is **algebraic of degree** d if there exists a polynomial of degree d with integer coefficients and α as a root. There does not exist such a polynomial of smaller degree.

Example.

- $d = 1$ is $\mathbb{Q}$.
- $d = 2$ is quadratic irrationals.

Theorem 82 (Liouville's theorem)**.** Let $\alpha \in \mathbb{R}$ be algebraic of degree d. Then for any $e \in \mathbb{R}_{>d}$, there are only finitely many $p/q \in \mathbb{Q}$ with

$$\left|\frac{p}{q} - \alpha\right| < \frac{1}{q^e}.$$

Proof. Let $P(x)$ be a polynomial of degree d with coefficients in $\mathbb{Z}$, with $P(\alpha) = 0$. Choose $\epsilon > 0$ such that the only root of $P(x)$ in $[\alpha - \epsilon, \alpha + \epsilon]$ is α. Write $P(x) = (x - \alpha) Q(x)$. $Q(x)$ is a polynomial of degree $d - 1$ with real coefficients, so in particular it is continuous, so there exists K such that $|Q(x)| \leq K$ for all $x \in [\alpha - \epsilon, \alpha + \epsilon]$. Assume that $|\alpha - p/q| < 1/q^e$. We may assume that q is large enough that $1/q^e < \epsilon$. Since P has integer coefficients and is of degree d, we have $|P(p/q)| \geq 1/q^d$. Note that $P(p/q) \neq 0$, or we could replace P by P' with $P(x) = (qx - p) P'(x)$. Since $|p/q - \alpha| < 1/q^e < \epsilon$, $p/q \in [\alpha - \epsilon, \alpha + \epsilon]$, so

$$\left|P\left(\frac{p}{q}\right)\right| = \left|\frac{p}{q} - \alpha\right| \left|Q\left(\frac{p}{q}\right)\right| \leq K \left|\frac{p}{q} - \alpha\right| < \frac{K}{q^e}.$$

So $K/q^e > |P(p/q)| \geq 1/q^d$, so $K > q^{e-d}$, so $K^{1/(e-d)} > q$. So there are only finitely many possible q, so only finitely many p/q. □

10.2 Constructing transcendentals

Recall that $\alpha \in \mathbb{C}$ is **algebraic** if it is algebraic of degree d, and otherwise it is called **transcendental**. The set of polynomials with integer coefficients is countable, so the set of algebraic numbers is countable. Since $\mathbb{R}$ is uncountable, transcendental numbers exist. Liouville's theorem gives a criterion, if for every $e > 0$, there are infinitely many p/q with $|\alpha - p/q| < 1/q^e$, then α cannot be algebraic.

Example. Let

$$\alpha = \sum_{n \geq 1} \frac{1}{10^{n!}}, \qquad \alpha_k = \sum_{n=1}^{k} \frac{1}{10^{n!}}.$$

$\alpha_k \in \mathbb{Q}$ has denominator $q = 10^{k!}$.

$$\begin{aligned} |\alpha - \alpha_k| &= \sum_{n=k+1}^{\infty} \frac{1}{10^{n!}} = \frac{1}{10^{(k+1)!}} \left(1 + \frac{1}{10^{(k+2)!-(k+1)!}} + \frac{1}{10^{(k+3)!-(k+1)!}} + \dots\right) \\ &< \frac{1}{10^{(k+1)!}} \left(1 + \frac{1}{10} + \frac{1}{100} + \dots\right) < \frac{2}{10^{(k+1)!}} = \frac{2}{q^{k+1}}. \end{aligned}$$

If $d \in \mathbb{Z}_{>0}$, and $k > d$, then $2/q^{k+1} < 1/q^d$. So there exists infinitely many $p/q = \alpha_k$ such that $|\alpha - p/q| < 1/q^d$. Taking d arbitrarily large, α is transcendental.

Lecture 23
Tuesday
27/11/18

10.3 Roth's theorem

Theorem 83 (Roth's theorem)**.** Suppose that α is algebraic. Then for any $\epsilon > 0$, there exist only finitely many $x/y \in \mathbb{Q}$ with

$$\left|\alpha - \frac{x}{y}\right| < \frac{1}{y^{2+\epsilon}}.$$

This can be used to show that many more numbers are transcendental than Liouville's theorem could.

Example. $\sum_{n\geq 1} 1/10^{3^n}$ is transcendental.

Example. We saw that if $d > 1$ is squarefree, then $x^2 - dy^2 = 1$ has infinitely many solutions with $x, y \in \mathbb{Z}$. Suppose now that $d > 1$, and consider $x^3 - dy^3 = 1$.

