Math 104C: Number Theory III (Spring 2015)

Alina Bucur

Contents

1 Review

1.1 Prime numbers

A prime number p has the following properties:

- p has no other divisors than 1 and p;
- $p \mid ab \implies p \mid a$ or $p \mid b$.

There are infinitely many primes. Every positive integer can be written uniquely as a product of primes.

1.2 Euclidean algorithm

The algorithm is used to find the *greatest common divisor* $d = (a, b)$ of two positive integers a and b. It also can be used to find integers r, s such that

$$d = ar + bs.$$

1.3 Congruences

Definition. *We say that two integers a and b are* congruent modulo *some integer n and write $a \equiv b \pmod{n}$ if $n \mid a - b$. (That is to say, a and b give the same remainder when divided by n.)*

Here a few properties of congruences:

- $a \equiv a \pmod{n}$
- $a \equiv b \pmod{n} \iff b \equiv a \pmod{n}$
- $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n} \implies a \equiv c \pmod{n}$
- $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n} \implies a \pm c \equiv b \pm d \pmod{n}, ac \equiv bd \pmod{n}$.
- $(a, n) = d$ and $ab \equiv ac \pmod{n} \implies b \equiv c \pmod{\frac{n}{d}}$.
- Given integers a and n, the equation $ax \equiv b \pmod{n}$ has solutions if $(a, n) \mid b$. Therefore it has solutions for all b iff $(a, n) = 1$. That is to say, if there exists an integer c such that $ac \equiv 1 \pmod{n}$. If such a c exists, it is unique modulo n, and the solution is $x = bc \pmod{n}$ is also unique modulo n.
- $a^{\phi(n)} \equiv 1 \pmod{n}$.

In addition to all these similarities to normal arithmetic operations (addition, subtraction, multiplication, division), there are similarities to linear algebra as well. For instance, the system of linear congruences

$$\begin{cases} a_{11}x_1 + \ldots + a_{1r}x_r \equiv b_1 \pmod{n} \\ \vdots \\ a_{r1}x_1 + \ldots + a_{rr}x_r \equiv b_r \pmod{n} \end{cases}$$

has unique solution $\pmod{n}$ iff $\det(a_{ij})$ and n are coprime.

Theorem 1.1 (Chinese Remainder Theorem)**.** *Assume that $m_1, \ldots, m_r$ are positive integer with the property that any two of them are relatively prime. Then, for any $a_1, \ldots, a_r \in \mathbb{Z}$, the system of equations*

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

has a unique solution $\pmod{m_1 \ldots m_r}$.

1.4 Groups

Definition. *A group G is a set endowed with an operation $\circ : G \times G \to G$ with following properties.*

(i) *$x \circ (y \circ z) = (x \circ y) \circ z$ for all $x, y, z \in G$ (associativity).*

(ii) *There exists an element $e \in G$ (called the unit of the group) such that $x \circ e = e \circ x = x$ for all $x \in G$.*

(iii) *For each $x \in G$ there exists an element $x^{-1} \in G$ (the inverse of x) such that*

$$x \circ x^{-1} = x^{-1} \circ x = e.$$

Definition. *We say that a group G is abelian (commutative) if $x \circ y = y \circ x$ for all $x, y \in G$.*

Theorem 1.2 (Lagrange)**.** *In a finite group G the order of every element is a divisor of the order of the group $\#G$. In particular*

$$x^{\#G} = e \text{ for all } x \in G.$$

Two particular cases are the following result in modular arithmetic.

Theorem 1.3 (Fermat)**.** *If p is a prime and a an integer not divisible by p, then*

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

Theorem 1.4 (Euler)**.** *If a and n are relatively prime nonzero integers, then*

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

1.5 Powers and roots in modular arithmetic

E.g. find x such that $x^5 \equiv 2 (\text{mod } 17)$.

1.6 Gaussian integers

The ring $\mathbb{Z}[i]$ is an Euclidean domain with respect to the norm $N(a + bi) = a^2 + b^2 = |z|^2$ where $z = a + bi \in \mathbb{C}$.

Theorem 1.5. *For any $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$ there exist $\gamma, \rho \in \mathbb{Z}[i]$ such that $\alpha = \beta\gamma + \rho$ and $0 \leq N(\rho) < N(\beta)$.*

Note that the norm map is completely multiplicative, i.e. $N(\alpha\beta) = N(\alpha)N(\beta)$.
The units in $\mathbb{Z}[i]$ are $\pm 1, \pm i$.

Theorem 1.6. *Let $p \in \mathbb{Z}$ be a (positive) prime. Its factorization in $\mathbb{Z}[i]$ is determined by its residue class modulo 4 as follows.*

(i) $2 = (1+i)(1-i) = -i(1+i)^2 = i(1-i)^2$ *and* $1+i = i(1-i)$ *represent the same prime ideal in* $\mathbb{Z}[i]$.

(ii) If $p \equiv 1 (\text{mod } 4)$ then $p = \pi\bar{\pi}$ where $\pi, \bar{\pi}$ are two prime gaussian integers that are complex conjugates, but not unit multiples. In particular, they generate different prime ideals in $\mathbb{Z}[i]$.

(iii) If $p \equiv 3 (\text{mod } 4)$ then p is prime in $\mathbb{Z}[i]$.

Theorem 1.7. *Every prime gaussian integer is a unit multiple of one of the following primes:*

(i) $1 + i$;

(ii) π *or* $\bar{\pi}$ *where* $N(\pi) = p$ *is a prime integer* $p \equiv 1 (\text{mod } 4)$;

(iii) a prime p in $\mathbb{Z}$ with $p \equiv 3 (\text{mod } 4)$. In this case $N(p) = p^2$.

1.7 Diophantine equations and congruences

We can try to show that a diophantine equation does not have solutions by showing that it has no solution modulo some integer n.

Example 1 $x^2 - 3y^2 = -1$

Looking at this equation modulo 3, we see that

$$x^2 \equiv -1 \pmod{3},$$

which we know it is impossible since $3 \nmid 1$.

Example 2 $x^2 - 7y^2 = -1$

This implies that $x^2 + 1 \equiv 0 \pmod 7$ and that is impossible since 7 is a prime and $7 \equiv 3 \pmod 4$.

Example 3 $x^2 - 15y^2 = 2$

This implies that $x^2 \equiv 2 \pmod 5$. But the only squares modulo 5 are $0, 1, 4$.

Example 4 $x^2 - 5y^2 = 3z^2$

Assume that we have a positive solution with $(x, y, z) = d$. Then $x = dx_1, y = dy_1, z = dz_1$ with $(x_1, y_1, z_1) = 1$ and

$$x_1^2 - 5y_1^2 = 3z_1^2.$$

In particular, $3 \mid x_1^2 - 5y_1^2$ and, since obviously $3 \mid 6y_1^2$, we get $3 \mid x_1^2 + y_1^2$. We know that this is only possible if $3 \mid x_1$ and $3 \mid y_1$. But then $9 \mid 3z_1^2$ and so $3 \mid z_1$. This cannot happen since $(x_1, y_1, z_1) = 1$.

2 Primes of the form $p = x^2 + ny^2$

This is recap from 104B. We proved that a prime p can be written as the sum of two squares if and only if $p = 2$ or $p \equiv 1 \pmod 4$. One direction was easy, but the other one was completely non-trivial. The proof consisted of two steps.

Reciprocity step: A prime $p \equiv 1 \pmod 4$, then it divides $N = a^2 + b^2$ with a and b relatively prime integers.

Descent step: If a prime p divides a number N of the form $N = a^2 + b^2$, where $(a, b) = 1$, then p itself can be written as $p = x^2 + y^2$ for some $(x, y) = 1$.

This step was based on the descent lemma which said that if a prime $q = x^2 + y^2$ divides a sum of squares $a^2 + b^2 = N$ with $(a, b) = 1$, then N/q can be written as a sum of relatively prime squares.

Furthermore, we used in an essential way the fact that if a number N is the sum of two squares, then all its prime divisors can be written as sums of two squares.

One can look at other questions of this type. For instance, Fermat himself stated (and Euler proved) the following two results.

Theorem 2.1. *A prime p is of the form $p = x^2 + 2y^2$ if and only if $p = 2$ or $p \equiv 1$ or 3* $\pmod 8$.

Theorem 2.2. *A prime p is of the form $p = x^2 + 3y^2$ if and only if $p = 3$ or $p \equiv 1$* $\pmod 3$.

2.1 Reciprocity step

We need to find congruence conditions which will guarantee that $p \mid x^2 + ny^2$ for some $(x, y) = 1$.

The problem is that we cannot adapt directly our proof from the $n = 1$ case (gaussian primes from 104B). This is because our proof was done in an ad-hoc manner. Namely, to recap, we said that if $p \equiv 1 \pmod 4$, then $\phi(p) = 4k$ for some integer k. Therefore the polynomial $X^{4k} - 1$ has $4k$ roots $\pmod p$. But

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

Since $X^{2k} - 1$ can have at most $2k$ roots $\pmod p$, it follows that there must exist an integer $(a, p) = 1$ such that $a^{2k} + 1 \equiv 0 \pmod p$. Thus $p \mid (a^k)^2 + 1^2$ and since a^k and 1 are relatively prime, we are done.

But this cannot be replicated directly for $n = 2$ for instance.

One more thing that is worth noticing. We have the following conjectures (due to Fermat).

- $n = 1 : p \equiv 1 \pmod 4 \implies p \mid a^2 + b^2$ for some $(a, b) = 1$.
- $n = 2 : p \equiv 1, 3 \pmod 8 \implies p \mid a^2 + 2b^2$ for some $(a, b) = 1$.
- $n = 3 : p \equiv 1 \pmod 3 \implies p \mid a^2 + 3b^2$ for some $(a, b) = 1$.

The key observation is that these are all congruences modulo $4n$. (The last one can be restated as $p \equiv 1, 7 \pmod{12}$.) And indeed, we are going to find conditions $\pmod{4n}$ that would ensure that a prime p is of the form $x^2 + ny^2$. A systematic approach was formulated in terms of Jacobi symbols (see Section 2.1.1).

2.1.1 Quadratic reciprocity

The Legendre symbol modulo an odd prime p is the function $\mathbb{Z} \to \mathbb{C}$ given by

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \mid a \\ 1 & \text{if } p \nmid a \text{ and } a \text{ is a quadratic residue } \pmod p\text{, i.e. } \exists x\, s.t.\, a \equiv x^2 \pmod p \\ -1 & \text{if } p \nmid a \text{ and } a \text{ is a quadratic nonresidue } \pmod p. \end{cases}$$

The Jacobi symbol is an extension of the Legendre symbol. Let m be an odd positive integer. The Jacobi symbol modulo m is given by

- if $m = 1$, the Jacobi symbol $\left(\frac{-}{1}\right) : \mathbb{Z} \to \mathbb{C}$ is the constant function 1;

- if $m > 1$, it has a decomposition as a product of (not necessarily distinct) primes $m = p_1 \cdots p_r$. The Jacobi symbol $\left(\frac{\ }{m}\right) : \mathbb{Z} \to \mathbb{C}$ is given by

$$\left(\frac{a}{m}\right) = \left(\frac{a}{p_1}\right) \cdots \left(\frac{a}{p_r}\right).$$

Note: The Jacobi symbol does not necessarily distinguish between quadratic residues and nonresidues.

