

M3P14 Number Theory

Lectured by Prof Toby Gee
 Typed by David Kurniadi Angdinata

Autumn 2018

$$\begin{aligned}
 \pi &= 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1 + \frac{1}{292 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{\ddots}}}}}}}} \\
 &= 0 + \frac{4}{1 + \frac{1^2}{2 + \frac{3^2}{2 + \frac{5^2}{\ddots}}}} = 0 + \frac{4}{1 + \frac{1^2}{3 + \frac{2^2}{5 + \frac{3^2}{\ddots}}}} = 3 + \frac{1^2}{6 + \frac{3^2}{6 + \frac{5^2}{6 + \frac{7^2}{\ddots}}}} \\
 &= 2 + \frac{2}{\frac{1}{1} + \frac{1}{\frac{1}{2} + \frac{1}{\frac{1}{3} + \frac{1}{\ddots}}}} = 2 + \frac{2}{1 + \frac{1 \cdot 2}{1 + \frac{2 \cdot 3}{1 + \frac{3 \cdot 4}{\ddots}}}} = 2 + \frac{4}{3 + \frac{1 \cdot 3}{4 + \frac{3 \cdot 5}{4 + \frac{5 \cdot 7}{\ddots}}}} \\
 &= 3 + \frac{1^3}{6 + \frac{1^3 + 2^3}{6 + \frac{1^3 + 2^3 + 3^3 + 4^3}{6 + \frac{1^3 + 2^3 + 3^3 + 4^3 + 5^3 + 6^3}{6 + \frac{1^3 + 2^3 + 3^3 + 4^3 + 5^3 + 6^3 + 7^3 + 8^3}{\ddots}}}}
 \end{aligned}$$

Syllabus

Prime numbers and factorisation. Euclid's algorithm and consequences. Congruences. The structure of $(\mathbb{Z}/n\mathbb{Z})^\times$. Primality testing and factorisation. RSA algorithm. Quadratic reciprocity. Sums of squares. Pell's equation. Continued fractions. Diophantine approximation. Primes in arithmetic progressions. Arithmetic functions. The distribution of prime numbers.

Contents

0	Introduction	4
1	Euclid's algorithm and unique factorisation	5
1.1	Divisibility	5
1.2	Euclid's algorithm	5
1.3	Unique factorisation	6
1.4	Linear diophantine equations	6
2	Congruences and modular arithmetic	7
2.1	Congruences	7
2.2	Linear congruence equations	8
2.3	The Chinese remainder theorem	8
3	The structure of $(\mathbb{Z}/n\mathbb{Z})^\times$	9
3.1	The Euler Φ function	9
3.2	Euler's theorem	9
4	Primality testing and factorisation	13
4.1	Factorisation	13
4.2	Testing primality	14
5	Public-key cryptography	16
5.1	Messages as sequences of classes modulo n	16
5.2	The Rivest-Shamir-Adleman (RSA) algorithm	16
5.3	Signing with RSA	16
5.4	Discrete logarithms	16
6	Quadratic reciprocity	17
6.1	Quadratic residues	17
6.2	Computing Legendre symbols	18
6.3	Proof of quadratic reciprocity	20
6.4	Jacobi symbols	21
7	Sums of squares	22
7.1	Sums of two squares	22
7.2	Sums of four squares - the ring of quaternions	23
7.3	Proof of Lagrange's theorem	24
7.4	Sums of three squares	25
8	Pell's equation	26
8.1	Pell's equation	26
8.2	Quadratic subrings of \mathbb{C}	26
8.3	Factorisation in quadratic rings	27
8.4	Back to Pell's equation	27
8.5	Constructing the fundamental 1-unit	28
8.6	The equation $x^2 - dy^2 = -1$	29
9	Continued fractions	30
9.1	Rational continued fractions	30
9.2	Infinite continued fractions	30
9.3	Best approximations	32
9.4	Returning to Pell's equation	33
9.5	Periodic continued fractions	35

10 Diophantine approximation	36
10.1 Liouville's theorem	36
10.2 Constructing transcendentals	36
10.3 Roth's theorem	36
11 Primes in arithmetic progressions	37
11.1 Primes in arithmetic progressions	37
11.2 Elementary results	37
11.3 Cyclotomic polynomials	38
11.4 Primes congruent to 1 mod n	39
12 Arithmetic functions	40
12.1 Dirichlet convolution	40
12.2 Möbius inversion	40
13 The distribution of prime numbers	41
13.1 Reminder of asymptotic notation	41
13.2 The prime number theorem	41
13.3 The Brun-Titchmarsh theorem and the Selberg sieve	43

0 Introduction

Lecture 1
Friday
05/10/18

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. The following are typical questions here.

- How do you tell if a number is prime?
- How many primes are there congruent to $a \pmod b$ for given a and 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.

- 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 , and $a \mid b$, 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 and b is the largest positive integer dividing a and b . Write it as (a, b) .

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

Note. The ring \mathbb{Z} is a principal ideal domain (PID). If $f_1, \dots, f_n \in R$, write (f_1, \dots, 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. 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. 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. \square

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. \square

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.

$$\begin{array}{lll}
 (120, 87) = (87, 33) & 120 = 87 + 33 & 3 = 12 - 9 \\
 = (33, 21) & 87 = 2(33) + 21 & = 12 - (21 - 12) \\
 = (21, 12) & 33 = 21 + 12 & = 2(12) - 21 \\
 = (12, 9) & 21 = 12 + 9 & = 2(33 - 21) - 21 \\
 = (9, 3) & 12 = 9 + 3 & = 2(33) - 3(21) \\
 = (3, 0) & 9 = 3(3) + 0 & = 2(33) - 3(87 - 2(33)) \\
 & & = 8(33) - 3(87) \\
 & & = 8(120 - 87) - 3(87) \\
 & & = 8(120) - 11(87).
 \end{array}$$

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. Idea is to write (a_n, b_n) for the sequence of pairs in Euclid's algorithm, and use downwards induction on n .¹ \square

¹Exercise

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 $b = n(rb) + (ab)s$, which is divisible by n . □

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

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 $n = r(na) + s(nb)$, which is divisible by ab . □

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

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

Proof. Induction on n , where $n = 2$ is Proposition 9. ² □

Proposition 11. Every $n \in \mathbb{Z}^*$ can be written as $\pm p_1 \dots p_r$ 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 $|n| = ab$ for $1 < a, b < |n|$, and by induction each of a and b is a product of primes. □

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

Proof. Existence is Proposition 11. Suppose that $n = p_1 \dots p_r = q_1 \dots q_s$, with p_i and q_i prime. Then without loss of generality suppose $r, s \geq 1$. Then $p_1 \mid p_1 \dots p_r$, so $p_1 \mid q_1 \dots q_s$. By Corollary 8, either $p_1 \mid q_1$ or $p_1 \mid q_2 \dots q_s$. 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 $p_2 \dots p_r = q_1 \dots q_{i-1} q_{i+1} \dots q_s$. 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_j , except q_i . □

1.4 Linear diophantine equations

Let $a, b, c \in \mathbb{Z}^*$. Want to solve

$$ax + by = c, \quad x, y \in \mathbb{Z}.$$

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

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

By Theorem 6, since $(a/(a, b), b/(a, b)) = 1$, we can find $r, s \in \mathbb{Z}$ with $ar/(a, b) + bs/(a, b) = 1$, 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 $x = rc/(a, b)$ and $y = sc/(a, b)$ is a solution. Then X and Y is another solution if and only if

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

See that the solutions are exactly

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

²Exercise

2 Congruences and modular arithmetic

2.1 Congruences

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

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

Example. If $a \equiv b \pmod n$ and $c \equiv d \pmod n$, then

$$a + c \equiv b + d \pmod n, \quad ac \equiv bd \pmod n.$$

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

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

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

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

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 $rs \equiv 1 \pmod n$. 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 $rx \equiv 1 \pmod n$, that is r is a unit. So

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{0 \leq i < n \mid (i, n) = 1\}.$$

Example. If p is a prime, then

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

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 modulo p is a unit.