- $d = e^3$ is a cube. Then $x^3 - dy^3 = x^3 - (ey)^3 = 1$, so either $(x, y) = (1, 0)$ or $(x, y) = (0, \pm 1)$ and $d = 1$.
- d is not a cube. Then $\sqrt[3]{d} \in \mathbb{R} \setminus \mathbb{Q}$ is algebraic, as it is a root of $X^3 - d = 0$. Suppose $x > 1$, so $x > \sqrt[3]{d}y$. Then

$$x - \sqrt[3]{d}y = \frac{x^3 - dy^3}{x^2 + x\sqrt[3]{d}y + \sqrt[3]{d^2}y^2} = \frac{1}{x^2 + x\sqrt[3]{d}y + \sqrt[3]{d^2}y^2} < \frac{1}{3\left(\sqrt[3]{d}y\right)^2} = \frac{1}{3\sqrt[3]{d^2}y^2}.$$

So $\left|x/y - \sqrt[3]{d}\right| < 1/3\sqrt[3]{d^2}y^3$. Choose any $0 < \epsilon < 1$. Then $3\sqrt[3]{d^2}y^3 < 1/y^{2+\epsilon}$ for any y sufficiently large. So Roth's theorem tells us that there are only finitely many solutions. Similarly if $x < 0$.

11 Primes in arithmetic progressions

Question is how are the prime numbers distributed mod n? Are there infinitely many primes congruent to $a \mod n$ for each a, n? Answer is no in general.

Example. There are finitely many primes congruent to $2 \mod 4$, or $0 \mod 2$.

If $(a, n) \neq 1$ then since any number is congruent to $a \mod n$ is divisible by (a, n), we can have at most one prime. If $(a, n) = 1$, there is no obvious obstruction.

Example. There are infinitely many primes congruent to $1 \mod 2$.

Theorem 84. If $(a, n) = 1$, then there are infinitely many primes congruent to $a \mod n$.

We will prove this for $a = 1$.

11.1 Elementary results

Theorem 85. There are infinitely many primes.

Proof. Let S be a finite set of primes, and let $Q = 1 + \prod_{p \in S} p$. Then $Q > 1$, so it has a prime factor q. Then $q \notin S$, so we are done. □

Theorem 86. There are infinitely many primes congruent to $3 \mod 4$.

Proof. Let S be a finite set of primes which are congruent to $3 \mod 4$. Let $Q = 2 + \prod_{p \in S} p^2$. Then $Q > 1$, and $Q \equiv 3 \mod 4$, so Q has a prime factor q which has $q \equiv 3 \mod 4$. Then $q \notin S$, so we are done. □

Lemma 87. Let x be even, and p a prime factor of $x^2 + 1$, then $p \equiv 1 \mod 4$.

Proof. Certainly p is odd. $x^2 + 1 \equiv 0 \mod p$, so $x^2 \equiv -1 \mod p$, so $\left(\frac{-1}{p}\right) = 1$, so $p \equiv 1 \mod 4$. □

Theorem 88. There are infinitely many primes congruent to $1 \mod 4$.

Proof. Let S be a finite set of primes congruent to $1 \mod 4$. Let $Q = 1 + 4\prod_{p \in S} p^2 = 1 + \left(2\prod_{p \in S} p\right)^2$. Then $Q > 1$, and if q is a prime factor of Q then $q \notin S$, and $q \equiv 1 \mod 4$ by Lemma 87. □

General ideal is to find a polynomial $P(x)$ such that every prime factor of $P(nx)$ is congruent to $a \mod n$, or at least one. Turns out that this can be done only when $a^2 \equiv 1 \mod n$. We will find such polynomials for $a = 1$.

Theorem 89. For any prime q, there are infinitely many primes congruent to $1 \mod q$.

Definition 90. The q-th **cyclotomic polynomial** is $\Phi_q(X) = (X^q - 1)/(X - 1) = X^{q-1} + \cdots + 1$.

Theorem 91. Let $p \neq q$ be prime, and let $a \in \mathbb{Z}$. Then $p \mid \Phi_q(a)$ if and only if a has order $q \mod p$.

Proof. a has order $q \mod p$ if and only if $a^q \equiv 1 \mod p$ and $a \not\equiv 1 \mod p$. If $p \mid \Phi_q(a)$ then $p \mid a^q - 1$. If also $a \equiv 1 \mod p$, then $\Phi_q(a) \equiv \Phi_q(1) \equiv q \not\equiv 0 \mod p$, a contradiction. Conversely if $a^q \equiv 1 \mod p$, $a \not\equiv 1 \mod p$, then $(a^q - 1)/(a - 1) \equiv 0 \mod p$. □

Lecture 24
Wednesday
28/11/18

Corollary 92. If $p \neq q$ is prime, and $a \in \mathbb{Z}$, and $p \mid \Phi_q(a)$, then $p \equiv 1 \mod q$.

Proof. By Theorem 91, a has order $q \mod p$. But $a^{p-1} \equiv 1 \mod p$, by Fermat's little theorem. So $q \mid p - 1$. □

Theorem 93. Let q be prime. Then there are infinitely many primes with $p \equiv 1 \mod q$.