Proposition 2.3. *Let m, n be positive odd integers and $a, b \in \mathbb{Z}$. Then*

(i) $\left(\frac{1}{m}\right) = 1;$

(ii) $\left(\frac{a}{m}\right) = 0 \iff (a, m) > 1;$

(iii) $a \equiv b \pmod{m} \implies \left(\frac{a}{m}\right) = \left(\frac{b}{m}\right);$

(iv) $\left(\frac{ab}{m}\right) = \left(\frac{a}{m}\right)\left(\frac{b}{m}\right);$

(v) $\left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right);$

(vi) $(a, m) = 1 \implies \left(\frac{a^2 b}{m}\right) = \left(\frac{b}{m}\right).$

Theorem 2.4. *Let m, n be positive odd integers. Then*

(i) $\left(\frac{-1}{m}\right) = (-1)^{\frac{m-1}{2}};$

(ii) $\left(\frac{2}{m}\right) = (-1)^{\frac{m^2-1}{8}};$

(iii) $\left(\frac{n}{m}\right) = (-1)^{\frac{m-1}{2}\frac{n-1}{2}}\left(\frac{m}{n}\right)$ *(quadratic reciprocity).*

Proposition 2.5. *If m, n are positive odd integers and is an integer with $D \equiv 0, 1 \pmod{4}$ such that $m \equiv n \pmod{D}$, then*

$$\left(\frac{D}{m}\right) = \left(\frac{D}{n}\right).$$

Theorem 2.6. *Let $D \equiv 0, 1 \pmod 4$ be a nonzero integer. Then there exists a* unique *group homomorphism $\chi_D : (\mathbb{Z}/D\mathbb{Z})^\times \to \{\pm 1\}$ such that*

$$\chi_D([p]) = \left(\frac{D}{p}\right) \textit{ (the Legendre symbol modulo } p\textit{) for all odd primes } p \nmid D.$$

Furthermore,

$$\chi_D([-1]) = \begin{cases} 1 & \textit{if } D > 0; \\ -1 & \textit{if } D < 0 \end{cases}$$

and, for $D \equiv 1 (\mathrm{mod}\ 4)$,

$$\chi_D([2]) = \begin{cases} 1 & \textit{if } D \equiv 1 \pmod 8; \\ -1 & \textit{if } D \equiv 5 \pmod 8. \end{cases}$$

Corollary 2.7. *Let n be a nonzero integer and let $\chi = \chi_{-4n} : (\mathbb{Z}/4n\mathbb{Z})^\times \to \{\pm 1\}$ be the group homomorphism defined in Theorem 2.6 when $D = -4n$. Let p be an odd prime, $p \nmid n$. The following are equivalent.*

(i) $p \mid a^2 + nb^2$ *for some integers* $(a, b) = 1$.

(ii) $\left(\frac{-n}{p}\right) = 1.$

(iii) $[p] \in \ker \chi \subset (\mathbb{Z}/4n\mathbb{Z})^\times$.

Note that this finishes the Reciprocity Step from Euler's strategy because if $\ker(\chi) = \{[\alpha], [\beta], [\gamma], \ldots\}$, Corollary 2.7 says that

$$p \mid a^2 + nb^2, (a, b) = 1 \iff p \equiv \alpha, \beta, \gamma, \ldots \pmod{4n}.$$

This is precisely the kind of condition we were looking for.

Again, it is easy to show that if the prime has the given form in terms of squares, then it lands in the desired congruence class. For the other direction, let us try to imitate the procedure from last quarter.

2.2 Descent step

We tackled the first component of our descent step for n by generalizing the descent lemma.

Lemma 2.8. *Fix $n \in \mathbb{Z}_{>0}$. Suppose M is an integer of the form $M = a^2 + nb^2$ with $(a, b) = 1$ and that $q = x^2 + ny^2$ is a prime divisor of M. Then there exist integers $(c, d) = 1$ such that $M/q = c^2 + nd^2$.*

Then we looked at the second component of the descent step. That is we would like to say that

$$p \text{ prime}, p \mid a^2+nb^2 \text{ with } (a,b)=1 \implies p = x^2+ny^2. \tag{2.1}$$

Without loss of generality, we can assume that

$$|a|,|b| \leq \frac{p}{2}.$$

Then, if p is odd

$$a^2+nb^2 < \frac{n+1}{4}p^2.$$

If $n \leq 3$, this implies that $a^2+nb^2 < p^2$ and therefore any prime divisor $q \neq p$ of a^2+nb^2 has to be $q < p$. Now completed the proof of the descent step for $n = 1,2,3$ in 104B. That is, we proved Theorems 2.1 and 2.2.
Note that (2.1) *cannot* hold in general. For instance, in the case $n = 5$ we see that $3 \mid 21 = 1^2+5\cdot 2^2$, but 3 cannot be written as x^2+5y^2. So we need to figure out how the prime divisors of a^2+nb^2 can be represented. The answer will come from Legendre's theory of reduced quadratic forms. (See Section 3.)

3 Quadratic forms

This marks the start of the new material for 104C. There are two examples we need to keep in mind. Euler made two conjectures regarding the cases $n = 5$ and $n = 14$. First, let's see what the reciprocity step says.
For $n = 5$, we need to look at congruence classes in $(\mathbb{Z}/20\mathbb{Z})^\times$. We can look at them one by one and, using Corollary 2.7, see that

$$p \mid a^2+5b^2, (a,b)=1 \iff p \equiv 1,3,7,9 \pmod{20}.$$

But here's Euler's conjecture (and of course, he had good numerical evidence for it). We have seen that not all divisors of a number of the form a^2+5b^2 can be written in the same form, which momentarily derailed our strategy. Indeed, things are more complicated in this case and we need to understand what forms the divisors of a^2+5b^2 can have.

Conjecture 3.1 (Euler)**.** *If $p \neq 5$ is an odd prime, then*

$$\begin{aligned} p = x^2+5y^2 &\iff p \equiv 1,9 \pmod{20} \\ 2p = x^2+5y^2 &\iff p \equiv 3,7 \pmod{20} \end{aligned}$$

The congruence classes break into two groups – $1,9$ and $3,7$ – that have very different representability properties. To see what's going on, recall that we have seen that not all divisors of a number of the form a^2+5b^2 can be written in the same form.
The case $n = 14$ is even more complicated.

Conjecture 3.2 (Euler)**.** *If $p \neq 7$ is an odd prime, then*

$$p = \begin{cases} x^2 + 14y^2 \\ 2x^2 + 7y^2 \end{cases} \iff p \equiv 1, 9, 15, 23, 25, 39 \pmod{56}$$
$$3p = x^2 + 14y^2 \iff p \equiv 3, 5, 13, 19, 27, 45 \pmod{56}$$

As in the previous case, the congruence classes modulo 56 that appear above are precisely the ones for which $\left(\dfrac{-14}{p}\right) = 1$. A new feature is that $x^2 + 14y^2$ and $2x^2 + 7y^2$ appear together. Another question is where the $2p$ in Conjecture 3.1 and $3p$ in Conjecture 3.2 come from. Why are they different multiples of p? Why 2 and 3 appear there, and not, say, 29? Gauss composition explains this phenomenon. What other condition is necessary to ensure $p = x^2 + 14y^2$? This is a much deeper question and the answer involves class field theory which is outside the scope of this class. For now, it should be clear that we need to know more about these quadratic polynomials.

Definition. *An integral binary quadratic form (for short, integral bqf) is a degree 2 homogeneous polynomial in two variables with integer coefficients, i.e. $f(x, y) = ax^2 + bxy + cy^2$, $a, b, c, \in \mathbb{Z}$.*

Note: One can define binary quadratic forms over any commutative ring R. In particular, they can be defined over $\mathbb{Q}$ or $\mathbb{R}$.

Definition. *An integral binary quadratic form $f(x, y) = ax^2 + bxy + cy^2$ is primitive if $(a, b, c) = 1$.*

Note: Any integral form is an integer multiple of a primitive form.

Definition. *Two bqf's $f(x, y)$ and $g(x, y)$ are equivalent if there are integers $\alpha, \beta, \gamma, \delta$ such that $f(x, y) = g(\alpha x + \beta y, \gamma x + \delta y)$ and $\alpha\delta - \beta\gamma = \pm 1$. In linear algebra terms, this just says that there exists a matrix $A \in \mathrm{GL}(2, \mathbb{Z})$ – the group of 2×2 invertible matrices with coefficients in $\mathbb{Z}$ –such that*

$$f(\vec{x}) = g(A\vec{x})$$

Note: You can think of the bqf $f(x, y) = ax^2 + bxy + cy^2$ as

$$f(\vec{x}) = {}^t\vec{x} \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \vec{x}$$

Proposition 3.3. *The above definition describes indeed an equivalence relation.*

Proof. Exercise. □

Definition. *We say that the equivalence of two bqf's is a proper equivalence if $\alpha\delta - \beta\gamma = 1$ (i.e. the matrix $A \in \mathrm{SL}(2,\mathbb{Z})$ – the subgroup of $\mathrm{GL}(2,\mathbb{Z})$ that consists of matrices with determinant equal to 1). It is called an improper equivalence otherwise (i.e. $\alpha\delta - \beta\gamma = -1 \iff \det A = -1$).*

Proposition 3.4. *Proper equivalence is indeed an equivalence relation.*

Proof. Exercise. □

Note: The terms "equivalence" and "proper equivalence" are due to Gauss. He had good reason to distinguish between the two notions.

3.1 Group actions

Let $(G, \cdot)$ be a group (with unit e) and X be a set.

Definition. *A* left action *of the group G on the set X is a map $G \times X \to X$ such that*

(*i*) *$e * x = x$ for all $x \in X$;*

(*ii*) *$(g_1 \cdot g_2) * x = g_1 * (g_2 * x)$ for all $x \in X$ and all $g_1, g_2 \in G$.*

Note that $g^{-1} * (g * x) = g * (g^{-1} * x) = x$ for all $x \in X$ and $g \in G$.

Definition. *A* right action *of G on X is a map $X \times G \to X$ such that*

(*i*) *$x \perp e = x$ for all $x \in X$;*

(*ii*) *$x \perp (g_1 \cdot g_2) = (x \perp g_1) \perp g_2$ for all $x \in X$ and all $g_1, g_2 \in G$.*

Note that $(x \perp g^{-1}) \perp g = (x \perp g) \perp g^{-1} = x$ for all $x \in X$ and $g \in G$.

Example 3.5. The symmetric group S_n acts on the set $X = \{1, \ldots, n\}$ by $\sigma * n = \sigma(n)$. This is a left action.

Example 3.6. A group G acts on itself by left multiplication (left action).

Example 3.7. Last quarter we defined

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (z) = \frac{az+b}{cz+d}.$$

This represents a left action of the group $\mathrm{GL}(2,\mathbb{R})$ of invertible 2×2 matrices with real coefficients on the set $\mathbb{C}$ of complex numbers.