2.2 Linear congruence equations

Consider the question of solving

$$ax \equiv b \pmod{c}, \quad a, b, c, x \in \mathbb{Z}.$$

This is equivalent to solving

$$ax + cy = b, \quad y \in \mathbb{Z}.$$

We saw yesterday that this has solutions if and only if $(a, c) \mid b$. Furthermore, there is a unique solution modulo $c/(a, c)$, 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 modulo c . If x_0 is one solution, the others are

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

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

Example.

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

2.3 The 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*

$$\mathbb{Z}/m_1 \dots m_n \mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/m_1 \mathbb{Z} \times \dots \times \mathbb{Z}/m_n \mathbb{Z}$$

is an isomorphism of rings. Consequently,

$$(\mathbb{Z}/m_1 \dots m_n \mathbb{Z})^\times \xrightarrow{\sim} (\mathbb{Z}/m_1 \mathbb{Z})^\times \times \dots \times (\mathbb{Z}/m_n \mathbb{Z})^\times$$

is an isomorphism of abelian groups.

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

Proof. The map

$$\mathbb{Z}/m_1 \dots m_n \mathbb{Z} \rightarrow \mathbb{Z}/m_1 \mathbb{Z} \times \dots \times \mathbb{Z}/m_n \mathbb{Z}$$

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 and 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 $x \pmod{m_1 \dots m_n}$ with $x \equiv a_i \pmod{m_i}$. Write

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

Choose q_i such that $q_i M_i \equiv 1 \pmod{m_i}$, 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

$$x \equiv a_i q_i M_i \equiv a_i \pmod{m_i}.$$

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

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

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

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

$$(\mathbb{Z}/mn\mathbb{Z})^\times \xrightarrow{\sim} (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times.$$

Φ is not strongly multiplicative, since $\Phi(4) = 2 \neq 1 = \Phi(2)\Phi(2)$. Write $n = \prod_i p_i^{a_i}$, where p_i are distinct primes. Then $\Phi(n) = \prod_i \Phi(p_i^{a_i})$. If p is prime then

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

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

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

3.2 Euler's theorem

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

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

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

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

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

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

Next, want 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 $n = 5$. Then $(\mathbb{Z}/5\mathbb{Z})^\times = \{1, 2, 3, 4\}$. This has order four. So it is either cyclic of order four or a product of two cyclic groups of order two. Since $2^2 = 4$, $2^3 = 3$, and $2^4 = 1$, $(\mathbb{Z}/5\mathbb{Z})^\times$ is cyclic of order four.

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

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 $\text{ord}_n g$ for the order of g in $(\mathbb{Z}/n\mathbb{Z})^\times$, the **order of g modulo n** .

Proposition 19. *If G is a group and g is an element of order a , then $g^n = 1$ if and only if $a \mid n$.*

Proof. If $n = ab$ then $g^n = (g^a)^b = 1^b = 1$. Conversely, write $n = ab + r$ with $0 \leq r < a$. Then $g^r = 1$ and since $r < a$ we have $r = 0$. \square

Lecture 4
Friday
12/10/18

In particular, if $(g, n) = 1$, then $g^{\Phi(n)} = 1$, by Euler's theorem, so Proposition 19 implies that $\text{ord}_n g \mid \Phi(n)$. 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 $\text{ord}_p g = \Phi(p) = p - 1$. We will do this by counting the number of elements of each order. The 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 modulo p for any prime p , namely ± 1 , but it can have more modulo n if n is composite. For example, if $n = 15$, then 4 and 11 are also solutions, since $X^2 - 1 \equiv 0 \pmod{n}$ if and only if $n \mid (X + 1)(X - 1)$, for example $15 \mid (4 + 1)(4 - 1)$.

Definition 20. $g \in \mathbb{Z}$ with $(g, p) = 1$ is a **primitive root** if $\text{ord}_p g = p - 1$, 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 $\deg P$, where $\deg P = 0$ is obvious. Let $\deg P = d$, and assume the result holds for degree at most $d - 1$. Let $P(X) = cX^d + \dots$ and $S(X) = P(X) - cX^{d-1}(X - \alpha)$. 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 $(X - \alpha)Q(X) = cX^{d-1}(X - \alpha) + S(X) = P(X)$. \square

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 $d = \deg P$, where $d = 1$ is obvious. If P has no roots, then we are done. Otherwise, let α be a root. By Lemma 21, $P(X) = (X - \alpha)Q(X)$, and $Q(X)$ has degree $d - 1$, so we are done by induction. \square

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 $\mathbb{Z}/p\mathbb{Z}$. By Fermat's little theorem, $X^{p-1} - 1$ has exactly $p - 1$ roots. Since $d \mid p - 1$, we can write

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

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

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

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

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

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

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

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

Now use the inductive hypothesis and Corollary 23. □

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

Proof. Consider three cases.

$n=1$. Theorem 25.

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

$$g^{p-1} \not\equiv 1 \pmod{p^2}.$$

Then $\#(\mathbb{Z}/p^2\mathbb{Z})^\times = \Phi(p^2) = p(p-1)$. So $\text{ord}_{p^2} g \mid p(p-1)$. On the other hand, $g^{\text{ord}_{p^2} g} \equiv 1 \pmod{p^2}$, so $g^{\text{ord}_{p^2} g} \equiv 1 \pmod{p}$, so $p-1 \mid \text{ord}_{p^2} g$, because $\text{ord}_p g = p-1$ by assumption. But $\text{ord}_{p^2} g \neq p-1$, as $g^{p-1} \not\equiv 1 \pmod{p^2}$. So $\text{ord}_{p^2} g = p(p-1)$, as required. Suppose now that

$$g^{p-1} \equiv 1 \pmod{p^2}.$$

It suffices to show that $(g+p)^{p-1} \not\equiv 1 \pmod{p^2}$, as we can then apply the analysis above with $g+p$ in place of g . By the binomial theorem,

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

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

$n \geq 2$. It suffices to show that if $\text{ord}_{p^2} g = p(p-1)$, then $\text{ord}_{p^n} g = p^{n-1}(p-1)$. We do this by induction on n . So assume that $\text{ord}_{p^n} g = p^{n-1}(p-1)$. Then $\text{ord}_{p^n} g \mid \text{ord}_{p^{n+1}} g$, and $\text{ord}_{p^{n+1}} g \mid \Phi(p^{n+1}) = p^n(p-1)$. So either $\text{ord}_{p^{n+1}} g = p^n(p-1)$, or $\text{ord}_{p^{n+1}} g = p^{n-1}(p-1)$. So we need to show that

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

To do this, consider $g^{p^{n-2}(p-1)}$ modulo p^{n-1} and modulo p^n . Since $\Phi(p^{n-1}) = p^{n-2}(p-1)$, by Euler's theorem, $g^{p^{n-2}(p-1)} \equiv 1 \pmod{p^{n-1}}$. Write $g^{p^{n-2}(p-1)} = 1 + p^{n-1}t$. Since $\text{ord}_{p^n} g = p^{n-1}(p-1)$ by assumption, $g^{p^{n-2}(p-1)} \not\equiv 1 \pmod{p^n}$, that is $p \nmid t$. Then

$$\begin{aligned} g^{p^{n-1}(p-1)} &= \left(g^{p^{n-2}(p-1)}\right)^p = (1 + p^{n-1}t)^p = 1 + p^n t + \binom{p}{2} p^{2(n-1)} t^2 + \dots + p^{p(n-1)} t^p \\ &\equiv 1 + p^n t \pmod{p^{n+1}}, \end{aligned}$$

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

$$p^{n+1} \mid p^{2n-1} = p^{2(n-1)+1} \mid \binom{p}{2} p^{2(n-1)}.$$

So $g^{p^{n-1}(p-1)} \not\equiv 1 \pmod{p^{n+1}}$, 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, since $1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod{8}$, so every element has order two.