Proof. Let S be a finite set of primes which are congruent to $1 \mod q$. Let $R = \prod_{p \in S} p$. Consider $\Phi_q(qR) \geq qR + 1 > 1$. Let p be a prime factor of $\Phi_q(qR)$. By Corollary 92, either $p = q$, or $p \equiv 1 \mod q$. Since $\Phi_q(qR) = (qR)^{q-1} + \cdots + 1 \equiv 1 \mod qR$, so $p \neq q$, $p \notin S$, and $p \equiv 1 \mod q$. □

11.2 Cyclotomic polynomials

Definition 94. Let $n \in \mathbb{Z}_{\geq 1}$. Then

$$\Phi_n(X) = \prod_{1 \leq a \leq n,\ (a,n)=1} \left(X - e^{\frac{2\pi a i}{n}}\right).$$

Lemma 95. For any n, we have

$$X^n - 1 = \prod_{d|n,\ d>0} \Phi_d(X).$$

Proof. Each side is a monic polynomial, so we just need to check that the roots are the same, with multiplicities. Left hand side are the n-th roots of unity, with multiplicity one each. Right hand side is Φ_d, the primitive d-th roots of unity, with multiplicity one. Each n-th root of unity is a primitive d-th root of unity for some unique $d \mid n$. The result follows. □

From this it is easy to deduce the following.

Lemma 96. For any $n \geq 1$, $\Phi_n(X) \in \mathbb{Z}[X]$.

Proof. By induction on n. If $n = 1$, $\Phi_1(X) = X - 1$. Assume that the result holds for all $d \mid n$, $d < n$. By Lemma 95, if we set $P(X) = \prod_{d|n,\ 0<d<n} \Phi_d(X)$, then $P(X) \in \mathbb{Z}[X]$, $P(X)$ is monic, and $X^n - 1 = \Phi_n(X) P(X)$. Write $\Phi_n(X) = \sum_i a_i X^i$, $P(X) = \sum_i b_i X^i$, and assume that not all $a_i \in \mathbb{Z}$. Let q be maximal with $a_q \notin \mathbb{Z}$. Let $e = \deg(P)$, so $P(X) = X^e + b_{e-1}X^{e-1} + \cdots + a_0$. Then the coefficient of X^{q+e} in $\Phi_n(X) P(X)$ is $a_q + a_{q+1}b_{e-1} + \cdots + a_{q+e}b_0$, where $a_{q-1}b_{e-1} + \cdots + a_{q+e}b_0 \in \mathbb{Z}$. Since $\Phi_n(X) P(X) = X^n - 1 \in \mathbb{Z}[X]$, this is a contradiction. □

Definition 97. Let F be any field, and let $P(X) \in F[X]$. Then $P'(X)$, the **derivative** of $P(X)$, is defined as follows. If $P(X) = \sum_{n=0}^{d} a_n X^n$, then $P'(X) = \sum_{n=1}^{d} n a_n X^{n-1}$.

Note that $(P+Q)' = P' + Q'$ and $(PQ)' = P'Q + PQ'$.

Lemma 98. Suppose that $(X-\alpha)^2$ divides $P(X)$. Then α is a root of both P and P'.

Proof. Write $P(X) = (X-\alpha)^2 R(X)$. Then

$$P'(X) = (X-\alpha)^2 R'(X) + 2(X-\alpha) R(X) = (X-\alpha)\left((X-\alpha) R'(X) + 2R(X)\right).$$

□

Corollary 99. If $p \nmid n$, then $\Phi_n(X)$ has no repeated roots mod p.

Proof. It suffices to show that $X^n - 1$ has no repeated roots mod p. The derivative of $X^n - 1$ is nX^{n-1}, so its only root is zero, which is not a root of $X^n - 1$. So we are done by Lemma 98. □

Note that if $n = p$, $X^p - 1 \equiv (X-1)^p \mod p$ and $\Phi_p(X) \equiv (X-1)^{p-1} \mod p$.

Theorem 100. Suppose $p \nmid n$ and $a \in \mathbb{Z}$. Then $p \mid \Phi_n(a)$ if and only if a has order exactly $n \mod p$.

Proof. Firstly suppose that a has order exactly n. Then a is a root of $X^n - 1 \mod p$, but not a root of $X^d - 1$ for any $d \mid n$, $d < n$. Since $\Phi_d(X) \mid X^d - 1$, a cannot be a root of $\Phi_d(X)$ for any $d \mid n$, $d < n$.

$$X^n - 1 = \Phi_n(X) \prod_{d|n,\ 0<d<n} \Phi_d(X), \tag{4}$$

so a is a root of $\Phi_n(X) \mod p$, that is $p \mid \Phi_n(a)$. Conversely, suppose that $p \mid \Phi_n(a)$. Then a is a root of $\Phi_n(X) \mod p$, so by 4, a is a root of $X^n - 1 \mod p$. We need to show that a is not a root of $X^d - 1$ for any $d \mid n$, $d < n$. Writing $X^d - 1 = \prod_{e|d} \Phi_e(X)$, a would be a root of $\Phi_e(X)$ for some $e \mid d \mid n$. So by (4), a is a root of both $\Phi_n(X)$ and $\Phi_e(X)$, so a is a repeated root of $X^n - 1 \mod p$. This contradicts Corollary 99. □

Corollary 101. If $p \nmid n$, and $a \in \mathbb{Z}$, then if $p \mid \Phi_n(a)$, then $p \equiv 1 \mod n$.