Example 3.8. The group $\mathrm{GL}(2,\mathbb{Z})$ acts on the right on the set of primitive bqfs via

$$(f \cdot A)(\vec{x}) = f(A\vec{x}).$$

Indeed,

$$f(\vec{x}) = {}^t\vec{x} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \vec{x},$$

hence

$$(f \cdot A)(\vec{x}) = f(A\vec{x}) = {}^t\vec{x}\ {}^tA \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} A\,\vec{x}. \tag{3.1}$$

Clearly, $f \cdot I_2 = f$ and

$$(f\cdot(AB))(\vec{x}) = {}^t\vec{x}\ {}^t(AB) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} (AB)\vec{x} = {}^t\vec{x}\,{}^tB \left({}^tA \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} A \right) B\vec{x} = ((f\cdot A)\cdot B)(\vec{x}).$$

Remark 3.9. If H is a subgroup of G and G acts (left or right) on X, then so does H. That is, $H \times X \to X$ given by the restriction of the map $G \times X \to X$ is an action of H on X.

Definition. *Assume G acts on X on the left. The orbit of an element $x \in X$ under the action of G is the set*

$$Gx = \{g \cdot x; g \in G\} \subset X.$$

If we are dealing with a right action, the orbit is

$$xG = \{x \cdot g; g \in G\} \subset X.$$

The main observation is that two orbits are either equal or disjoint. This in turn implies that being in the same orbit defines an equivalence relation on X.

Proposition 3.10. *If $x, y \in X$ then either $Gx = Gy$ or $Gx \cap Gy = \emptyset$.*

Proof. If $z \in Gx \cap Gy$ there exist $g, h \in G$ such that $z = gx = hy$. Thus $x = g^{-1}hy \in Gy \implies Gx \subset Gy$. Similarly, $Gy \subset Gx$. □

Corollary 3.11. *Assume the group G acts on the set X. Then the following is an equivalence relation:*

$$x \sim y \iff x, y \text{ are in the same } G\text{-orbit} \iff \text{there exists } g \in G \text{ s.t. } y = gx.$$

Proof. $x = ex$ so $x \sim x$. If $x \sim y$ then $y = gx$ for some $g \in G$. Thus $x = g^{-1}y$ and therefore $y \sim x$. If $x \sim y$ and $y \sim z$ then $y = gx$ and $z = hy$ for some $g, h \in G$ and so $z = (hg)x \implies x \sim z$. □

Definition. *The set of equivalence classes are denoted by $G\backslash X$ if we are dealing with a left action or by X/G if we are dealing with a right action.*

3.2 Back to quadratic forms

Example 3.12. The forms $ax^2 + bxy + cy^2$ and $ax^2 - bxy + cy^2$ are always (improperly) equivalent via $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. However sometimes they are properly equivalent (e.g. $2x^2 \pm 2xy + 3y^2$) and sometimes they are not (e.g. $3x^2 \pm 2xy + 5y^2$).

Definition. *An integer m is represented by a integral bqf $f(x, y)$ if the equation*

$$f(x, y) = m$$

has an integer solution (x, y). If we can find an integer solution with x, y relatively prime, we say that m is properly represented by $f(x, y)$.

Example 3.13. A bqf $f(x, y) = ax^2 + bxy + cy^2$ properly represents both $a = f(\pm 1, 0)$ and $c = f(0, \pm 1)$.

The question we are trying to answer is which primes p can be (properly) represented by the (primitive) integral bqf $x^2 + ny^2$.

Lemma 3.14. *If m is an integer represented by the bqf $f(x, y)$, then m can be written as $m = d^2 m'$ where $m', d \in \mathbb{Z}$ and m' is properly represented by $f(x, y)$.*

Proof. Since $m = f(x, y)$ for some integers x, y with $d = (x, y)$, it follows that $m = d^2 f(x', y')$ where $x = dx', y = dy'$. But then $(x', y') = 1$ and the result follows by setting $m' = f(x', y')$. □

Lemma 3.15. *A bqf $f(x, y)$ properly represents an integer m if and only if $f(x, y)$ is properly equivalent to $mx^2 + bxy + cy^2$ for some $b, c \in \mathbb{Z}$.*

Proof. Assume m is properly represented by $f(x, y)$. The there exist relatively prime integers α, γ such that $f(\alpha, \gamma) = m$. Since $(\alpha, \gamma) = 1$, there exist integers β, δ such that

$$\alpha\delta - \beta\gamma = 1 \iff \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}).$$

Then

$$f(\alpha x + \beta y, \gamma x + \delta y) = f(\alpha, \gamma)x^2 + [f(\alpha, \delta) + f(\beta, \gamma)]\, xy + f(\beta, \delta)y^2,$$

which is of the desired form since $f(\alpha, \gamma) = m$.
Conversely, note that $g(x, y) = mx^2 + bxy + cy^2$ properly represents m because $m = g(1, 0)$. □

Proposition 3.16. *(i) Two equivalent bqf's represent the same integers.*

(ii) Two equivalent bqf's properly represent the same integers.

(iii) If a bqf $f(x, y)$ is equivalent to a primitive bqf, then $f(x, y)$ itself is primitive.

Proof. Exercise. □

Definition. *The discriminant of the bqf* $ax^2 + bxy + cy^2$ *is the integer* $D = b^2 - 4ac$.

Proposition 3.17. *Two equivalent forms have the same discriminant.*

Proof. Exercise. □

The discriminant D has a strong effect on the behavior of the bqf $f(x, y) = ax^2+bxy+cy^2$. We have

$$4af(x, y) = (2ax + by)^2 - Dy^2. \tag{3.2}$$

Thus, if $D < 0$, then $4af(x, y) \geq 0$, so the form represents either only nonnegative integers if $a > 0$ or only nonpositive integers if $a < 0$. (Note that we cannot have $a = 0$, since that would make $D = b^2$ which cannot be negative.)

On the other hand, if $D > 0$ then

$$f(b, -2a) = -aD$$

and

$$f(1, 0) = a$$

have opposite signs whenever $a \neq 0$. When $a = 0$, $D = b^2 > 0$ so $b \neq 0$ and $f(x, y) = bxy + cy^2 = y(bx + cy)$. Then $f(-c+1, b) = b(-bc+b+bc) = b^2 = D > 0$ and $f(-c-1, b) = -b^2 = -D < 0$. Therefore $f(x, y)$ represents both positive and negative integers.

Definition. *A bqf* $f(x, y) = ax^2 + bxy + cy^2$ *is called*

- *positive definite if* $D < 0, a > 0$. *It cannot represent negative integers.*
- *negative definite if* $D < 0, a < 0$. *It cannot represent positive integers.*
- *indefinite if* $D > 0$. *It represents both positive and negative integers* ($D > 0$).

Note: The above notions are invariant under equivalence.

Examples:

- $x^2 + ny^2$ is positive definite when $n > 0$.
- $x^2 + 3xy + y^2$ is indefinite
- $-x^2 + 3xy - 13y^2$ is negative definite.

The discriminant D influences the form in one other way. Since $D = b^2 - 4ac$ it follows that

$$D \equiv b^2 \pmod 4 \equiv 0, 1 \pmod 4.$$

Proposition 3.18. *Let $D \equiv 0, 1 \pmod 4$ be an integer and m be an odd integer such that $(m, D) = 1$. Then m is properly represented by a primitive bqf of discriminant D if and only if D is a quadratic residue modulo m.*

Proof. First assume m is properly represented by a primitive form $g(x, y)$. By Lemma 3.15, $g(x, y)$ is properly equivalent to a form $f(x, y) = mx^2 + bxy + cy^2$ where $b, c \in \mathbb{Z}$. By Proposition 3.16, $f(x, y)$ is also primitive and by Proposition 3.17, m is properly represented by $f(x, y)$. The discriminant of $f(x, y)$ is $D = b^2 - 4mc \equiv b^2 \pmod m$, so D is a quadratic residue modulo m.
Conversely, suppose $D \equiv b^2 \pmod m$. Since m is odd, we assume that D and b have the same parity (replace b by $b+m$ if necessary). Since $D \equiv 0, 1 \pmod 4$ it follows that $4 \mid D-b^2$ and thus

$$D \equiv b^2 \pmod{4m}.$$

Hence there exists $c \in \mathbb{Z}$ such that $D = b^2 - 4mc$.
Therefore $f(x, y) = m^2 + bxy + cy^2$ properly represents m (by Lemma 3.15) and has discriminant D. Since $(m, D) = 1$ it follows that $f(x, y)$ is primitive. □

Corollary 3.19. *Let $n \in \mathbb{Z}$ and p be an odd prime that does not divide n. Then $\left(\frac{-n}{p}\right) = 1$ if and only if p is represented by a primitive form of discriminant $-4n$.*

Proof. This follows immediately from Proposition 3.18, since p is prime and therefore $-4n$ is a quadratic residue modulo p if and only if $1 = \left(\frac{-4n}{p}\right) = \left(\frac{-n}{p}\right)$. □

This corollary gives an inkling into how to represent the primes that divide numbers of the form $a^2 + nb^2, (a, b) = 1$. Namely, we have seen that these primes are the ones for which $\left(\frac{-n}{p}\right) = 1$. Corollary 3.19 tells us that such primes are represented by some bqf of discriminant $-4n$.

The problem is that there are too many bqf of discriminant $-4n$. For instance, all the forms that appear in Euler's Conjecture 3.2 have discriminant -56. Or, apply the proof of Proposition 3.18 to $n = 3$ (so $D = -12$) and $m = 13$. Since $\left(\frac{-3}{13}\right) = 1$, Proposition 3.18 implies that 13 is represented by some bqf of discriminant -12. Going through the proof, we have to find b even such that

$$D \equiv b^2 \pmod{4m} \iff -12 \equiv b^2 \pmod{52}.$$

Going through $-12 \pmod{52} = \{\ldots, -12, 40, 92, 144 = 12^2, \ldots\}$ we see that we can take $b = 12$. Next, we need to find c such that

$$D = b^2 - 4mc \iff -12 = 144 - 52c \iff c = 3.$$

Thus 13 is represented by the bqf $f(x, y) = 13x^2 + 12xy + 3y^2$ (which has indeed discriminant -12). This is not exactly enlightening. What we need is a way to produce simpler bqf's that represent a given integer.

From now on we restrict our attention to primitive, positive definite binary quadratic forms. Happily enough, the forms $x^2 + ny^2 (n > 0)$ that we care about are indeed primitive and positive definite.

Definition. *A primitive positive definite bqf $ax^2 + bxy + cy^2$ is* reduced *if*

$$0 \leq |b| \leq a \leq c \quad \text{and} \quad b \geq 0 \text{ if either } |b| = a \text{ or } a = c. \tag{3.3}$$

Note: The integers a, c must be positive since the form is positive definite.

Examples:

- If $n > 0$, then $x^2 + ny^2$ is reduced.
- $2x^2 + 7y^2$ is reduced.
- $13x^2 + 12xy + 3y^2$ is primitive and positive definite, but not reduced.

Theorem 3.20. *Any primitive positive definite bqf is properly equivalent to a* unique *reduced form.*

Proof. Our proof has three steps.
Step 1 We show that a given primitive, positive definite bqf $f(x, y)$ is equivalent to a primitive positive definite bqf $f(x, y) = ax^2 + bxy + cy^2$ with $0 \leq |b| \leq a \leq c$.