Lemma 27. For $n \geq 0$ we have

$$5^{2^n} \equiv 1 + 2^{n+2} \pmod{2^{n+3}}.$$

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

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

where $t + 2^{n+1}t^2$ is odd. □

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

$$(\mathbb{Z}/2^n\mathbb{Z})^\times \xrightarrow{\sim} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z}.$$

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 $\{1, \dots, g^{\text{ord } g - 1}\}$ generated by g . Consider the natural map

$$\langle -1 \rangle \times \langle 5 \rangle \rightarrow (\mathbb{Z}/2^n\mathbb{Z})^\times.$$

This is injective, because if $\pm 1 (5)^s \equiv 1 \pmod{2^n}$ then in particular $\pm 1 (5)^s \equiv 1 \pmod{4}$ so $\pm 1 \equiv 1 \pmod{4}$, so we must have $5^s \equiv 1 \pmod{2^n}$, that is $5^s = 1$ in $\langle 5 \rangle$. Then $\langle -1 \rangle$ has order 2 and $\langle 5 \rangle$ has order $\text{ord}_{2^n} 5 = 2^{n-2}$ by Lemma 27. So $\langle -1 \rangle \times \langle 5 \rangle$ has order $2(2^{n-2}) = 2^{n-1} = \Phi(2^n) = \#(\mathbb{Z}/2^n\mathbb{Z})^\times$. So the map $\langle -1 \rangle \times \langle 5 \rangle \rightarrow (\mathbb{Z}/2^n\mathbb{Z})^\times$ 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 $r \geq 1$.

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 $1, \dots, g^{p-2}$ 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 $\text{ord}_p g \mid p-1$ and $\text{ord}_p g \nmid (p-1)/q$. If this holds for all $q \mid p-1$, then $\text{ord}_p g = p-1$, 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 $p-1 = 30 = (2)(3)(5)$. Then g is a primitive root if and only if

$$g^{15} \neq 1, \quad g^{10} \neq 1, \quad g^6 \neq 1.$$

- 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 modulo 31.

Lecture 6
Wednesday
17/10/18

4 Primality testing and factorisation

The 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 a paper or a calculator? If $n \leq 400$ and n is composite, then it 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, 11. Use that $10 \equiv 1 \pmod{3}$ and $10 \equiv -1 \pmod{11}$. So

$$\sum_i a_i 10^i \equiv \sum_i a_i \pmod{3}, \quad \sum_i a_i 10^i \equiv \sum_i a_i (-1)^i \pmod{11}.$$

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

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

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

Example.

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

Factor four digit numbers by an algorithm due to Fermat. The idea is to first check for small prime factors by hand, say $p = 2, \dots, 19$. 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 and b odd and $a \leq b$, then

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

If you know $(a+b)/2$ and $(b-a)/2$, you can recover a and 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 $n = 6077$. Then $77^2 < 6077 < 78^2$, so

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

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

Lecture 7
Friday
19/10/18

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.

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

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

Since $41^2 \equiv 2^5 \pmod{1649}$ and $43^2 \equiv (2)^3 (5)^2 \pmod{1649}$,

$$80^2 \equiv (41)^2 (43)^2 = 1763^2 \equiv 114^2 \pmod{1649}.$$

Then

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

In fact, $1649 = (17)(97)$. Better for this last step would be to have computed

$$(194, 1649) = 97, \quad (34, 1649) = 17.$$

Can do this quickly using Euclid's algorithm. To make this into an efficient algorithm, need to have a way given x_1, \dots, x_r to find a subset whose product is a square. If we know the prime factorisation for the x_i , we can write

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

Want to choose $\epsilon_i = 0, 1$ 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

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

Let

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

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

$$(\epsilon_1 \ \epsilon_2) \begin{pmatrix} 5 & 0 \\ 3 & 2 \end{pmatrix} \equiv (0 \ 0) \pmod{2} \iff (\epsilon_1 \ \epsilon_2) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = (0 \ 0)$$

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. The 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 $m^2 \equiv n \pmod{p}$. This gives two possible values for $m \pmod{p}$. 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. $m^2 = 1649 \equiv 2 \pmod{3}$ has no solutions.

4.2 Testing primality

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

Lemma 30. Let $n > 1$ be congruent to 3 mod 4. Then n is prime if and only if

$$a^{\frac{n-1}{2}} \equiv \pm 1 \pmod{n}, \quad (a, n) = 1.$$

Lecture 8
Tuesday
23/10/18

Proof.

- If n is prime, then $a^{n-1} \equiv 1 \pmod{n}$ by Fermat's little theorem, so $(a^{(n-1)/2})^2 \equiv 1 \pmod{n}$, so $a^{(n-1)/2} \equiv \pm 1 \pmod{n}$.
- Suppose firstly that $n = p^k$ with p prime, and $k \geq 2$. Try

$$a = 1 + p.$$

Then

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

by the binomial theorem. If $a^{(n-1)/2} \equiv \pm 1 \pmod{n}$, then

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

so

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

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,

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

Choose a such that

$$a \equiv -1 \pmod{r}, \quad a \equiv 1 \pmod{s}.$$

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

$$a^{\frac{n-1}{2}} \equiv -1 \pmod{r}, \quad a^{\frac{n-1}{2}} \equiv 1 \pmod{s}.$$

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

□

Lemma 31. Suppose that $n \equiv 3 \pmod{4}$ is composite. Then the set of $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ which satisfy $a^{(n-1)/2} \equiv \pm 1 \pmod{n}$ is a proper subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$.

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

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

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

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 $\frac{3}{4}$ of the numbers $1 \leq a \leq n-1$ satisfy $a^{(n-1)/2} \not\equiv \pm 1 \pmod{n}$. 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 $a^{(n-1)/2} \not\equiv \pm 1 \pmod{n}$ is at least $1 - \left(\frac{1}{4}\right)^x$. This gives a probabilistic algorithm to check if n is prime in polynomial time. If you assume the generalised Riemann hypothesis (GRH) you can find some

$$1 \leq a \leq \left\lceil 2(\log n)^2 \right\rceil, \quad a^{\frac{n-1}{2}} \not\equiv \pm 1 \pmod{n}.$$

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 modulo n

How do we turn messages into numbers in $\mathbb{Z}/n\mathbb{Z}$? The 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 modulo n .

5.2 The Rivest-Shamir-Adleman (RSA) algorithm

Now apply some function $f : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$, and then tell whoever you are trying to communicate with the result of this computation. Then they should apply some other function $g : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$, 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. The idea is to choose two large prime numbers p and q and set $n = pq$. Choose $(e, \Phi(n)) = 1$. Find d such that

$$de \equiv 1 \pmod{\Phi(n)} = (p-1)(q-1) = n - (p+q) + 1.$$

Publish n and e , and you keep $p, q, \Phi(n), d$ secret. Let $f(x) = x^e \pmod{n}$ and $g(x) = x^d \pmod{n}$. Then

$$(x^e)^d \equiv x^{de} \equiv x \pmod{n},$$

because $de \equiv 1 \pmod{\Phi(n)}$ and $x^{\Phi(n)} \equiv 1 \pmod{n}$. 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

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

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 is a problems class.