Proof. a has order $n \mod p$ by Theorem 100, so $n \mid p - 1$, by Fermat's little theorem. □

11.3 Primes congruent to $1 \mod n$

We are now in a position to prove the following.

Theorem 102. If $n \in \mathbb{Z}_{\geq 1}$, there are infinitely many primes p with $p \equiv 1 \mod n$.

Proof. Let S be a finite set of primes congruent to $1 \mod n$, and let $R = \prod_{p \in S} p$. For each k, let Q_k be $\Phi_n(knR) \in \mathbb{Z}$. Note that not all Q_k are ± 1, since $\Phi_n(X)$ is a non-constant polynomial. Thus choose k large enough that $Q_k > 1$, so there is a prime p dividing Q_k. Since Q_k divides $(knR)^n - 1$, no prime dividing n or R can divide Q_k. Thus p is not in S, and by Corollary 101 p is congruent to $1 \mod n$. □

Lecture 25 is a problem class.

Lecture 25
Friday
30/11/18

Lecture 26
Tuesday
04/12/18

12 Arithmetic functions

An arithmetic function is a function $f : \mathbb{Z}_{\geq 1} \to \mathbb{C}$, such as Φ.

12.1 Dirichlet convolution

The set of arithmetic functions is a ring in the following way. Addition is $(f+g)(n) = f(n) + g(n)$. Multiplication is **Dirichlet convolution** $f * g$,

$$(f*g)(n) = \sum_{d|n,\ d\geq 1} f(d)\, g\left(\frac{n}{d}\right) = \sum_{a,b\geq 1,\ ab=n} f(a)\, g(b).$$

We have $f*g = g*f$ and $f*(g*h) = (f*g)*h$, and both are given by

$$(f*g*h)(n) = \sum_{a,b,c\geq 1,\ abc=n} f(a)\, g(b)\, h(c).$$

$f*(g+h) = f*g + f*h$. There exists a multiplicative unit ϵ, that is $f*\epsilon = \epsilon * f = f$. This is easy to figure out. We need

$$f(n) = (f*\epsilon)(n) = \sum_{ab=n} f(a)\,\epsilon(b).$$

Example.

$$f(4) = f(4)\,\epsilon(1) + f(2)\,\epsilon(2) + f(1)\,\epsilon(4).$$

Forces $\epsilon(1) = 1$, $\epsilon(2) = \epsilon(4) = 0$. So

$$\epsilon(n) = \begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases}.$$

12.2 Möbius inversion

Möbius function $\mu : \mathbb{Z}_{\geq 1} \to \mathbb{C}$ is defined as follows. $\mu(1) = 1$. If $n = p_1 \dots p_k$ is a product of distinct prime factors, then $\mu(n) = (-1)^k$. Otherwise $\mu(n) = 0$.

Lemma 103. If 1 is the function $1(n) = 1$ for all n, then $1 * \mu = \epsilon$.

Proof. $\epsilon(1) = (1*\mu)(1) = 1 \times 1$. If $n > 1$, we just have to check that

$$\sum_{d|n} \mu(d) = \sum_{ab=n} 1(a)\,\mu(b) = 0.$$

Let $p_1, \dots, p_k$ be the distinct primes dividing n. Then

$$\sum_{d|n} \mu(d) = \sum_{(\epsilon_1,\dots,\epsilon_k),\ \epsilon_i\in\{0,1\}} (-1)^{\epsilon_1+\dots+\epsilon_k} = \left(\sum_{\epsilon_1=0}^{1} (-1)^{\epsilon_1}\right)\dots\left(\sum_{\epsilon_k=0}^{1} (-1)^{\epsilon_k}\right) = 0,$$

where $d = \prod_{i=1}^{k} p_i^{\epsilon_i}$. □

Proposition 104 (Möbius inversion)**.** If f and g are arithmetic functions then

$$g = f*1 \qquad \Longleftrightarrow \qquad f = g*\mu.$$

Proof. $(f*1)*\mu = f*(1*\mu) = f*\epsilon = f$, by Lemma 103, and $(g*\mu)*1 = g*(\mu*1) = g*\epsilon = g$. □

Example. Let $Id(n) = n$. Then $Id = \Phi * 1$. That is, $n = \sum_{d|n} \Phi(d)$. So $\Phi = Id * \mu$. So

$$\Phi(n) = \sum_{d|n} \mu(d)\,\frac{n}{d} = n\sum_{d|n} \frac{\mu(d)}{d}.$$

13 The distribution of prime numbers

Let $\pi(X)$ be the number of primes p such that $p \leq X$.