Among all the forms properly equivalent to $g(x, y)$ – which we already know that have to be primitive and positive definite – choose the one with the minimal coefficient of xy. That is, choose $f'(x, y) = a'x^2 + b'xy + c'y^2$ such that $|b'|$ is minimal. Assume by contradiction that $a' < |b'|$. Then, for any integer m,

$$g'(x, y) = g'(x + my, y) = a'x^2 + (2a'm + b')xy + (c' + a'm^2)y^2$$

is properly equivalent to our $g(x, y)$. Since $a' < |b'|$ we can choose $m \in \mathbb{Z}$ (think quotient of division of $|b'|$ by $2a$) such that $0 \leq |2a'm + b'| < |b'|$. This contradicts the minimality of $|b'|$, so $|b'| \leq a'$. Similarly, we get $|b| \leq c'$. If $a' \leq c'$, choose $f(x, y) = f'(x, y)$ (and $b = b'$, $|b| \leq a = a' \leq c' = c$). If $a' > c'$, take $f(x, y) = f'(-y, x) = a'y^2 - bxy + c'x^2$ and $b = -b'$, $|b| = |b'| < a = c' < a' = c$. (i.e. interchange a' and c' and change the sign of b'). Note that $(x, y) \mapsto (-y, x)$ induces a proper equivalence since

$$\det \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = 1.$$

Step 2 We show that a primitive positive definite bqf $f(x, y) = ax^2 + bxy + cy^2$ with $0 \leq |b| \leq a \leq c$ is properly equivalent to a reduced one.

The form $f(x, y)$ is already reduced unless $b < 0$ and $-b = a$ or $a = c$. But then $f'(x, y) = ax^2 - bxy + cy^2$ is reduced and all we have to show is that $f(x, y)$ and $f'(x, y)$ are properly equivalent.

If $a = -b :\ A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$ and

$$f(A\vec{x}) = f(x+y,y) = a(x+y)^2 - a(x+y)y + cy^2 = ax^2 + axy + cy^2 = f'(x,y).$$

If $a = c :\ A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$ and

$$f(B\vec{x}) = f(-y,x) = ay^2 - bxy + ax^2 = f'(x,y).$$

Step 3 We show that two reduced forms cannot be properly equivalent.
Let $f(x,y) = ax^2 + bxy + cy^2$ with $|b| \leq a \leq c$. Since $f(x,y)$ is positive definite, for any integers x, y we have

$$f(x,y) \geq (a - |b| + c)\min(x^2,y^2) \text{ (exercise!)}$$

Therefore

$$f(x,y) \geq a - |b| + c \geq a \text{ whenever } xy \neq 0. \tag{3.4}$$

On the other hand $f(x,0) = ax^2$ and $f(0,y) = cy^2$. As we have seen in Example 3.13, a is properly represented by $f(x,y)$ and (3.4) implies that a is the smallest nonzero value of $f(x,y)$. Moreover, if $c > a$ then c is the next smallest positive value of $f(x,y)$. *Therefore the coefficients of x^2 and y^2 of a reduced form are the smallest positive integers properly represented by any equivalent form.* (This observation is due to Legendre.) For simplicity, assume $f(x,y) = ax^2 + bxy + cy^2$ is a reduced from with $|b| < a < c$. (The other cases are left as exercise.) From what we discussed above, it follows that $a < c < a - |b| + c$ are the smallest numbers properly represented by $f(x,y)$.

$$\begin{aligned} \textbf{Claim}\quad & f(x,y) = a,\ (x,y) = 1 \iff x = \pm 1, y = 0 \\ & f(x,y) = c,\ (x,y) = 1 \iff x = 0, y = \pm 1. \end{aligned} \tag{3.5}$$

Assume that $g(x,y) = a'x^2 + b'xy + c'y^2$ is a reduced form equivalent to $f(x,y)$. Since $f(x,y)$ and $g(x,y)$ represent the same numbers and are reduced, they must have the same coefficient of x^2 by Legendre's observation. So $a = a'$. On the other hand, $c' \geq a$. Assume that $c' = a$. Then, by (3.5), the equation $g(x,y) = a$ has 4 proper solutions $\pm(1,0), \pm(0,1)$. But the equation $f(x,y) = a$ has only 2 proper solutions (contradiction). Hence $c' > a$ and by applying again Legendre's observation, it follows that $c = c'$. Since the two bqf's have the same discriminant, it follows that $|b'| = |b|$. Thus

$$g(x,y) = ax^2 \pm bxy + cy^2.$$

It remains to show that $f(x,y) = g(x,y)$ when we make the stronger assumption that the two bqf's are *properly equivalent.* That is, we now assume that we have

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \quad g(x,y) = f(\alpha x + \beta y, \gamma x + \delta y).$$

Then

$$a = g(1,0) = f(\alpha,\gamma) \quad c = g(0,1) = f(\beta,\delta).$$

Since $\det A = 1$, we have $\alpha\delta - \beta\gamma = 1$, so $(\alpha,\gamma) = 1$ and $(\beta,\delta) = 1$. By (3.5), it follows that $(\alpha,\gamma) = \pm(1,0)$ and $(\beta,\delta) = \pm(0,1)$. Since $\det A = 1$, it follows that

$$A = \pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and therefore $f(x,y) = g(x,y)$.

□

Note: Now we can justify the examples we gave of properly equivalent and improperly equivalent forms. Namely, $3x^2 \pm 2xy + 5y^2$ (which we know are equivalent) are both reduced, and therefore they cannot be properly equivalent. Thus they are improperly equivalent. On the other hand, $2x^2 \pm 2xy + 3y^2$ are equivalent, but only $2x^2 + 2xy + 3y^2$ is reduced. By the proof of Theorem 3.20, the two forms properly equivalent.

From now on, all bqf's will be primitive, positive definite and equivalence will be proper.

Definition. *The class number $h(D)$ is the number of classes of primitive positive definite forms of discriminant $D < 0$.*

Note: By Theorem 3.20, $h(D)$ is equal to the number of reduced forms of discriminant D. A priori this number has no reason to be finite. However, suppose $ax^2 + bxy + cy^2$ to be a reduced form of discriminant $D < 0$. Since $|b| \le a$ we have $b^2 \le a^2$. Combining this with $a \le c$, we get

$$-D = 4ac - b^2 \ge 4a^2 - b^2 \ge 4a^2 - a^2 = 3a^2 \implies 0 \le a \le \sqrt{\frac{-D}{3}}.$$

If D is fixed, then the above relation and the fact $|b| \le a$ imply that there are only finitely many choices for a and b. Moreover, each such choice fixes c since $D = b^2 - 4ac$.
Thus there are only finitely many reduced forms of discriminant D and we have proved the following result.

Theorem 3.21. *Let $D \in \mathbb{Z}_{<0}$. The class number $h(D)$ is finite and is equal to the number of reduced forms of discriminant D.*

Here are a couple of examples computed using the algorithm described above. We will need to use some of them later on, and I might explain them in class. But it would be a good idea to work as many of them as you can on your own.

D	$h(D)$	reduced forms of discriminant D
-4	1	x^2+y^2
-8	1	x^2+2y^2
-12	1	x^2+3y^2
-20	2	$x^2+5y^2, 2x^2+2xy+3y^2$
-28	1	x^2+7y^2
-56	4	$x^2+14y^2, 2x^2+7y^2, 3x^2\pm 2xy+5y^2$
-108	3	$x^2+27y^2, 4x^2\pm 2xy+7y^2$
-256	4	$x^2+64y^2, 4x^2+4xy+17y^2, 5x^2\pm 2xy+13y^2$

We can now go back to the Descent Step in Euler's strategy.

Theorem 3.22. *Let $n \in \mathbb{Z}_{>0}$ and p be an odd prime such that $p \nmid n$. Then $\left(\dfrac{-n}{p}\right) = 1$ if and only if p is represented by one of the $h(-4n)$ reduced forms of discriminant $-4n$.*

Proof. This is an immediate consequence of Corollary 3.19 and Theorem 3.20. □

This result completely settles the Descent Step. We just need to put it together with the Reciprocity Step, and see what we get. But rather than looking at the case of bqf's of discriminant $-4n$, we will state a result that applies to all discriminants $D < 0$.

Theorem 3.23. *Let D be a negative integer such that $D \equiv 0, 1 \pmod 4$. Let $\chi = \chi_D : (\mathbb{Z}/D\mathbb{Z})^{\times} \to \{\pm 1\}$ be the group homomorphism defined in Theorem 2.6 and p be an odd prime, $p \nmid D$. Then $p \pmod D \in \ker \chi$ if and only if p is represented by one of the $h(D)$ reduced forms of discriminant D.*

Proof. We have seen that

$$p \pmod D \in \ker \chi \iff \left(\frac{D}{p}\right) = 1.$$

We also know that

$$\left(\frac{D}{p}\right) = 1 \iff D \text{ is a quadratic residue modulo } p.$$

By Proposition 3.18 this is equivalent to the fact that p is represented by a primitive positive definite form of discriminant D. The result now follows from Theorem 3.20. □

This theorem tells us that there is a congruence condition $p \equiv \alpha, \beta, \ldots \pmod D$ which gives necessary and sufficient conditions for an odd prime $p \nmid D$ to be represented by a form of discriminant D. Since we know how to find the reduced forms of a given discriminant and quadratic reciprocity makes it easy to find the congruence classes $\alpha, \beta, \ldots \pmod D$ such that $\left(\frac{D}{p}\right) = 1$, we now have a complete and effective form of Euler's strategy.

Example 3.24. $D = -4:$ the only reduced form is $x^2 + y^2$. On the other hand we know that

$$\left(\frac{-1}{p}\right) = 1 \iff p \equiv 1 \pmod 4.$$

Thus it follows immediately from Theorem 3.23 that $p \neq 2$ is of the form $x^2 + y^2$ if and only if $p \equiv 1 \pmod 4$.

In other words, now we get a two line proof of a fact that had taken a week to prove last quarter.

$D = -8:$ again we have only one reduced form of discriminant -8, namely $x^2 + 2y^2$. And we know that

$$\left(\frac{-2}{p}\right) = 1 \iff p \equiv 1, 3 \pmod 8.$$

Theorem 3.23 implies that $p \neq 2$ is of the form $x^2 + 2y^2$ if and only if $p \equiv 1, 3 \pmod 8$. I won't remind you how long that took to prove!

$D = -12:$ the only reduced form is $x^2 + 3y^2$ and it is easy to see find the congruence classes for p so that $\left(\frac{-3}{p}\right) = 1$.

We can go further than Fermat.

Proposition 3.25. *If p is a prime, then*

$$p = x^2 + 7y^2 \iff p = 7 \text{ or } p \equiv 1, 9, 11, 15, 23, 25 \pmod{28}.$$

Proof. Exercise. □

Each time we made use of the fact that there is only one reduced form of discriminant $-4n$, i.e. that $h(-4n) = 1$. Unfortunately, the list of $n > 0$ for which this happens is rather short.