Lecture 9
Wednesday
24/10/18

5.3 Signing with RSA

If you have functions $f, g : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ with $f \circ g = g \circ f = \text{id}$, 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

$$x^r \equiv a \pmod{n},$$

write $x = g^y$, and the congruence becomes equivalent to

$$yr \equiv \log_g(a) \pmod{\Phi(n)}.$$

Unfortunately, or fortunately for cryptography, computing \log_g is believed to be a hard problem. In particular, there is 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 $g^x \pmod{p}$. If they later input y , you compute g^y , and check it matches what you stored. If it does then $y \equiv x \pmod{p-1}$.

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

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

Proof. The map

$$\begin{array}{ccc} (\mathbb{Z}/p\mathbb{Z})^\times & \longrightarrow & (\mathbb{Z}/p\mathbb{Z})^\times \\ x & \longmapsto & x^2 \end{array}$$

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

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

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

Proof. Let H be the image of

$$\begin{array}{ccc} (\mathbb{Z}/p\mathbb{Z})^\times & \longrightarrow & (\mathbb{Z}/p\mathbb{Z})^\times \\ x & \longmapsto & x^2 \end{array},$$

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 $(\mathbb{Z}/p\mathbb{Z})^\times = H \cup 1 + H$. \square

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

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

Proposition 35 can be restated as saying that

$$\begin{array}{ccc} (\mathbb{Z}/p\mathbb{Z})^\times & \longrightarrow & \{\pm 1\} \\ a & \longmapsto & \left(\frac{a}{p}\right) \end{array}$$

is a group homomorphism, that is

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

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

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

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

Proof. Let g be a primitive root modulo p , and write $a \equiv g^r \pmod{p}$ for $0 \leq r < p-1$. Now $(g^{(p-1)/2})^2 = g^{p-1} \equiv 1 \pmod{p}$. So $g^{(p-1)/2} \equiv \pm 1 \pmod{p}$. Since g is a primitive root, $g^{\frac{p-1}{2}} \not\equiv 1 \pmod{p}$, so $g^{(p-1)/2} \equiv -1 \pmod{p}$. So

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

But

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

□

6.2 Computing Legendre symbols

Proposition 38. -1 is a square modulo p if and only if $p = 2$ or $p \equiv 1 \pmod{4}$.

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

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

so in fact

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

Then

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

□

Proposition 39 (Gauss' lemma).

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

that is

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

Proof.

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

by Euler's criterion. Let $q = (p-1)/2$, and let

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

Subtracting p from every term which is bigger than q ,

$$Q \equiv (2)(4)\dots(-3)(-1) \equiv (-1)^r q! \pmod{p},$$

where r is the number of odd integers in $1, \dots, q$. Since $p \nmid q!$, we have $2^{(p-1)/2} \equiv (-1)^r \pmod{p}$. Now the following holds.³

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

□

³Exercise

Example. If $p \equiv 1 \pmod{8}$, say $p = 1 + 8n$, then $q = 4n$. Odd integers in $1, \dots, 4n$ are $1, 3, \dots, 4n-3, 4n-1$, so $r = 2n$.

Example.

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

Theorem 40 (Law of quadratic reciprocity). *If p and 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 $p \equiv q \equiv 3 \pmod{4}$, 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 modulo 5 are 1 and 4. So

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

- What is $\left(\frac{3}{p}\right)$ for $p \neq 3$? If $p \equiv 1 \pmod{4}$, then

$$\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right) = \begin{cases} 1 & p \equiv 1 \pmod{3} \\ -1 & p \equiv -1 \pmod{3} \end{cases}.$$

If $p \equiv -1 \pmod{4}$, then

$$\left(\frac{3}{p}\right) = -\left(\frac{p}{3}\right) = \begin{cases} 1 & p \equiv -1 \pmod{3} \\ -1 & p \equiv 1 \pmod{3} \end{cases}.$$

So

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

For example, $\left(\frac{3}{7}\right) = -1$, since QRs are $1, 2, 4$, and $\left(\frac{3}{11}\right) = 1$, since $5^2 \equiv 3 \pmod{11}$.

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

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

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

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

6.3 Proof of quadratic reciprocity

Lecture 12
Wednesday
31/10/18

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

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

Proof of Theorem 40. 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 $x \in (\mathbb{Z}/pq\mathbb{Z})^\times$. 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

$$\{(x, y) \mid 1 \leq x \leq \frac{p-1}{2}, 1 \leq y \leq q-1\}.$$

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

$$\{(x, y) \mid 1 \leq x \leq p-1, 1 \leq y \leq \frac{q-1}{2}\}.$$

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

$$\{1 \leq i \leq \frac{pq-1}{2} \mid (i, pq) = 1\}.$$

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

$$\prod_{i=1, (i,pq)=1}^{\frac{pq-1}{2}} i = \left(\prod_{i=1, (i,p)=1}^{\frac{pq-1}{2}} i \right) / \left(\prod_{i=1, (i,p)=1, q|i}^{\frac{pq-1}{2}} i \right), \quad (1)$$

$$\prod_{i=1, (i,p)=1}^{\frac{pq-1}{2}} i = \left(\prod_{i=1, (i,p)=1}^{p\left(\frac{q-1}{2}\right)} i \right) \left(\prod_{i=p\left(\frac{q-1}{2}\right)+1, (i,p)=1}^{p\left(\frac{q-1}{2}\right)+\frac{p-1}{2}} i \right), \quad (2)$$

$$\prod_{i=1, (i,p)=1, q|i}^{\frac{pq-1}{2}} i = \prod_{j=1, (j,p)=1}^{\frac{p-1}{2}} qj = q^{\frac{p-1}{2}} \left(\frac{p-1}{2} \right)!. \quad (3)$$

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

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

□

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

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

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

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 $a \pmod{b}$.
3. $\left(\frac{a^2}{b}\right) = 1$.
4. $\left(\frac{-1}{b}\right) = (-1)^{(b-1)/2}$.
5. $\left(\frac{2}{b}\right) = (-1)^{(b^2-1)/8}$.
6. If $a, b > 0$ are both odd

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

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, and 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 and 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

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

that is need

$$(b_1^2 - 1) + (b_2^2 - 1) \equiv (b_1 b_2)^2 - 1 \pmod{16},$$

that is $(b_1^2 - 1)(b_2^2 - 1) \equiv 0 \pmod{16}$. This is true because $b_1^2 \equiv b_2^2 \equiv 1 \pmod{4}$. □

Example. Since

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

7411 is not a square modulo 9283.

Lecture 13
Friday
02/11/18

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

$$n = x^2 + y^2, \quad x, y \in \mathbb{Z}.$$

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

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

Definition 45. The **Gaussian integers** $\mathbb{Z}[i]$ are the subring of \mathbb{C} consisting of

$$a + bi, \quad a, b \in \mathbb{Z}.$$

The **norm** is defined by

$$\begin{aligned} N : \mathbb{Z}[i] &\longrightarrow \mathbb{Z}_{\geq 0} \\ a + bi &\longmapsto a^2 + b^2 \end{aligned}$$

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

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

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

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

□

Theorem 47 (Fermat's two square theorem). If $p \equiv 1 \pmod{4}$ 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 $ab = 0$ implies that $a = 0$ or $b = 0$, and there exists a function $N : R \rightarrow \mathbb{Z}_{\geq 0}$ 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. Then $\mathbb{Z}[i]$ 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. The units in $\mathbb{Z}[i]$ are ± 1 and $\pm i$.⁴ Two elements of $\mathbb{Z}[i]$ are **associates** if their ratio is a unit, that is z and w are associates if $z = uw$ for $u = \pm 1, \pm i$.