13.1 Reminder of asymptotic notation

- $A \ll B$ means there exists a constant $C > 0$ such that $|A| \leq CB$. For example if $x \geq 1$, $x \ll x^2 \ll e^x/x^{100}$.
- $B \gg A$ means $A \ll B$.
- $A = O(B)$ means $A \ll B$.
- $A \ll_k B$ means $A \ll B$ with the constant C depending on k. For example, $kx \ll_k x$.
- $A = o(B)$ means for all $\epsilon > 0$ we have $|A| \leq \epsilon B$ as some other specified parameter becomes large enough. For example, $1/\log x = o(1)$ as $x \to \infty$.
- $A \sim B$ means $A = (1 + o(1))B$.

13.2 The prime number theorem

Theorem 105 (Prime number theorem)**.**

$$\pi(X) \sim \frac{X}{\log X},$$

as $X \to \infty$.

Theorem 106. There exist constants $0 < c_1 < 1 < c_2$ such that for all sufficiently large X,

$$c_1 \frac{X}{\log X} \leq \pi(X) \leq c_2 \frac{X}{\log X}.$$

This gives

$$\pi(X) = O\left(\frac{X}{\log X}\right).$$

Lecture 27
Wednesday
05/12/18

Proof.

- Firstly consider the lower bound. We will prove that for some $C_1 > 1$, we have

$$\prod_{p \leq 2n} \geq C_1^n. \tag{5}$$

Given (5), we have

$$(2n)^{\pi(2n)} \geq \prod_{p \leq 2n} p \geq C_1^n.$$

Taking logarithms,

$$\pi(2n) \geq \left(\frac{1}{2}\log C_1\right)\frac{2n}{\log 2n}.$$

This gives the lower bound if $X = 2n \in \mathbb{Z}$ is even, but since $\pi(X+1) - \pi(X) \leq 1$, it is easy to get the lower bound for all X. We will prove (5) by considering the prime factors of

$$\binom{2n}{n} = \prod_{p \leq 2n} p^{v_p(n)}.$$

Claim that

1. if $p > \sqrt{2n}$ then $v_p(n) \leq 1$,
2. for all $p \leq 2n$, $p^{v_p(n)} \leq 2n$, and
3. $\prod_{p \leq 2n} p^{v_p(n)} \geq 4^n / (2n+1)$.

Suppose 1 to 3 are true. Then

$$\begin{aligned}
\frac{4^n}{2n+1} &\leq \prod_{p \leq 2n} p^{v_p(n)} && \text{by 3} \\
&= \prod_{p \leq \sqrt{2n}} p^{v_p(n)} \prod_{\sqrt{2n} < p \leq 2n} p^{v_p(n)} \\
&\leq (2n)^{\pi\left(\sqrt{2n}\right)} \prod_{\sqrt{2n} < p \leq 2n} p^{v_p(n)} && \text{by 2} \\
&\leq (2n)^{\pi\left(\sqrt{2n}\right)} \prod_{\sqrt{2n} < p \leq 2n} p && \text{by 1} \\
&\leq (2n)^{\pi\left(\sqrt{2n}\right)} \prod_{p \leq 2n} p \\
&\leq (2n)^{\sqrt{2n}} \prod_{p \leq 2n} p.
\end{aligned}$$

So

$$\prod_{p \leq 2n} p \geq \frac{4^n}{(2n+1)(2n)^{\sqrt{2n}}}.$$

(Exercise: show that for n sufficiently large, and any $4 > C_1$, the right hand side is at least C_1^n, that is if $K > 1$, $K^n \geq (2n+1)(2n)^{\sqrt{2n}}$ for all n sufficiently large)

1. In the first example sheet question 11, the exact power of p dividing $m!$ is $\sum_{i=1}^{\infty} \lfloor m/p^i \rfloor$. So

$$v_p(n) = \sum_{i=1}^{\infty} \left(\left\lfloor \frac{2n}{p^i} \right\rfloor - 2 \left\lfloor \frac{n}{p^i} \right\rfloor \right),$$

since

$$\binom{2n}{n} = \frac{(2n)!}{n!n!}.$$

For any $x \in \mathbb{R}$, $\lfloor 2x \rfloor - 2\lfloor x \rfloor \geq 0$, and in fact $\lfloor 2x \rfloor - 2\lfloor x \rfloor = 0$ or $\lfloor 2x \rfloor - 2\lfloor x \rfloor = 1$. If $p > \sqrt{2n}$, then $p^2 > 2n$, so all terms in the sum vanish if $i \geq 2$, so the sum is at most one.

2. Note that the terms in the sum are zero as soon as $p^i > 2n$, that is

$$i > \frac{\log 2n}{\log p}.$$

So

$$v_p(n) \leq \frac{\log 2n}{\log p},$$

that is $p^{v_p(n)} \leq 2n$.