Theorem 3.26 (Landau)**.** *Let n be a positive integer. Then*

$$h(-4n) = 1 \iff n = 1, 2, 3, 4, 7.$$

Proof. We will follow Landau. In a nutshell, the idea is that we already know a reduced from of discriminant $-4n$, namely $x^2 + ny^2$. So if we produce another one, that means that $h(-4n) > 1$. We already know that $h(-4) = 1$, so we can assume that $n > 1$. If n is *not* a prime power, it means that n has at least two distinct prime divisors p and q. Therefore

$$n = p^r q^s m, \text{ with } r, s \geq 1 \text{ and } (m, p) = (m, q) = 1.$$

Choose $a = \min(p^r, q^s)$ and $c = m \cdot \max(p^r, q^s)$.Then $n = ac$, $c > a > 1$ and $(a, c) = 1$. Therefore the form

$$ax^2 + cy^2$$

is reduced of discriminant $-4n$.

If $n = 2^r$ and $r \geq 4$, then

$$4x^2 + 4xy + (2^{r-2} + 1)y^2$$

is reduced (note that $4 \leq 2^{r-2} + 1$ since $r \geq 4$) of discriminant $-4n$.

If $n = 2^3$, then we follow the algorithm for finding reduced forms of discriminant $D = -4 \cdot 8 = -32$. We know that

$$0 < a \leq \sqrt{\frac{32}{3}} \implies 1 \leq a \leq 3.$$

If $a = 3$, then $|b| \leq 3$. But if $b = \pm 3$, this means that we have to find $c \in \mathbb{Z}$ such that $-32 = 9 - 12c$ which is impossible. For $b = \pm 2$ we get $-32 = 4 - 12c$, so $c = 3$. But then only $3x^2 + 2xy + 3y^2$ is reduced. However this is enough for our purposes, because it shows that $h(-32) > 1$. (In fact, $h(-32) = 2$, which is left as an exercise.)

Of the powers of 2 this leaves us with $n = 2, 4$. We have already seen what happens for $n = 2$. The case $n = 4$ is left as an exercise.

If $n = p^r$ with p and odd prime, then $n + 1$ is even. So if $n + 1$ is a *not* a power of 2, then $n + 1 = ac$ with $1 < a < c$ and $(a, c) = 1$. It follows that

$$ax^2 + 2xy + cy^2$$

is reduced of discriminant $-4n$. If $n + 1 = 2^s$ and $s \geq 6$, then

$$8x^2 + 6xy + (2^{s-3} + 1)y^2$$

is reduced (indeed, $8 < 2^{s-3} + 1$ in this case) of discriminant $-4n$.
If $s = 5$, then $n + 1 = 32$, so $n = 31$ which is an odd prime. We go through our algorithm again to find reduced forms with discriminant $-4n = -124$. We have

$$0 < a \leq \sqrt{\frac{124}{3}} \implies 1 \leq a \leq 5.$$

Now we need to find an integer solution to the equation $-124 = b^2 - 4ac$ with $|b| \leq a \leq c$. First note the b has to be even. If $a = 5$ and $b = \pm 4$ the equation becomes $-124 = 16 - 20c$, so $c = 7$. The forms

$$5x^2 \pm 4xy + 7y^2$$

are both reduced of the given discriminant, so $h(-4n) \geq 3$. (In fact we have equality, a fact that I leave for you to prove!).
For $s = 4$ we get $n + 1 = 16 \implies n = 15$ which is not a prime power.
For $s = 3$ we get $n+1 = 8 \implies n = 7$ and we have seen in Proposition 3.25 that $h(-28) = 1$.
For $s = 2$ we get $n + 1 = 4 \implies n = 3$ and we have seen in Example 3.24 that $h(-12) = 1$.
For $s = 1$ we get back to $n = 1$. □

Note: The case $n = 4$ is included in the $p = x^2 + y^2$ case since one of the x, y has to be even and the other odd in order for p to be odd.

3.3 Elementary genus theory

Landau's Theorem 3.26 makes it clear that we need some new ideas for dealing with the case $h(-4n) > 1$. Let us consider the following example.

Example 3.27. Take the case $n = 5$. First let us determine the reduced form of discriminant $D = -20$. We have seen that they need to satisfy

$$0 \le |b| \le a \le \sqrt{\frac{20}{3}} \implies 0 \le |b| \le a \le 2,$$

and $-20 = b^2 - 4ac$ so b is even.

- $a = 2$: then $-20 = b^2 - 8c$.
  If $b = 2$, then $c = 3$ and we get the reduced form $2x^2 + 2xy + 3y^2$.
  If $b = 0$, the diophantine equation has no solution $c \in \mathbb{Z}$.
- $a = 1$: then $b = 0$ and $20 = -4c$, so $c = 5$. This yields the familiar $x^2 + 5y^2$.

Therefore $h(-20) = 2$ and the two reduced form are

$$2x^2 + 2xy + 3y^2 \quad \text{and} \quad x^2 + 5y^2.$$

Here Theorem 3.23 and quadratic reciprocity tell us that, if $p \neq 5$ is an odd prime

$$p \equiv 1, 3, 7, 9 \pmod{20} \iff \left(\frac{-5}{p}\right) = 1 \iff p = x^2 + 5y^2 \text{ or } p = 2x^2 + 2xy + 3y^2.$$

We can see from this example that what we need is a method to separate reduced forms of the same discriminant. The basic idea is due to Lagrange: consider the congruence classes in $(\mathbb{Z}/D\mathbb{Z})^\times$ represented by a single form and group together the forms that represent the same congruence classes. This is precisely the basic idea of genus theory.

To clarify what we mean, we look again at the case $D = -20$. We will plug in for x, y all the values in $\mathbb{Z}/D\mathbb{Z}$ and for each pair compute the value of the two reduced forms of genus D. Then we throw out the pairs that give values that are not in $(\mathbb{Z}/D\mathbb{Z})^\times$. To shorten our computation, note that if both x, y are divisible by 2 or 5, then so will both $x^2 + 5y^2$ and $2x^2 + 2xy + 3y^2$. So at least one of them has to be relatively prime to 20.

$x \pmod{20}$	$y \pmod{20}$	x^2+5y^2		$2x^2+2xy+3y^2$	
0	± 1	5	(throw this out)	3	(keep this one)
0	± 3	5	(throw this out)	7	(keep this one)
0	± 7	5	(throw this out)	7	(keep this one)
0	± 9	5	(throw this out)	3	(keep this one)
± 1	0	1	(keep this one)	2	(throw this out)
± 3	0	9	(keep this one)	18	(throw this out)
± 7	0	9	(keep this one)	18	(throw this out)
± 9	0	1	(keep this one)	2	(throw this out)
1	1	6	(throw this out)	7	(keep this one)
1	2	1	(keep this one)	18	(throw this out)
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

Continuing this table one sees that

$$\begin{aligned} &x^2+5y^2 && \text{represents } 1,9 && \text{in } (\mathbb{Z}/20\mathbb{Z})^\times; \\ \text{and} \\ &2x^2+2xy+3y^2 && \text{represents } 3,7 && \text{in } (\mathbb{Z}/20\mathbb{Z})^\times. \end{aligned} \tag{3.6}$$

Repeating the same procedure for $D=-56$, we get that

$$\begin{aligned} &x^2+14y^2,\ 2x^2+7y^2 && \text{represent } 1,9,15,23,25,39 && \text{in } (\mathbb{Z}/56\mathbb{Z})^\times; \\ \text{and} \\ &3x^2 \pm 2xy+5y^2 && \text{represent } 3,5,13,19,27,45 && \text{in } (\mathbb{Z}/56\mathbb{Z})^\times. \end{aligned} \tag{3.7}$$

Definition. *We say that two primitive positive definite bqf's of discriminant D have the same* genus *if they represent the same congruence classes in* $(\mathbb{Z}/D\mathbb{Z})^\times$.

Note: Since equivalent forms represent the same integers, they are in the same genus. In particular, each genus consists of a finite number of proper classes of forms.

In (3.6), we have seen that for $D=-20$ there are 2 genera, each consisting of a single class. Combining the same (3.6) with Theorem 3.23 we obtain that for an odd prime $p \neq 5$,

$$\begin{aligned} p = x^2+5y^2 &\iff && p \equiv 1,9 \pmod{20} \\ p = 2x^2+2xy+3y^2 &\iff && p \equiv 3,7 \pmod{20} \end{aligned} \tag{3.8}$$

So we now have a proof for the first part of Euler's Conjecture 3.1.

On the other hand, (3.7) shows that for $D=-56$ there are also 2 genera, but now each genus consists of 2 classes. Combining it with Theorem 3.23 we obtain that for an odd prime

$p \neq 7$,

$$\begin{aligned} p = x^2 + 14y^2 \text{ or } p = 2x^2 + 7y^2 &\iff \quad p \equiv 1, 9, 15, 23, 25, 39 \pmod{56} \\ p = 3x^2 \pm 2xy + 5y^2 &\iff \quad p \equiv 3, 5, 13, 19, 27, 45 \pmod{56} \end{aligned} \tag{3.9}$$

This proves first part of Euler's Conjecture 3.2.

In both these cases, what made the whole thing work was the fact the the two genera represent disjoint sets of values in $(\mathbb{Z}/D\mathbb{Z})^\times$. We must show that this phenomenon holds in general. To that end, we start with a result of Gauss.

Lemma 3.28. *Given a form $f(x, y)$ and an integer $M \neq 0$, the bqf $f(x, y)$ properly represents numbers relatively primes to M.*

Proof. Let $f(x, y) = ax^2 + bxy + cy^2$. We know that $(a, b, c) = 1$ – since all our forms are primitive – so no prime can divide all of them. Let p be an arbitrary prime. There are three possibilities.

- $p \mid a$ and $p \mid c$: then $p \nmid b$. Therefore if $p \nmid x$ and $p \nmid y$, then $p \nmid f(x, y)$.
- $p \nmid a$: choose x, y such that $p \nmid x$ and $p \mid y$. Then $p \nmid f(x, y)$.
- $p \nmid c$: choose x, y such that $p \mid x$ and $p \nmid y$. Then $p \nmid f(x, y)$.

If $M = \pm 1$, the result is obvious. If not, then $M = \pm p_1^{a_1} \dots p_r^{a_r}$ with $p_1, \dots, p_r$ distinct primes. By the Chinese Remainder Theorem we can choose x, y subject to the above conditions for each of the $p_i, 1 \leq i \leq r$. Then $p_i \nmid f(x, y)$ for all i and therefore $m = f(x, y)$ is relatively prime to M. The result follows since, by Lemma 3.14, $m = d^2 m'$ for some m' that is properly represented by $f(x, y)$. □

Definition. *For a negative integer $D \equiv 0, 1 \pmod 4$ the* principal form *of discriminant D is*

$$x^2 - \frac{D}{4} y^2 \qquad \text{if } D \equiv 0 \pmod 4$$

$$x^2 + xy + \frac{1 - D}{4} y^2 \qquad \text{if } D \equiv 1 \pmod 4.$$

Note: These forms have indeed discriminant D and they are reduced.

Proposition 3.29. *Given a negative integer $D \equiv 0, 1 \pmod 4$, denote by $\chi = \chi_D : (\mathbb{Z}/D\mathbb{Z})^\times \to \{\pm 1\}$ the group homomorphism defined in Theorem 2.6. Let $f(x, y)$ be a bqf of discriminant D.*

(i) The values in $(\mathbb{Z}/D\mathbb{Z})^\times$ represented by the principal form of discriminant D form a subgroup $H \subset \ker \chi \subset (\mathbb{Z}/D\mathbb{Z})^\times$.