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 $p \mid q$ implies that $q = pv$ 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 $N(p) \mid q^2$, so $N(p) = q$ 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 $N(p) = q$. □

⁴Exercise

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 $p \equiv 1 \pmod{4}$, 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 $n^2 \equiv -1 \pmod{p}$. If $p = 2$ obvious, and if $p \equiv 1 \pmod{4}$,

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

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 $n \pm i = p(c \pm di)$, so $1 = pd$, a contradiction. \square

Remark. If $p \equiv 3 \pmod{4}$ then p is not a sum of two squares, even modulo four.

Remark. In practice, to go from $n^2 + 1 \equiv 0 \pmod{p}$ to finding a and 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*

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

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. \square

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

$$a + bi + cj + dk, \quad a, b, c, d \in \mathbb{R},$$

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

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

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

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

In particular, if m and 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 $p \equiv 3 \pmod{4}$ is prime, then there exist x and y such that*

$$x^2 + y^2 + 1 \equiv 0 \pmod{p}.$$

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

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

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 $p \equiv 3 \pmod{4}$, $\left(\frac{-1}{p}\right) = -1$ 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 and y such that

$$x^2 \equiv a \pmod{p}, \quad y^2 \equiv -(a+1) \pmod{p}.$$

Then $x^2 + y^2 \equiv -1 \pmod{p}$. □

By Lemma 54, there exist $x, y \in \mathbb{Z}$ such that

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

for some r . Since the congruence $x^2 + y^2 + 1 \equiv 0 \pmod{p}$ only depends on x and y modulo p , we can find x and y with $-p/2 < x, y < p/2$. Then

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

Proposition 55. *Suppose that*

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

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

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

Proposition 55 implies that 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 $x \equiv y \pmod{2}$ and $z \equiv w \pmod{2}$. Then take

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

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

$$x \equiv a \pmod{r}, \quad y \equiv b \pmod{r}, \quad z \equiv c \pmod{r}, \quad w \equiv d \pmod{r}.$$

Then

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

Write $a^2 + b^2 + c^2 + d^2 = rr'$. Then $rr' < 4(r/2)^2 = r^2$, 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 $1 \leq r' < r$. Then

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

Then

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

So take

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

□

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

$$\left\{ \frac{a + bi + cj + dk}{2} \mid a \equiv b \equiv c \equiv d \pmod{2} \right\}.$$

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

We will find it useful to write

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

This suggests that we should look at a ring like

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

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 $a_0 + \cdots + a_n\pi^n$ for $a_i \in \mathbb{Z}$ and n arbitrary.
- Also $\mathbb{Z}[\sqrt[3]{2}]$ is not just the set $\{a + b\sqrt[3]{2} \mid a, b \in \mathbb{Z}\}$, because this set does not contain $(\sqrt[3]{2})^2 = \sqrt[3]{4}$.
- Also $\mathbb{Z}[\frac{1}{p}]$ contains $1/p^n$ for all n , so in fact $\mathbb{Z}[\frac{1}{p}] = \{a/p^n \mid a \in \mathbb{Z}, n \geq 0\}$.

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

Lecture 16 is a problems class.

Lecture 16
Friday
09/11/18

8.2 Quadratic subrings of \mathbb{C}

Definition 59. Say that $\alpha \in \mathbb{C}$ is an **algebraic integer of degree two** if it is a root of a polynomial

$$X^2 + aX + b, \quad a, b \in \mathbb{Z}, \quad \alpha \notin \mathbb{Z}.$$

Example.

- $\alpha = i$ is a root of $X^2 + 1$.
- $\alpha = \sqrt{d}$ is a 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 it is closed under multiplication. But

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

□

Lecture 17
Tuesday
13/11/18

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

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 and α^* are the roots of $X^2 + aX + b$, we have $\alpha + \alpha^* = -a$ and $\alpha\alpha^* = b$. 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 $N : \mathbb{Z}[\alpha] \rightarrow \mathbb{Z}$ is multiplicative. Then $N(x + y\alpha) = 0$ if and only if $x = y = 0$.⁵ 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. If $\alpha = \sqrt{d}$, then $N(\sqrt{d}) = (\sqrt{d})(-\sqrt{d}) = -d < 0$, and

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

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

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

If $\mathbb{Z}[\alpha]$ is imaginary quadratic, then $\mathbb{Z}[\alpha]^\times = \mathbb{Z}[\alpha]^{\times,1}$ is finite. What are the possibilities for this group?⁶

What is $\mathbb{Z}[\sqrt{d}]^{\times,1}$? 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 $\mathbb{Z}[\sqrt{d}]^{\times,1}$. Then

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

Proof. If $x, y > 0$ then $x + y\sqrt{d} > y\sqrt{d} \geq \sqrt{d} > 1$. Then $x - y\sqrt{d} = 1/(x + y\sqrt{d}) \in (0, 1)$. 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. \square

⁵Exercise

⁶Exercise

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

Proof. $z - 1/z = x + y\sqrt{d} - (x - y\sqrt{d}) = 2y\sqrt{d}$, so just need to check that $z > z'$ if and only if $z - 1/z > z' - 1/z'$. But $z - 1/z$ is increasing, since its derivative is $1 + 1/z^2 > 0$. \square

Suppose that there exists $z \in \mathbb{Z}[\sqrt{d}]^{\times,1}$, 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

$$\epsilon = x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}]^{\times,1}, \quad x, y > 0,$$

with y as small as possible. Call ϵ the **fundamental 1-unit** of $\mathbb{Z}[\sqrt{d}]$.

Proposition 64. Suppose that $\mathbb{Z}[\sqrt{d}]^{\times,1} \neq \{\pm 1\}$, and let ϵ be the fundamental 1-unit. Then every element of $\mathbb{Z}[\sqrt{d}]^{\times,1}$ 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 $z \in \mathbb{Z}[\sqrt{d}]^{\times,1}$, 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 $1 \leq z\epsilon^{-n} < \epsilon$, and $N(z\epsilon^{-n}) = N(z)N(\epsilon)^{-n} = 1$. So $z\epsilon^{-n} \in \mathbb{Z}[\sqrt{d}]^{\times,1}$. So by the choice of ϵ , and Lemma 63, we have $z\epsilon^{-n} = 1$, that is $z = \epsilon^n$. \square

Example. Let $d = 2$ and $x^2 - 2y^2 = 1$. Then $y = 2$ and $x = 3$ is a solution. So $\epsilon = 3 + 2\sqrt{2}$. Then $\epsilon^2 = (3 + 2\sqrt{2})^2 = 17 + 12\sqrt{2}$, and $17^2 - 2(12)^2 = 1$.

8.5 Constructing the fundamental 1-unit

The idea is that if $x^2 - dy^2 = 1$ for $x, y > 0$, then $x/y \approx \sqrt{d}$. Then $|x - y\sqrt{d}| = 1/|x + y\sqrt{d}|$, which is small. So one way to try to find 1-units is to find rational numbers which are good approximations to \sqrt{d} . Want to make $|x/y - \sqrt{d}|$ 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

$$\left| \frac{x}{y} - \alpha \right| < \frac{C}{y^n},$$

where C and 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 and y with $|x/y - \alpha| < 1/y^2$, as we now show.