3.

$$4^n = 2^{2n} = (1+1)^{2n} = \sum_{i=0}^{2n} \binom{2n}{i} \leq (2n+1)\binom{2n}{n},$$

so

$$\prod_{p \leq 2n} p^{v_p(n)} = \binom{2n}{n} \geq \frac{4^n}{2n+1}.$$

- Claim that there exists $C_2 > 1$ such that for all X sufficiently large, we have

$$\prod_{\frac{X}{2} \leq p \leq X} p \leq C_2^X. \tag{6}$$

Suppose we know (6). Then

$$C_2^X \geq \prod_{\frac{X}{2} \leq p \leq X} p \geq \left(\frac{X}{2}\right)^{\pi(X)-\pi\left(\frac{X}{2}\right)}.$$

Taking logarithms,

$$\pi(X) \leq \pi\left(\frac{X}{2}\right) + \frac{X \log C_2}{\log \frac{X}{2}}. \tag{7}$$

Suppose that X is large enough that (6) holds for $X, \ldots, X/2^{m-1}$. Substituting $X, \ldots, X/2^{m-1}$ into (7), and summing,

$$\pi(X) \leq \pi\left(\frac{X}{2^m}\right) + 2 \log C_2 \sum_{i=1}^{m} \frac{\frac{X}{2^i}}{\log \frac{X}{2^i}}.$$

Now fix X and choose m to be largest possible with $2^m \leq \sqrt{X}$. Then $X/2^m \geq \sqrt{X}$, so (6) is indeed valid for $X, \ldots, X/2^{m-1}$ provided that X is sufficiently large. Since m is maximal such that $2^m \leq \sqrt{X}$, we have $2^m \geq \sqrt{X}/2$. So

$$\pi\left(\frac{X}{2^m}\right) \leq \frac{X}{2^m} \leq 2\sqrt{X}.$$

So substituting into the above,

$$\pi(X) \leq 2\sqrt{X} + 2 \log C_2 \sum_{i=1}^{m} \frac{\frac{X}{2^i}}{\log \frac{X}{2^i}} \leq 2\sqrt{X} + \frac{2 \log C_2}{\frac{1}{2} \log X} \sum_{i=1}^{m} \frac{X}{2^i} \leq 2\sqrt{X} + (4 \log C_2) \left(\frac{X}{\log X}\right).$$

This gives our upper bound, because $\sqrt{X} \ll X/\log X$. Now remains to prove (6). We saw above that if $n \in \mathbb{Z}$ then

$$\prod_{n < p \leq 2n} p \leq \binom{2n}{n} \leq 4^n = \sum_{i=0}^{2n} \binom{2n}{i}.$$

Take $n = \lfloor X/2 \rfloor$. Then $2n \leq X$, and we get

$$\prod_{\frac{X}{2} < p \leq 2\left\lfloor \frac{X}{2} \right\rfloor} p \leq 2^{2n} \leq 2^X.$$

So

$$\prod_{\frac{X}{2} < p \leq X} p \leq X 2^X < C_2^X,$$

for X sufficiently large, for any $C_2 > 2$.

□

Lecture 28 is a problem class.

Lecture 28
Friday
07/12/18

Lecture 29
Tuesday
11/12/18

13.3 The Brun-Titchmarsh theorem and the Selberg sieve

What can we say about the number of primes p with $X < p \leq X + Y$? That is, $\pi(X+Y) - \pi(X)$. Think of Y being fixed for a moment. Best possible lower bound is zero.

Example.

$$n! + 2, \qquad \ldots, \qquad n! + n$$

is a sequence of consecutive composite numbers.

It was conjectured, in 1920s, by Hardy and Littlewood that

$$\pi(X+Y) \leq \pi(X) + \pi(Y),$$

that is $\pi(X+Y) - \pi(X) \leq \pi(Y)$. This is no longer believed.

Theorem 107.

$$\pi(X+Y) - \pi(X) \leq \frac{(2+o(1))\,Y}{\log Y},$$

where $o(1)$ is as $Y \to \infty$ and X is fixed.

In

$$X+1, \qquad \ldots, \qquad X+Y,$$

about half of these are divisible by two, about a third of these are divisible by three, about a sixth of these are divisible by six. If $p_1, \ldots, p_k$ are primes, the error term is 2^k, so can only consider the first $\log Y$ primes, which gives Theorem 107 for $Y/\log\log Y$. Selberg's idea is to weight the inclusion-exclusion count.