(ii) The values in $(\mathbb{Z}/D\mathbb{Z})^\times$ *represented by* $f(x,y)$ *form a coset of* H *in* $\ker\chi$.

Note: Since cosets are disjoint, this says that different genera represent disjoint sets of values in $(\mathbb{Z}/D\mathbb{Z})^\times$. It also means that each genus corresponds to an element of the quotient group $(\ker\chi)/H$.

Proof. We start by proving that if a number $(m,D)=1$ is represented by a form $g(x,y)$ of discriminant D, then $[m]\in\ker\chi$. By Lemma 3.14, $m=d^2m'$ where m' is properly represented by $g(x,y)$. Then

$$\chi([m])=\chi([d^2m'])=\chi([d])^2\chi([m'])=\chi([m']).$$

On the other hand, Proposition 3.18 implies that D is a quadratic residue modulo m', so there exist integers $b,c\in\mathbb{Z}$ such that $D=b^2-cm'$. Note that $(b,m')=1$. If m' is odd, then

$$\chi([m])=\chi([m'])=\left(\frac{D}{m'}\right)=\left(\frac{b^2-cm'}{m'}\right)=\left(\frac{b^2}{m'}\right)=\left(\frac{b}{m'}\right)^2=1.$$

If m' is even, then D must be odd and Lemma 3.15 implies that m' is (properly) represented by a form $m'x^2+b'xy+c'y^2$ of discriminant D. Hence

$$D=(b')^2-4m'c'\equiv(b')^2 \pmod 8.$$

But b' has to be odd, and the only odd square modulo 8 is 1. Hence $D\equiv 1 \pmod 8$, and this implies that $\chi([2])=1$. Therefore, if we write $m'=2^am''$ we have

$$\chi([m])=\chi([m'])=\chi([2])^a\chi([m''])=\chi([m''])=1 \text{ (as before)}.$$

Now that our claim is proved, let us go back to the first statement of the Proposition. By definition

$$H=\{[m];\ m \text{ is represented by the principal form of discriminant } D\}.$$

The above claim shows that the set H is a subset of $\ker\chi$. We have to show that H contains the identity (and this is trivial since the principal form evaluated at $(1,0)$ yields precisely 1) and is closed under multiplication.

- When $D=-4n$, the principal form is x^2+ny^2. But we know that

  $$(x^2+ny^2)(u^2+nv^2)=(xu+nyv)^2+n(xv-yu)^2.$$

  Therefore the product of two representable integers is also representable.

- When $D=1-4n$, the principal form is $x^2+xy+ny^2$. We have

  $$4(x^2+xy+ny^2)\equiv 4x^2+4xy+y^2 \pmod D \equiv (2x+y)^2 \pmod D. \qquad (3.10)$$

  Let $H'=\{[m]^2;[m]\in(\mathbb{Z}/D\mathbb{Z})^\times\}$ the subgroup of squares in $(\mathbb{Z}/D\mathbb{Z})^\times$. Then (3.10) shows that $H=H'$ and therefore H is closed under multiplication.

For the second statement of the Proposition, we again treat the two cases separately. If $D = -4n$, then taking $M = 4n$ in Lemma 3.28 we obtain that $f(x, y)$ properly represents some integer a relatively prime to D. By Lemma 3.15, we get that $f(x, y)$ is properly equivalent to a bqf of the form $ax^2 + b'xy + cy^2$ of discriminant D. Since representability is stable under equivalence of forms, we can assume that $f(x, y) = ax^2 + b'xy + cy^2$.
But $-4n = D = (b')^2 - 4ac$, so b' is even. Therefore

$$f(x, y) = ax^2 + 2bxy + cy^2 \text{ and } n = ac - b^2.$$

Therefore

$$af(x, y) = (ax + by)^2 + ny^2.$$

Since $(a, 4n) = 1$ it follows that the values of $f(x, y)$ in $(\mathbb{Z}/4n\mathbb{Z})^\times$ lie in the coset $[a]^{-1}H$. Conversely, if $[m] \in [a]^{-1}H$, then $[ac] \in H$, so there exist integers u, v such that

$$am \equiv u^2 + nv^2 \pmod{4n}.$$

Choose $x, y \in \mathbb{Z}$ such that

$$\begin{cases} ax + by & \equiv u \pmod{4n} \\ y & \equiv v \pmod{4n}. \end{cases}$$

Note that we can do this since $(a, 4n) = 1$. Then

$$af(x, y) = (ax + by)^2 + ny^2 \equiv u^2 + nv^2 \pmod{D} \equiv am \pmod{D}.$$

Again we use the fact that $(a, D) = 1$ to obtain

$$f(x, y) \equiv m \pmod{D} \implies [m] \text{ is represented by } f(x, y).$$

The case $D \equiv 1 \pmod 4$ is similar (exercise!) and the result is proved. □

Definition. *With the notation from Proposition 3.29, let H' be a coset of H in* $\ker\chi$. *The* genus *of the coset H' consists of all the forms of discriminant D that represent the values of H' modulo D.*
The genus containing the principal form is called the principal genus.

We have proved the following result.

Theorem 3.30. *Let $D < 0$ be an integer such that $D \equiv 0, 1 \pmod 4$ and $p \nmid D$ be an odd prime. With the notation from Proposition 3.29, let H' be a coset of H in $\ker\chi$. Then $[p] = p \pmod{D} \in H'$ if and only if p is represented by a reduced form of discriminant D in the genus of H'.*

This is the main result of our elementary genus theory. It generalizes (3.8) and (3.9), and it shows that there are always congruence conditions which characterize the primes that can be represented by some bqf in a given genus.

For us, the most interesting situation regards the principal genus, since for $D = -4n$ the principal form is $x^2 + ny^2$.

Corollary 3.31. *Let $n \in \mathbb{Z}> 0$ and $p \nmid n$ and odd prime. Then p is represented by a form of discriminant $-4n$ in the principal genus if and only if there exist an integer β such that*

$$p \equiv \beta^2 \text{ or } \beta^2 + n \pmod{4n}.$$

Proof. If y is even, then $x^2 + ny^2 \equiv x^2 \pmod{4n}$.
On the other hand, if y is odd, then $x^2 + ny^2 \equiv x^2 + n \pmod{4n}$.

□

4 Eulers' conjecture for $n = 5$ and $n = 14$

4.1 Euler's conjecture for $n = 5$

Euler's Conjecture 3.1 states that if p is a prime, $p \neq 2, 5$, then

$$\begin{aligned} p = x^2 + 5y^2 &\iff p \equiv 1, 9 \pmod{20} \\ 2p = x^2 + 5y^2 &\iff p \equiv 3, 7 \pmod{20} \end{aligned}$$

We have proved that for $p \neq 2, 5$ we have

$$\begin{aligned} p = x^2 + 5y^2 &\iff p \equiv 1, 9 \pmod{20} \\ p = 2x^2 + 2xy + 3y^2 &\iff p \equiv 3, 7 \pmod{20} \end{aligned}$$

To go from the second to the first, we proceed as follows.

$$p \equiv 3, 7 \pmod{20} \iff p = 2x^2 + 2xy + 3y^2 \iff 2p = 4x^2 + 4xy + 6y^2 = (2x + y)^2 + 5y^2.$$

In fact, something more general is true. Fermat notices that if p and q are both of the form $2x^2 + 2xy + 3y^2$, then pq is of the form $x^2 + 5y^2$. It comes down to the fact that

$$(2x^2 + 2xy + 3y^2)(2u^2 + 2uv + 3v^2) = (2xu + xy + yu + 3yv)^2 + 5(xv - yu)^2. \tag{4.1}$$

4.2 Euler's conjecture for $n = 14$

Euler's conjecture conjn14 states that if $p \neq 7$ is an odd prime, then

$$\begin{aligned} p = \begin{cases} x^2 + 14y^2 \\ 2x^2 + 7y^2 \end{cases} &\iff p \equiv 1, 9, 15, 23, 25, 39 \pmod{56} \\ 3p = x^2 + 14y^2 &\iff p \equiv 3, 5, 13, 19, 27, 45 \pmod{56} \end{aligned}$$

We have proved that for $p \neq 2, 7$ we have

$$p = \begin{cases} x^2 + 14y^2 \\ 2x^2 + 7y^2 \end{cases} \iff p \equiv 1, 9, 15, 23, 25, 39 \pmod{56}$$

$$p = 3x^2 + \pm 2xy + 5y^2 \iff p \equiv 3, 5, 13, 19, 27, 45 \pmod{56}$$

As 3 is represented by $3x^2 + \pm 2xy + 5y^2$, the second part of Euler's conjecture follows from the fact that if $a^2 + 2bxy + cy^2$ is a form of discriminant $D = -4n$, then

$$(ax^2 + 2bxy + cy^2)(au^2 + 2buv + cv^2) = (axu + bxy + byu + cyv)^2 + n(xv - yu)^2. \qquad (4.2)$$

4.3 Further remarks

We now see that the coefficient 2 of $2p$ that appears in Conjecture 3.1 could be replaced by any integer represented by the form $2x^2 + 2xy + 3y^2$. Similarly, the coefficient 3 from second part of Conjecture 3.2 could be replaced by any integer represented by the form $3x^2 \pm 2xy + 5y^2$. But Legendre's observation from the proof of Theorem 3.20 shows that 2 and 3 are in some sense the best possible coefficients, as they are the smallest integers represented by the bqf's in question.

The two identities (4.1) and (4.2) are special cases of Gauss's law for the composition of binary quadratic forms.

5 Algebraic numbers and algebraic integers

Definition. *A complex number $z \in \mathbb{C}$ is called* algebraic *(over $\mathbb{Q}$) if it is the root of some equation*

$$a_n z^n + \ldots + a_1 z + a_0 = 0$$

with coefficients $a_0, \ldots, a_n \in \mathbb{Q}, a_n \neq 0$. Otherwise z is called transcendental.

Note that there is no loss in generality if we suppose that $a_n = 1$.

Example 5.1. Any rational number is algebraic. If $r \in \mathbb{Q}$, then $\sqrt[n]{r}$ is algebraic. Any quadratic irrational is algebraic. But π and e are transcendental.

Definition. *For an algebraic number $\alpha \in \mathbb{C}$, we define its* minimal polynomial *over $\mathbb{Q}$ to be the monic polynomial $m_\alpha(X) \in \mathbb{Q}[X]$ that has α as a root. (Recall that monic means the leading coefficient is 1.)*

Remark 5.2. The minimal polynomial m_α is unique, irreducible, and it divides any polynomial $f(X) \in \mathbb{Q}[X]$ that has the property that $f(\alpha) = 0$.

Definition. *The* degree of an algebraic number $\alpha \in \mathbb{C}$ *is the degree of its minimal polynomial.*

Example 5.3.
- The degree of 5 is 1.
- In fact, the degree of any rational number is 1.
- The degree of $\sqrt{5}$ is 2.
- The degree of i is 2. ($m_i = X^2 + 1$)

Definition. *An algebraic number $\alpha \in \mathbb{C}$ is called an* algebraic integer *if it satisfies an equation*

$$z^n + \ldots + a_1 z + a_0 = 0$$

with coefficients $a_0, \ldots, a_{n-1} \in \mathbb{Z}$.