Theorem 65 (Dirichlet's theorem). 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

$$1 \leq q < Q, \quad |p - q\alpha| < \frac{1}{Q}.$$

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

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

The set

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

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 $1 \leq q < Q$. \square

Lecture 18
Wednesday
14/11/18

Corollary 66. For any $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, there exist infinitely many pairs $p, q \in \mathbb{Z}$ such that

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{q^2}.$$

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' and q' such that

$$\left| \alpha - \frac{p'}{q'} \right| < \frac{1}{(q')^2}, \quad \left| \alpha - \frac{p'}{q'} \right| < \left| \alpha - \frac{p}{q} \right|.$$

Choose Q such that $1/Q < |\alpha - p/q|$. By Theorem 65, there exist p' and q' with

$$1 \leq q' < Q, \quad \left| \alpha - \frac{p'}{q'} \right| < \frac{1}{Qq'} < \frac{1}{(q')^2}.$$

Also

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

as required. \square

We can now show the following.

Theorem 67. If $d > 1$ is squarefree, then there exist x and 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 $|p_i/q_i - \sqrt{d}| < 1/q_i^2$, that is $|p_i - q_i\sqrt{d}| < 1/q_i$. Then

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

So

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

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

$$p_i \equiv p_0 \pmod{M}, \quad q_i \equiv q_0 \pmod{M},$$

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

$$N(p_i + q_i\sqrt{d}) = N(p_j + q_j\sqrt{d}) = M, \quad p_i \equiv p_j \pmod{M}, \quad q_i \equiv q_j \pmod{M}.$$

Then

$$\frac{p_i - q_i\sqrt{d}}{p_j - q_j\sqrt{d}} = \frac{(p_i - q_i\sqrt{d})(p_j + q_j\sqrt{d})}{M} = \frac{(p_i p_j - dq_i q_j) + (p_i q_j - p_j q_i)\sqrt{d}}{M},$$

$$p_i q_j \equiv p_j q_i \pmod{M}, \quad p_i p_j - dq_i q_j \equiv p_i^2 - dq_i^2 = M \equiv 0 \pmod{M}.$$

So

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

so $(p_i - q_i\sqrt{d}) / (p_j - q_j\sqrt{d}) \in \mathbb{Z}[\sqrt{d}]^{\times, 1}$, as required. \square

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

$x^2 - dy^2 = -1$ has a solution if and only if there exists $u \in \mathbb{Z}[\sqrt{d}]^{\times}$ 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 $N(v) = -1$ if and only if $N(v) = N(u)$, if and only if $N(v/u) = 1$.

Example. If $d = 3$, there are no solutions, as $X^2 \equiv -1 \pmod{3}$ has no solutions.

Lecture 19
Friday
16/11/18

9 Continued fractions

9.1 Rational continued fractions

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

$$\frac{p}{q} = a_0 + r_0, \quad a_0 = \left\lfloor \frac{p}{q} \right\rfloor \in \mathbb{Z}, \quad 0 \leq r_0 < 1.$$

If $r_i \neq 0$, write

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

Eventually get some $r_n = 0$. Write

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

Example.

$$\frac{40}{19} = 2 + \frac{2}{19}, \quad \frac{19}{2} = 9 + \frac{1}{2} \quad \implies \quad \frac{40}{19} = 2 + \frac{1}{9 + \frac{1}{2+0}}.$$

9.2 Infinite continued fractions

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

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

Define sequences a_i and r_i by

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

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

$$\alpha = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots}}}.$$

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

$$\begin{aligned} a_0 &= 1, & r_0 &= \sqrt{3} - 1, & \frac{1}{r_0} &= \frac{1}{\sqrt{3} - 1} = \frac{\sqrt{3} + 1}{2} = 1 + \frac{\sqrt{3} - 1}{2}, \\ a_1 &= 1, & r_1 &= \frac{\sqrt{3} - 1}{2}, & \frac{1}{r_1} &= \frac{2}{\sqrt{3} - 1} = \sqrt{3} + 1 = 2 + (\sqrt{3} - 1), \\ a_2 &= 2, & r_2 &= \sqrt{3} - 1 = r_0, & \frac{1}{r_2} &= \frac{1}{\sqrt{3} - 1} = \frac{1}{r_0}, \end{aligned}$$

so

$$a_i = \begin{cases} 1 & i > 0 \text{ odd} \\ 2 & i > 0 \text{ even} \end{cases}.$$

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

$$[a_0; a_1, \dots, a_n] = a_0 + \frac{1}{a_1 + \frac{1}{\dots + \frac{1}{a_n}}}.$$

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

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

Assuming that no $q_i = 0$, we have $[a_0; a_1, \dots, a_n] = p_n/q_n$.

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 and q'_i for $0 \leq i \leq n-1$ by applying the definition to the sequence

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

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

$$\left[a_0; a_1, \dots, a_{n-2}, a_{n-1} + \frac{1}{a_n} \right] = \frac{p'_{n-1}}{q'_{n-1}}.$$

By definition,

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

So we only need to show that $p'_{n-1}/q'_{n-1} = p_n/q_n$, and

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

□

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 $a_0, a_1, \dots \in \mathbb{R}$ is an infinite sequence with $a_i \geq 1$ for all i , say that p_i/q_i is the i -th **convergent** to

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots}}}.$$

Lemma 69. For all n ,

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

Proof. Obvious for $n = 1$. For the 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. If $a_i \in \mathbb{Z}$, then $p_i, q_i \in \mathbb{Z}$, and Lemma 69 implies that $(p_n, q_n) = 1$. In general, Lemma 69 implies that

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

If $a_i \geq 1$ for all $i \geq 1$, then the sequence q_i increases exponentially. So $\sum_{i=1}^n 1/q_i q_{i-1}$ converges, so that (p_n/q_n) is a Cauchy sequence, so it converges.

Lemma 70. Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and let $[a_0; a_1, a_2, \dots]$ 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 .

$n = 0$. $a_0 = \lfloor \alpha \rfloor < \alpha$ and $p_0/q_0 = a_0/1 = a_0$.

n odd. By induction, we have $[a_1; a_2, \dots, a_n] < 1/(\alpha - a_0)$, since $\alpha = a_0 + 1/(a_1 + 1/\dots)$. That is, $\alpha - a_0 < 1/[a_1; a_2, \dots, a_n]$, that is

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

n even. The same argument with $>$.

□

Corollary 71. Assume $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and $a_0, a_1, \dots \in \mathbb{Z}$ be coming from its continued fraction. Let $p_n/q_n = [a_0; a_1, \dots, a_n]$ be the n -th convergent. Then

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

In particular, $p_n/q_n \rightarrow \alpha$ as $n \rightarrow \infty$.

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,

$$\left| \frac{p_n}{q_n} - \alpha \right| < \left| \frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}} \right| \leq \frac{1}{q_n q_{n+1}},$$

by Lemma 69.

□

Note. $1/q_n q_{n+1} < 1/q_n^2$, 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 and r_i by

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

If $i \geq 1$,

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

Lemma 72. For all n ,

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

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

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

□

Corollary 73. For all n ,

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

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 $|\alpha q_n - p_n| = r_n|\alpha q_{n-1} - p_{n-1}| < |\alpha q_{n-1} - p_{n-1}|$, so

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

□

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

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

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

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

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 and v have opposite signs, else

$$|k| = |uq_n| + |vq_{n+1}| \geq q_n + q_{n+1} > q_{n+1}.$$

If $v = 0$, then $h = up_n$ and $k = uq_n$, and everything is easy. If $v \neq 0$, then write

$$k\alpha - h = u(\alpha q_n - p_n) + v(\alpha q_{n+1} - p_{n+1}).$$

Then u and v have opposite signs. By Lemma 70, $\alpha q_n - p_n$ and $\alpha q_{n+1} - p_{n+1}$ 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 $|k\alpha - h|/|k| \geq |q_n\alpha - p_n|/q_n$, that is $|\alpha - h/k| \geq |\alpha - p_n/q_n|$. \square

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

$$\frac{h}{k} = \frac{p_n}{q_n},$$

for some n .