Proof. Let $\lambda_1, \lambda_2, \cdots \in \mathbb{R}$ be any sequence with $\lambda_1 = 1$. Let $R < Y$ be fixed for now. Later we will choose $R = Y^{1/2-\epsilon}$. Set

$$\nu(n) = \left(\sum_{d|n,\ d \leq R} \lambda_d \right)^2 \geq 0.$$

Suppose that p is prime, and $p > R$. Then by definition, $\nu(p) = \lambda_1^2 = 1$.

$$\pi(X+Y) - \pi(X) = \sum_{X < p \leq X+Y} 1 \leq \pi(R) + \sum_{X < n \leq X+Y} \nu(n) \leq R + \sum_{X < n \leq X+Y} \nu(n).$$

Now have to choose λ_i to minimise $\sum_{X \leq n \leq X+Y} \nu(n)$.

$$\begin{aligned}
\sum_{X < n \leq X+Y} \nu(n) &= \sum_{X < n \leq X+Y} \left(\sum_{d|n,\ d \leq R} \lambda_d \right)^2 \\
&= \sum_{X < n \leq X+Y} \left(\sum_{d_1|n,\ d_1 \leq R} \lambda_{d_1} \right) \left(\sum_{d_2|n,\ d_2 \leq R} \lambda_{d_2} \right) \\
&= \left(\sum_{d_1, d_2 \leq R} \lambda_{d_1} \lambda_{d_2} \right) \left(\sum_{X < n \leq X+Y,\ d_1|n,\ d_2|n} 1 \right) \\
&= \left(\sum_{d_1, d_2 \leq R} \lambda_{d_1} \lambda_{d_2} \right) \left(\frac{Y(d_1, d_2)}{d_1 d_2} + O(1) \right),
\end{aligned}$$

since $lcm(d_1, d_2) = d_1 d_2 / (d_1, d_2)$. Putting this together,

$$\pi(X+Y) - \pi(X) \leq Y \sum_{d_1, d_2 \leq R} \frac{\lambda_{d_1} \lambda_{d_2} (d_1, d_2)}{d_1 d_2} + R + O(1) \sum_{d_1, d_2 \leq R} |\lambda_{d_1} \lambda_{d_2}|,$$

where the leading term is

$$Y \sum_{d_1,d_2\leq R} \frac{\lambda_{d_1}\lambda_{d_2}\left(d_1,d_2\right)}{d_1d_2},$$

and the error term is

$$R + O\left(1\right) \sum_{d_1,d_2\leq R} \left|\lambda_{d_1}\lambda_{d_2}\right|.$$

Now choose λ_i such that $\lambda_1 = 1$, in such a way as to minimise the leading term. Then choose $R = Y^c$ for $c < 1/2$. Check that for any $\epsilon > 0$, we have $\lambda_d \ll_\epsilon d^\epsilon$. Then

$$\sum_{d_1d_2} \left|\lambda_{d_1}\lambda_{d_2}\right| \leq R^{2+2\epsilon} = Y^{2c(1+\epsilon)}.$$

Choose $\epsilon < 1/2c - 1$, then $Y^{2c(1+\epsilon)} \ll Y/\log Y$. Write $\overrightarrow{\lambda} = (\lambda_1, \lambda_2, \ldots)$.

$$Q\left(\overrightarrow{\lambda}\right) = \sum_{d_1,d_2\leq R} \frac{\lambda_{d_1}\lambda_{d_2}\left(d_1,d_2\right)}{d_1d_2}.$$

Want to minimise this subject to $\lambda_1 = 1$. Want to diagonalise $Q\left(\overrightarrow{\lambda}\right)$. Use, a slight variant of, Möbius inversion. For any m, $m = \sum_{d|m} \Phi\left(d\right)$. Take $m = (d_1, d_2)$. Then $(d_1, d_2) = \sum_{\delta|(d_1,d_2)} \Phi\left(\delta\right)$.

$$Q\left(\overrightarrow{\lambda}\right) = \sum_{d_1,d_2\leq R} \frac{\lambda_{d_1}\lambda_{d_2}\left(d_1,d_2\right)}{d_1d_2} = \sum_{\delta\leq R} \Phi\left(\delta\right) \left(\sum_{\delta|d,\ d\leq R} \frac{\lambda_d}{d}\right)^2,$$

by using that

$$\delta \mid d_1, \qquad \delta \mid d_2 \qquad \Longleftrightarrow \qquad \delta \,\Big|\, \frac{d_1d_2}{(d_1,d_2)}.$$

Set $u_\delta = \sum_{\delta|d,\ d\leq R} \lambda_d/d$. Then

$$Q\left(\overrightarrow{\lambda}\right) = \sum_{\delta\leq R} \Phi\left(\delta\right) u_\delta^2.$$

Lecture 30
Wednesday
12/12/18

Claim that

$$\frac{\lambda_d}{d} = \sum_{d|\delta,\ \delta\leq R} \mu\left(\frac{\delta}{d}\right) u_\delta. \tag{8}$$

Right hand side is

$$\sum_{d|\delta,\ \delta\leq R} \mu\left(\frac{\delta}{d}\right) \left(\sum_{\delta|d',\ d'\leq R} \frac{\lambda_{d'}}{d'}\right) = \sum_{d'\leq R} \frac{\lambda_{d'}}{d'} \left(\sum_{d|\delta|d'} \mu\left(\frac{\delta}{d}\right)\right).$$

So we need to show that

$$\sum_{d|\delta|d'} \mu\left(\frac{\delta}{d}\right) = \begin{cases} 1 & d = d' \\ 0 & \text{otherwise.} \end{cases}.$$