Note that this is the same as saying that m_α has integer coefficients.

Example 5.4.
- 5 is an algebraic integer.
- In fact, any integer is an algebraic integer.
- $\sqrt{5}$ is an algebraic integer.
- A rational number is an algebraic integer if and only if it is an integer to begin with.
- i is an algebraic integer.

If we have a field $F \subset \mathbb{C}$ that contains $\mathbb{Q}$ we can look at the set of algebraic integers in F.

Definition. *The set of algebraic integers in F*

$$\mathcal{O}_F = \{\alpha \in F; \alpha \textit{ is an algebraic integer}\}$$

is called the ring of integers of F or the integral closure of $\mathbb{Z}$ in F.

It is a nontrivial fact that $\mathcal{O}_F$ is a ring with respect to the $+$ and $\cdot$ of the field F.

Proposition 5.5. *Let d be a square-free integer and set $F = \mathbb{Q}(\sqrt{d})$ we have*

$$\mathcal{O}_F = \begin{cases} \mathbb{Z}[\sqrt{d}] & d \equiv 2, 3 \pmod 4 \\ \\ \mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right] & d \equiv 1 \pmod 4). \end{cases}$$

There is also a notion of discriminant of a number field, and in this case it is given by

$$D = \begin{cases} 4d & d \equiv 2, 3 \pmod 4; \\ d & d \equiv 1 \pmod 4. \end{cases}$$

Remark 5.6. Note that $\mathbb{Q}(\sqrt{d}) = \mathbb{Q}(\sqrt{D})$. We denote this field by F.

- If $d \equiv 2, 3 \pmod 4$ then $\mathcal{O}_F = \mathbb{Z}[\sqrt{d}]$, the discriminant of F is $D = 4d$ and the minimal polynomial of $\sqrt{d}$ is
$$t^2 - d = t^2 - \frac{D}{4}.$$
By thinking of $t = x/y$ and clearing denominators we recover the principal for of discriminant D, namely
$$x^2 - \frac{D}{4}y^2.$$

- If $d \equiv 1 \pmod 4$ then $\mathcal{O}_F = \mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$, the discriminant of F is $D = d$ and the minimal polynomial of $\frac{-1+\sqrt{d}}{2}$ is
$$t^2 + t + \frac{1-d}{4} = t^2 + t + \frac{1-D}{4}.$$
By thinking of $t = x/y$ and clearing denominators we recover the principal for of discriminant D, namely
$$x^2 + xy + \frac{1-D}{4}y^2.$$

Proof of Proposition 5.5. Note that if $\alpha = x + y\sqrt{d} \in \mathbb{Q}(\sqrt{d})$, then the minimal polynomial of α is

$$m_\alpha = \begin{cases} t - \alpha = t - x & \alpha \in \mathbb{Q} \iff y = 0; \\ t^2 - 2xt + (x^2 - dy^2) & \alpha \notin \mathbb{Q} \iff y \neq 0. \end{cases}$$

Hence if $\alpha \in \mathbb{Z}[\sqrt{d}]$, then $m_\alpha \in \mathbb{Z}$, so α is an algebraic integer. Hence $\mathbb{Z}[\sqrt{d}] \subset \mathcal{O}_F$. If $d \equiv 1 \pmod 4$, we can find more algebraic integers. Namely, every $\alpha \in \mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$ is of the form
$$\alpha = a + b\frac{-1+\sqrt{d}}{2}, \text{ with } a, b \in \mathbb{Z} \implies \alpha = \left(a - \frac{b}{2}\right) + \frac{b}{2}\sqrt{d}.$$
If $b = 0$, then $\alpha = a \in \mathbb{Z}$ is an algebraic integer. If $b \neq 0$, the minimal polynomial of α is
$$m_\alpha = t^2 - (2a-b)t + \left(a - \frac{b}{2}\right)^2 - \frac{b^2 d}{4} = t^2 - (2a-b)t + \left(a^2 - ab + b^2\frac{1-d}{4}\right) \in \mathbb{Z}[t]$$
and therefore α is an algebraic integer. Hence, $\mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right] \subset \mathcal{O}_F$ for $d \equiv 1 \pmod 4$.

Conversely, we want to show that $\mathcal{O}_F$ is contained in $\mathbb{Z}[\sqrt{d}]$ if $d \equiv 2, 3 \pmod 4$ or contained in $\mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$ if $d \equiv 1 \pmod 4$.

Let $\alpha \in \mathcal{O}_F$. Hence $\alpha = x + y\sqrt{d}$ for some $x, y \in \mathbb{Q}$ and $m_\alpha \in \mathbb{Z}[t]$.

If $y = 0$, then $m_\alpha = t - x$, so $x \in \mathbb{Z}$, and thus $\alpha \in \mathbb{Z}$. If $y \neq 0$, then $m_\alpha = t^2 - 2xt + (x^2 - dy^2) \in \mathbb{Z}[t]$ so $2x, x^2 - dy^2$ are both integers. Since $2x \in \mathbb{Z}$, we have $4x^2 \in \mathbb{Z}$, and this implies that $4dy^2 \in \mathbb{Z}$. Since d is square-free, it follows that y can have denominator at most 2 (since d cannot clear any part of the denominator of y^2.) and $2y \in \mathbb{Z}$.

Hence there exist integers a, b such that $x = a/2$ and $y = b/2$, which means that $\alpha = \frac{a+b\sqrt{d}}{2}$ and $m_\alpha = t^2 - at + \frac{a^2-b^2d}{4}$.

Since $m_\alpha \in \mathbb{Z}[t]$ we know that

$$a^2 - b^2 d \equiv 0 \pmod 4. \tag{5.1}$$

Since $a^2, b^2 \equiv 0, 1 (\text{mod } 4)$, for $d \equiv 2, 3 (\text{mod } 4)$ the only solution of equation (5.1) is $a^2, b^2 \equiv 0 \pmod 4$, so a, b are both even, and $\alpha \in \mathbb{Z}[\sqrt{d}]$. If $d \equiv 1 (\text{mod } 4)$, (5.1) implies that $a^2 \equiv b^2 \pmod 4$. It means that a, b must have the same parity, so

$$\alpha = \frac{a+b\sqrt{d}}{2} = \frac{a+b}{2} + b\frac{-1+\sqrt{d}}{2} \in \mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right].$$

□

6 The order of arithmetic functions

Given that prime numbers are the building blocks of all numbers, it is remarkable that there are so many simple problems we havent got answers to yet. Here are a few examples.

1. **The Goldbach conjecture:** Every even number greater than 2 is the sum of two prime numbers. For example, $6 = 3 + 3, 12 = 5 + 7, 100 = 47 + 53$. This was proposed in 1742 by Christian Goldbach, and has now been verified for numbers up to 4×10^18. A recent paper (May 2013) by Harald Helfgott has shown that every odd number greater than 5 is the sum of at most three primes (`http://www.truthiscool.com/prime-numbers-the-271-year-old-puzzle-resolved`).

2. **The twin prime conjecture:** There are infinitely pairs of primes with a difference of two. For example, some twin prime pairs are $(3, 5), (5, 7), (11, 13), (17, 19), (41, 43)$. The largest known twin primes have over 200,000 digits in them – but can we always find a bigger one? More generally, Polignac's conjecture states that for any even number n there are infinitely pairs of primes separated by n. This has not been proven or disproven for any value of n. However, a paper by Yitang Zhang in April 2013 was the first to show that the conjecture is true for some value of n less than 70 million. Later the same year, James Maynard proved that there are infinitely many primes separated by at most 600. Under certain other conjectures, one can prove that the smaller "gap" between consecutive primes is at most 6.

3. **The Fermat prime conjecture:** There are infinitely many Fermat primes. A Fermat prime is a prime of the form $2^{2^n}+1$; for example $5 = 2^{2^1}+1$. Fermat conjectured that all numbers of this form were prime, but in 1735 Euler showed that $2^{2^5}+1$ was a composite number. Nobody has found a Fermat prime other than up to $n = 4$, but it is unknown whether there might be infinitely many.

4. **The Mersenne prime conjecture:** There are infinitely many Mersenne primes. A Mersenne prime is a prime of the form $2^n - 1$; for example $7 = 2^3 - 1$. A Mersenne number can only be prime if n itself is prime. Forty-eight such Mersenne primes are known (as of April 2015), including the largest prime ever discovered (in January 2013), but it is unknown whether there are infinitely many such numbers.

5. **Distributions of primes:** Prime numbers are distributed randomly, i.e. they have the same distributions as any set of randomly generated integers.

6. **The Riemann Hypothesis:** It is known that the distribution of prime numbers are is governed by the distribution of the (non-trivial) zeros of a certain complex function called the Riemann zeta function. The Riemann Hypothesis (formulated by Riemann himself) states that every (non-trivial) zero of the Riemann zeta function has real part equal to $1/2$. This is one of the six remaining Clay Institute Millennium Prize Problems worth \$1 million. Amazingly enough, this problem has implications in other sciences, e.g. statistical mechanics. There are whole papers written about the influence of the Riemann zeta function on physics, e.g. `http://arxiv.org/abs/1101.3116`.

In order to answer questions about primes, one needs some way to figure out how fast certain arithmetic functions grow. For instance, in order to learn about the distribution of primes, we can look at the function

$$\pi(x) = \text{ the number of prime numbers } p \text{ up to } x, \quad x \geq 0.$$

We will discuss this function later. Here are some important arithmetic functions.

Example 6.1. The functions $\phi(n), \sigma(n)$ and $\tau(n)$ are all multiplicative, but not completely multiplicative functions.

Example 6.2. The Möbius function $\mu : \mathbb{Z}_{>0} \to \{0, \pm 1\} \subset \mathbb{C}$ is given by

$$\mu(n) = \begin{cases} 1 & n = 1, \\ (-1)^r & n = p_1 \dots p_r \text{ where } p_1, \dots, p_r \text{ are distinct primes, i.e. } n \text{ is square-free,} \\ 0 & n \text{ is not square-free.} \end{cases}$$

This is again a multiplicative, but not completely multiplicative function.

Theorem 6.3. *If $f(n)$ is a multiplicative function, then so is the function $F(n) = \sum_{d|n} f(d)$.*

Proof. Exercise. □

Proposition 6.4.