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

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

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

9.4 Returning to Pell's equation

Pell's equation is $x^2 - dy^2 = 1$. If (x, y) is a solution, then $|\sqrt{d} - x/y|$ 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 $1 = x^2 - dy^2 = r^2(p_n^2 - dq_n^2)$, so $r = 1$. By Corollary 75, it suffices to prove that $|\sqrt{d} - x/y| < 1/2y^2$. Since $x - y\sqrt{d} = 1/(x + y\sqrt{d}) > 0$, so $x > y\sqrt{d}$, and $x/y > \sqrt{d}$. So

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

Lecture 21
Wednesday
21/11/18

- 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 $y - x/\sqrt{d} = (1/d) / (y + x/\sqrt{d}) > 0$. So $y > x/\sqrt{d}$, so

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

So Corollary 75 implies that y/x is a convergent for the continued fraction of $1/\sqrt{d}$. Then $\left[1/\sqrt{d} \right] = 0$, so the continued fraction for $1/\sqrt{d}$ is of the form $[0; a_0, a_1, \dots]$. The next step is $1 / \left(1/\sqrt{d} \right) = \sqrt{d}$. So if $\sqrt{d} = [a_0; a_1, a_2, \dots]$, then $1/\sqrt{d} = [0; a_0, a_1, \dots]$, since

$$\sqrt{d} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots}}}, \quad \frac{1}{\sqrt{d}} = 0 + \frac{1}{a_0 + \frac{1}{a_1 + \frac{1}{\dots}}}.$$

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.

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

Definition 77. We say that $[a_0; a_1, a_2, \dots]$ 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 $N = 0$.

Fact 78. The following are facts.

- The continued fraction of \sqrt{d} is eventually periodic.
- In fact, it is of the form

$$[a_0; \overline{a_1, \dots, a_{m-1}, 2a_0}].$$

- a_1, \dots, 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 $n \equiv -1 \pmod{m}$. If $n = lm - 1$, then

$$p_n^2 - dq_n^2 = (-1)^{lm}.$$

- The fundamental 1-unit is

$$\begin{cases} p_{m-1} + q_{m-1}\sqrt{d} & m \text{ even} \\ p_{2m-1} + q_{2m-1}\sqrt{d} & m \text{ odd} \end{cases}.$$

- There is a solution to $x^2 - dy^2 = -1$ if and only if m is odd, in which case the solutions are

$$(x, y) = (p_n, q_n), \quad n \equiv m - 1 \pmod{2m}.$$

Example.

- Let $x^2 - 43y^2 = \pm 1$, so $m = 10$ is even, so there are no solutions to $x^2 - 43y^2 = -1$. The smallest solution for $x^2 - 43y^2 = 1$ is p_9 and q_9 . Then

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

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

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

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

so $18^2 - 13(5)^2 = -1$ is the smallest solution, and $N(18 + 5\sqrt{13}) = -1$, so

$$N\left((18 + 5\sqrt{13})^2\right) = N(649 + 180\sqrt{13}) = 1.$$

In fact, it follows from our facts that this is the fundamental 1-unit, that is $p_9 + q_9\sqrt{13}$.

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.

Lecture 22
Friday
23/11/18

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

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

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 $[a_0; a_1, \dots, a_{d-1}, \alpha]$. 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

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

where β has a 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 $1/\gamma$ and $\gamma + n$ are quadratic irrationals 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, and there does not exist such a polynomial of smaller degree.

Example. $d = 1$ is \mathbb{Q} and $d = 2$ is the 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 $P(X) = (X - \alpha)Q(X)$. Then $Q(X)$ 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 $x \in [\alpha - \epsilon, \alpha + \epsilon]$. Assume that $|p/q - \alpha| < 1/q^e$. 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 $|p/q - \alpha| < \epsilon$, $p/q \in [\alpha - \epsilon, \alpha + \epsilon]$, so

$$\frac{1}{q^d} \leq \left| P\left(\frac{p}{q}\right) \right| = \left| \frac{p}{q} - \alpha \right| \left| Q\left(\frac{p}{q}\right) \right| \leq K \left| \frac{p}{q} - \alpha \right| < \frac{K}{q^e}.$$

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

10.2 Constructing transcendentals

Recall that $\alpha \in \mathbb{C}$ is **algebraic** if it is algebraic of some degree d , and otherwise it is **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 $\alpha = \sum_{n=1}^{\infty} 10^{-n!}$ and $\alpha_k = \sum_{n=1}^k 10^{-n!}$. Then α_k is rational with denominator $q = 10^{k!}$, and

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

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

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 $|\alpha - x/y| < 1/y^{2+\epsilon}$.*

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

Example. $\sum_{n=1}^{\infty} 10^{-3^n}$ 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 $(x, y) = (0, 1)$ and $d = 1$.
- d is not a cube. Then $\sqrt[3]{d} \in \mathbb{R} \setminus \mathbb{Q}$ is algebraic, as $X^3 - d = 0$. Suppose $x > 1$, so $x > \sqrt[3]{d}y$. Then

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

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

Lecture 23
Tuesday
27/11/18

11 Primes in arithmetic progressions

11.1 Primes in arithmetic progressions

A question is how are the prime numbers distributed modulo n ? Are there infinitely many primes congruent to $a \pmod n$ for each a and n ? The answer is no in general.

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

If $(a, n) \neq 1$ then since any number is congruent to $a \pmod n$ 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 $1 \pmod 2$.

Theorem 84 (Dirichlet's theorem). *If $(a, n) = 1$, then there are infinitely many primes congruent to $a \pmod n$.*

We will prove this for $a = 1$.

11.2 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. \square

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

Proof. Let S be a finite set of primes which are congruent to $3 \pmod 4$. Let

$$Q = 2 + \prod_{p \in S} p^2.$$

Then $Q > 1$, and $Q \equiv 3 \pmod 4$, so Q has a prime factor q which has $q \equiv 3 \pmod 4$. Then $q \notin S$, so we are done. \square

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

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

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

Proof. Let S be a finite set of primes congruent to $1 \pmod 4$. 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 $q \equiv 1 \pmod 4$ by Lemma 87. \square

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

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

Definition 90. The q -th cyclotomic polynomial is

$$\Phi_q(X) = \frac{X^q - 1}{X - 1} = X^{q-1} + \cdots + 1.$$

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 $q \pmod p$.