The sum is equal to

$$\sum_{m \,\Big|\, \frac{d'}{d}} \mu\left(m\right) = \left(1 * \mu\right)\left(\frac{d'}{d}\right) = \epsilon\left(\frac{d'}{d}\right).$$

The condition that $\lambda_1 = 1$ translates via (8) to the condition that $1 = \sum_{\delta\leq R} \mu\left(\delta\right) u_\delta$. The Cauchy-Schwarz inequality is $|ab| \leq |a|\,|b|$, that is

$$\sum_i a_ib_i \leq \left(\sum_i a_i^2\right)^{\frac{1}{2}} \left(\sum_i b_i^2\right)^{\frac{1}{2}},$$

with equality if and only if there exists λ such that $b_i = \lambda a_i$ for all i. So

$$1 = \sum_{\delta \leq R} \mu(\delta) u_\delta \leq \left(\sum_{\delta \leq R} \Phi(\delta) u_\delta^2 \right)^{\frac{1}{2}} \left(\sum_{\delta \leq R} \frac{\mu(\delta)^2}{\Phi(\delta)} \right)^{\frac{1}{2}}.$$

So

$$Q\left(\overrightarrow{\lambda}\right) = \sum_{\delta \leq R} \Phi(\delta) u_\delta^2 \geq \frac{1}{D},$$

where $D = \sum_{\delta \leq R} \mu(\delta)^2 / \Phi(\delta)$. Equality holds when $u_\delta = \mu(\delta) / D\Phi(\delta)$. We are going to show that $D \geq \log R + O(1)$. Since $R = Y^c$, this gives us a leading term of

$$\frac{Y}{\log R} = \frac{Y}{\log Y^c} = \frac{1}{c}\left(\frac{Y}{\log Y}\right).$$

$c < 1/2$ gives $1/c > 2$.

$$D = \sum_{\delta \leq R} \frac{\mu(\delta)^2}{\Phi(\delta)} = \sum_{\delta \leq R,\ \delta \text{ squarefree}} \frac{1}{\Phi(\delta)}.$$

If δ is squarefree, write $\delta = p_1 \dots p_k$. Then

$$\Phi(\delta) = (p_1 - 1) \dots (p_k - 1) = p_1 \dots p_k \left(1 - \frac{1}{p_1}\right) \dots \left(1 - \frac{1}{p_k}\right).$$

So

$$D = \sum_{\delta \leq R,\ \delta \text{ squarefree}} \frac{1}{\delta} \prod_{p \mid \delta} \left(1 - \frac{1}{p}\right)^{-1}.$$

Now, $(1 - 1/p)^{-1} = 1 + 1/p + \dots$. So

$$D = \sum_{\delta \leq R,\ \delta \text{ squarefree}} \frac{1}{\delta} \prod_{p \mid \delta} \left(1 + \frac{1}{p} + \dots\right) \geq \sum_{n \leq R} \frac{1}{n} = \log R + O(1),$$

by taking $n \leq R$, and writing $n = p_1^{a_1} \dots p_m^{a_m}$ and $\delta = p_1 \dots p_m \leq R$ squarefree, so

$$\frac{1}{n} = \frac{1}{\delta}\left(\frac{1}{p_1^{a_1 - 1}} \cdots \frac{1}{p_m^{a_m - 1}}\right).$$

The only thing remaining is to show that $\lambda_d \ll_\epsilon d^\epsilon$. Recall that $u_\delta = \mu(\delta) / D\Phi(\delta)$. So

$$\lambda_d = d \sum_{d \mid \delta,\ \delta \leq R} \mu\left(\frac{\delta}{d}\right) u_\delta = \frac{d}{D} \sum_{d \mid \delta,\ \delta \leq R} \frac{\mu\left(\frac{\delta}{d}\right)\mu(\delta)}{\Phi(\delta)} = \frac{d}{D} \sum_{d \mid \delta,\ \delta \leq R,\ \delta \text{ squarefree}} \frac{\mu\left(\frac{\delta}{d}\right)\mu(\delta)}{\Phi(\delta)}.$$

Write $\delta' = \delta / d$. Since $\delta = \delta' d$, and δ is squarefree, we have $(\delta', d) = 1$, so $\Phi(\delta) = \Phi(\delta')\Phi(d)$. So

$$|\lambda_d| \leq \frac{d}{\Phi(d) D} \sum_{\delta' \leq R,\ \delta' \text{ squarefree}} \frac{1}{\Phi(\delta')} = \frac{d}{\Phi(d)}.$$

Need to show that $\Phi(d) \gg_\epsilon d^{1-\epsilon}$ if d is squarefree,

$$\Phi(d) = \prod_{p \mid d} (p - 1).$$

If p is sufficiently large, then $p - 1 \geq p^{1-\epsilon}$. If p is not sufficiently large, then $(p-1)/p > 0$ can be regarded as a constant. $\square$