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

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

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

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

Proof. Let S be a finite set of primes which are congruent to $1 \pmod q$. 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 $p \equiv 1 \pmod q$. Since $\Phi_q(qR) = (qR)^{q-1} + \cdots + 1 \equiv 1 \pmod{qR}$, so $p \neq q$, $p \notin S$, and $p \equiv 1 \pmod q$. \square

11.3 Cyclotomic polynomials

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

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

Lemma 95. For any n , we have

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

Proof. Each side is a monic polynomial, so we just need to check that the roots are the same, with multiplicities. The left hand side are the n -th roots of unity, with multiplicity one each. The 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. \square

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 $d \mid n$ for $d < n$. By Lemma 95, if we set

$$P(X) = \prod_{d \mid n, 0 < d < n} \Phi_d(X),$$

then $P(X) \in \mathbb{Z}[X]$, $P(X)$ is monic, and $X^n - 1 = \Phi_n(X)P(X)$. Write

$$\Phi_n(X) = \sum_i a_i X^i, \quad P(X) = \sum_i b_i X^i,$$

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 $P(X) = X^e + b_{e-1}X^{e-1} + \cdots + b_0$. Then the coefficient of X^{q+e} in $\Phi_n(X)P(X)$ is

$$a_q + a_{q+1}b_{e-1} + \cdots + a_{q+e}b_0, \quad a_{q+1}b_{e-1} + \cdots + a_{q+e}b_0 \in \mathbb{Z}.$$

Since $\Phi_n(X)P(X) = X^n - 1 \in \mathbb{Z}[X]$, this is a contradiction. \square

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. $(P+Q)' = P' + Q'$ 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

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

□

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

Proof. It suffices to show that $X^n - 1$ has no repeated roots modulo 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. If $n = p$,

$$X^p - 1 \equiv (X - 1)^p \pmod{p}, \quad \Phi_p(X) \equiv (X - 1)^{p-1} \pmod{p}.$$

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 $n \pmod{p}$.

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

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

so a is a root of $\Phi_n(X) \pmod{p}$, that is $p \mid \Phi_n(a)$. Conversely, suppose that $p \mid \Phi_n(a)$. Then a is a root of $\Phi_n(X) \pmod{p}$, so by (4), a is a root of $X^n - 1 \pmod{p}$. We need to show that a is not a root of $X^d - 1$ for any $d \mid n$ for $d < n$. Writing

$$X^d - 1 = \prod_{e \mid d} \Phi_e(X),$$

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 $X^n - 1 \pmod{p}$. This contradicts Corollary 99. □

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

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

11.4 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 $p \equiv 1 \pmod{n}$.

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 $Q_k = \Phi_n(knR) \in \mathbb{Z}$. 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 problems class.

Lecture 25
Friday
30/11/18

12 Arithmetic functions

Lecture 26
Tuesday
04/12/18

An **arithmetic function** is a function $f : \mathbb{Z}_{\geq 1} \rightarrow \mathbb{C}$, 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

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

Then $f * (g + h) = f * g + f * h$. 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. $f(4) = f(4)\epsilon(1) + f(2)\epsilon(2) + f(1)\epsilon(4)$. This forces $\epsilon(1) = 1$ and $\epsilon(2) = \epsilon(4) = 0$.

So

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

12.2 Möbius inversion

The **Möbius function** $\mu : \mathbb{Z}_{\geq 1} \rightarrow \mathbb{C}$ is defined as follows.

$$\mu(n) = \begin{cases} 1 & n = 1 \\ (-1)^k & n = p_1 \dots p_k \text{ is a product of distinct prime factors} \\ 0 & \text{otherwise} \end{cases}.$$

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

$$\sum_{d|n} \mu(d) = \sum_{(\epsilon_1, \dots, \epsilon_k), \epsilon_i=0,1} (-1)^{\epsilon_1 + \dots + \epsilon_k} = \left(\sum_{\epsilon_1=0}^1 (-1)^{\epsilon_1} \right) \dots \left(\sum_{\epsilon_k=0}^1 (-1)^{\epsilon_k} \right) = 0, \quad n = \prod_{i=1}^k p_i^{\epsilon_i}.$$

□

Proposition 104 (Möbius inversion). *If f and g are arithmetic functions then $g = f * 1$ if and only if $f = g * \mu$.*

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

$A \ll B$, or $A = O(B)$, means there exists a constant $C > 0$ such that $|A| \leq CB$, and $B \gg A$ means $A \ll B$.

Example. If $x \geq 1$, $x \ll x^2 \ll e^x/x^{100}$.

$A \ll_k B$ means $A \ll B$ with the constant C depending on k .

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.

Example. $1/\log x = o(1)$ as $x \rightarrow \infty$.

$A \sim B$ means $A = (1 + o(1))B$.

13.2 The prime number theorem

Theorem 105 (Prime number theorem).

$$\pi(X) \sim \frac{X}{\log X}, \quad X \rightarrow \infty.$$

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 implies that

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

Proof.

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

$$\prod_{p \leq 2n} \geq C_1^n. \quad (5)$$

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

Lecture 27
Wednesday
05/12/18

Suppose 1 to 3 are true. Then

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

So

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

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. ⁷

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

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

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

$$i > \frac{\log 2n}{\log p} \implies v_p(n) \leq \frac{\log 2n}{\log p} \implies p^{v_p(n)} \leq 2n.$$

- 3.

$$4^n = 2^{2n} = (1+1)^{2n} = \sum_{i=0}^{2n} \binom{2n}{i} \leq (2n+1) \binom{2n}{n} \implies \prod_{p \leq 2n} p^{v_p(n)} = \binom{2n}{n} \geq \frac{4^n}{2n+1}.$$

- 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

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

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 $X, \dots, X/2^{m-1}$. Substituting $X, \dots, X/2^{m-1}$ into (7), and summing,

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

⁷Exercise

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 $X, \dots, X/2^{m-1}$ 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,

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

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

$$\prod_{\frac{X}{2} < p \leq 2 \lfloor \frac{X}{2} \rfloor} p \leq 2^{2n} \leq 2^X \implies \prod_{\frac{X}{2} < p \leq X} p \leq X 2^X < C_2^X,$$

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

□

Lecture 28 is a problems class.

Lecture 28
Friday
07/12/18

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. The best possible lower bound is zero.

Lecture 29
Tuesday
11/12/18

Example. $n! + 2, \dots, n! + n$ 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 $Y \rightarrow \infty$ and X is fixed.

In $X + 1, \dots, X + Y$, about half of these are divisible by two, about a third of these are divisible by three, and about a sixth of these are divisible by six. If p_1, \dots, p_k are primes, the error term is 2^k , so can only consider the first $\log Y$ primes, which implies Theorem 107 for $Y/\log \log Y$. Selberg's idea is to weight the inclusion-exclusion count.

Proof. Let $\lambda_1, \lambda_2, \dots \in \mathbb{R}$ be any sequence with $\lambda_1 = 1$. Let $R < Y$ be fixed for now. Later we will choose $R = Y^{\frac{1}{2}-\epsilon}$. 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, $\nu(p) = \lambda_1^2 = 1$, so

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

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

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

since $\text{lcm}(d_1, d_2) = d_1 d_2 / (d_1, d_2)$. 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

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

and the error term is

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

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 $c < \frac{1}{2}$. Check that for any $\epsilon > 0$, we have $\lambda_d \ll_\epsilon d^\epsilon$. Then

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

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

$$Q(\vec{\lambda}) = \sum_{d_1, d_2 \leq R} \frac{\lambda_{d_1} \lambda_{d_2} (d_1, d_2)}{d_1 d_2}.$$

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

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

by using that $\delta | d_1$ and $\delta | d_2$ if and only if $\delta | d_1 d_2 / (d_1, d_2)$. Set $u_\delta = \sum_{\delta|d, d \leq R} \lambda_d / d$. Then

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

Claim that

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

The 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

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

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

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

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

$$Q(\vec{\lambda}) = \sum_{\delta \leq R} \Phi(\delta) u_\delta^2 \geq \frac{1}{D}, \quad D = \sum_{\delta \leq R} \frac{\mu(\delta)^2}{\Phi(\delta)}.$$

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

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

$c < \frac{1}{2}$ implies that $1/c > 2$, so

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

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

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

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

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

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

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

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

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

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, where $\Phi(d) = \prod_{p|d} (p-1)$. 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