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

Proof. Define $M(n) = \begin{cases} 1 & n = 1, \\ 0 & n > 1. \end{cases}$ This is clearly a multiplicative function. Since $\mu(n)$ is multiplicative, Theorem 6.3 tells us that the left hand side function in the statement of the proposition is also multiplicative. Thus, it is enough to show that the two sides agree on prime powers. Let $n = p^a$, with $a > 0$. Then

$$\sum_{d|p^a} \mu(d) = \mu(1) + \mu(p) + \mu(p^2) + \cdots + \mu(p^a) = 1 + (-1) + 0 + \cdots + 0 = 0 = M(p^a).$$

Also, for $n = 1$ we have $\sum_{d|1} \mu(d) = \mu(1) = 1 = M(1)$, and the proof is complete. □

Theorem 6.5 (Möbius inversion)**.** *If $f : \mathbb{Z}_{>0} \to \mathbb{C}$ is any function (not necessarily multiplicative) and $F : \mathbb{Z}_{>0} \to \mathbb{C}$ is given by $F(n) = \sum_{d|n} f(d)$, then*

$$f(n) = \sum_{d|n} F(d)\mu\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d)\left(\frac{n}{d}\right) = \sum_{d_1 d_2 = n} F(d_1)\mu(d_2) = \sum_{d_1 d_2 = n} \mu(d_1)F(d_2).$$

Proof. It is clear that the four summation are the same. Then we see that

$$\sum_{d_1 d_2 = n} \mu(d_1)F(d_2) = \sum_{d_1 d_2 = n} \mu(d_1) \sum_{d|d_2} f(d) = \sum_{d_1 d | n} \mu(d_1) f(d) = \sum_{d|n} f(d) \sum_{d_1 | \frac{n}{d}} \mu(d_1). \tag{6.1}$$

Proposition 6.4 implies that $\sum_{d_1|\frac{n}{d}} \mu(d_1) = \begin{cases} 1 & d = n \\ 0 & d < n. \end{cases}$

Hence only the term with $d = n$ is nonzero in (6.1), and thus

$$\sum_{d_1 d_2 = n} \mu(d_1)F(d_2) = \sum_{d|n} f(d) \sum_{d_1|\frac{n}{d}} \mu(d_1) = f(n).$$

□

Proposition 6.6. *(i)* $\displaystyle\sum_{d|n} \phi(d) = n.$

(ii) $\displaystyle\frac{\phi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d}.$

Proof. The first part is left as an exercise. The second part follows from the first and Möbius inversion. □

Example 6.7. The von Mangoldt function $\mu : \mathbb{Z}_{>0} \to \mathbb{R} \subset \mathbb{C}$ is given by

$$\Lambda(n) = \begin{cases} \log p & n = p^a \text{ where } p \text{ is a prime,} \\ \\ 0 & \text{otherwise.} \end{cases}$$

This function is neither multiplicative nor additive. (Recall from HW3 that an additive function has the property that $f(mn) = f(m) + f(n)$ whenever m, n are relatively prime.) However it plays an important role in number theory and it has some of the same properties that mimic those of $\phi(n)$.

Proposition 6.8. *For all positive integers n we have*

$$\sum_{d|n} \Lambda(d) = \log n$$

and

$$\Lambda(n) = -\sum_{d|n} \mu(d) \log d.$$

Proof. For $n = 1$ both equations are immediate. For $n > 1$, we work with the prime decomposition $n = p_1^{a_1} \dots p_r^{a_r}$. The only divisors d of n for which the von Mangoldt function is nonzero are the ones that are prime powers, i.e. of the form p_i^b for some $1 \leq i \leq r$ and $1 \leq b \leq a_i$. Therefore

$$\sum_{d|n} \Lambda(d) = \sum_{b_1=1}^{a_1} \log p_1 + \cdots + sum_{b_r=1}^{a_r} \log p_r = \log\left(p_1^{a_1} \dots p_r^{a_r}\right) = \log n.$$

Thus the first identity is proved. Möbius inversion implies that

$$\Lambda(n) = \sum_{d|n} \mu(d) \log\left(\frac{n}{d}\right) = \log n \left(\sum_{d|n} \mu(d)\right) - \sum_{d|n} \mu(d) \log d.$$

Proposition 6.4 implies that the first term is always zero and the proof is complete. □

6.1 Big O and small o notation

Definition. *For two functions $f : D \to \mathbb{C}$ and $g : D \to \mathbb{R}_{\geq 0}$ defined on a set of real numbers D we say that*

$$f(x) = O(g(x))$$

if there exist positive constants A, M such that

$$|f(x)| \geq Ag(x) \text{ for all } x \in D, x > M.$$

We might also write $f \ll g$ or $g \gg f$.

We can make the same definition for $D \subset \mathbb{C}$, but then we need to have $|x| > M$.

Definition. *For two functions $f, g : D \to \mathbb{C}$ with $D \subset \mathbb{C}$ we say that*

$$f(x) = o(g(x))$$

if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

Definition. *We write $f(x) \asymp g(x)$ if $f \ll g$ and $g \ll f$.*

Definition. *We write $f(x) \sim g(x)$ if $\lim_{x\to\infty} \dfrac{f(x)}{g(x)} = 1$.*

Example 6.9.

- $x = o(x^2)$
- $\log x = o(x^\varepsilon)$ for any $\varepsilon > 0$
- $\log x^k = o(x^\varepsilon)$ for any $\varepsilon > 0$
- If $f(x)$ is continuous and defined everywhere, then $f(x) = O(1) \iff f(x)$ is bounded
- $x^2 + 19 \asymp 13x^2 - 7000$
- $x^2 + 19 \sim x^2 - 7000$
- $\lfloor x \rfloor = x + O(1) \sim x$.

Theorem 6.10. *There exists a constant $\gamma \in (0,1)$ such that*

$$\sum_{n\leq x} \frac{1}{n} = \log x + \gamma + O\left(\frac{1}{n}\right).$$

The number γ is called Euler's constant.

Proof. This is Theorem 6.10 in Leveque's book. Please read the proof in there. □

Theorem 6.11 (Partial summation). *Let $c_1, c_2, \dots$ be a sequence of complex numbers. Define*

$$S(x) = \sum_{n\leq x} c_n \quad \textit{for all real numbers } x.$$

Let n_0 be a positive integer and with the property that $c_j = 0$ for all $j < n_0$. Let $f : [n_0, \infty) \to \mathbb{C}$ a function that has a continuous derivative. Then

$$\sum_{n\leq x} c_n f(n) = S(x)f(x) - \int_{n_0}^{x} S(t)f'(t)dt.$$

Proof. This is Theorem 6.15 in Leveque's book with $\lambda_n = n$. Please read the proof in there. □

Proposition 6.12.

$$\sum_{n \leq x} \log n = x \log x - x + O(\log x).$$

Proof. We will apply the partial summation technique to the constant sequence $c_n = 1$ and the function $f(x) = \log x, x \geq 1$. Then $S(x) = \lfloor x \rfloor$ and we get

$$\sum_{n \leq x} \log n = \lfloor x \rfloor \log x - \int_1^x \frac{\lfloor t \rfloor}{t} dt$$

Since $\lfloor x \rfloor = x + O(1)$, the first term is $\lfloor x \rfloor \log x = x \log x + O(\log x)$ and the second term is

$$\int_1^x \frac{\lfloor t \rfloor}{t} dt = \int_1^x \frac{t + O(1)}{t} dt = \int_1^x 1 dt + O\left(\int_1^x \frac{dt}{t} \right) = x - 1 + O(\log x) = x + O(\log x).$$

The statement follows by subtracting the two terms. □

Note: From now on the index p on summation (or products) will mean a sum over primes. For instance, $\sum_{p \leq x}$ means sum over the primes up to x.

In particular,

$$\pi(x) = \sum_{p \leq x} 1 = O(x).$$

Around 1800, enough tables of primes had been computed to figure out how $\pi(x)$ grows. The following is a prediction made by Gauss. It was later (in 1895) proven by de la

Theorem 6.13. *[Prime Number Theorem - weak version] As $x \to \infty$,*

$$\pi(x) \sim \frac{x}{\log x}.$$

This asymptotic was later (in 1895) proven independently by two mathematicians, Hadamard and de la Vallée-Poussin. But the first substantive progress towards proving the Prime Number Theorem was made by Chebycheff in 1850 when he managed to show that

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

Later better approximations for $\pi(x)$ were found, but this is definitely the beginning of the story. The best possible result would be equivalent to the Riemann Hypothesis.

Here are two important functions in analytic number theory. They both appear in Chebycheff's work.

Definition.

$$\theta(x) = \sum_{p \le x} \log p$$

and

$$\psi(x) = \sum_{p^a \le x} \log p.$$

Note that $\psi(x) = \sum_{n \le x} \Lambda(n)$.

Theorem 6.14 (Chebycheff)**.** *There exist positive constants A, B such that*

$$Ax < \theta(x) < Bx.$$

In other words, $\theta(x) \asymp x$.

Ramanujan's proof. Ramanujan's proof starts with the observation that if we have a decreasing sequence of non-negative numbers

$$a_0 \ge a_1 \ge a_2 \ge$$

that tend to 0, then

$$a_0 - a_1 \le \sum_{n=0}^{\infty} (-1)^n a_n \le a_0 - a_1 + a_2.$$

Indeed,

$$\sum_{n=0}^{\infty} (-1)^n a_n = a_0 - a_1 + (a_2 - a_3) + \cdots \ge a_0 - a_1$$

and

$$\sum_{n=0}^{\infty} (-1)^n a_n = a_0 - a_1 + a_2 - (a_3 - a_4) - (a_5 - a_6) - \cdots \le a_0 - a_1 + a_2.$$

Define $T(x) = \sum_{n \le x} \psi\left(\frac{x}{n}\right)$. Then

$$\sum_{n \le x} \log n = \sum_{n \le x} \left(\sum_{p^a | n} \log p \right) = \sum_{m \le x} \psi\left(\frac{x}{m}\right) = T(x).$$

It follows from Proposition 6.12 that

$$T(x) = x \log x - x + O(\log x).$$

Hence

$$T(x) - 2T\left(\frac{x}{2}\right) = x \log x - x + O(\log x) - 2\left(\frac{x}{2} \log \frac{x}{2} - \frac{x}{2} + O(\log x)\right) = x \log 2 + O(\log x).$$

On the other hand, from the definition on $T(x)$ we see that

$$T(x) - 2T\left(\frac{x}{2}\right) = \sum_{n \le x} (-1)^{n-1} \psi\left(\frac{x}{n}\right).$$

Since $a_{n-1} = \psi\left(\frac{x}{m}\right)$ is a decreasing sequence of nonnegative numbers that converges to 0, the observation at the beginning of the proof tells us that

$$\psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) \ge x \log 2 + O(\log x) \tag{6.2}$$

and

$$\psi(x) - \psi\left(\frac{x}{2}\right) \le x \log 2 + O(\log x) \tag{6.3}$$

We can apply (6.3) to $x/2, x/4, \ldots$ By applying to $\frac{x}{2^k}$ we get that

$$\psi\left(\frac{x}{2^k}\right) - \psi\left(\frac{x}{2^{k+1}}\right) \le \frac{x}{2^k} \log 2 + O(\log x - k \log 2) = \frac{x}{2^k} \log 2 + O(\log x).$$

We sum over $0 \le k \le \log_2(x) = \frac{\log x}{\log 2}$. Since $\psi(y) = 0$ for $y < 2$, and $\displaystyle\sum_{k=0}^{\infty} \frac{1}{2^k} = 2$ the summation yields

$$\psi(x) \le 2x \log 2 + O(\log^2 x).$$

For each k we add have one instance of $O(\log x)$. Since we have less than $\log x$ term, the contribution of the error terms is $\log x O(\log x) = O(\log^2 x)$. Since $\theta(x) \le \psi(x)$, we have proved that $\theta(x) \le 2x \log 2 + O(\log^2 x)$, so there exists a constant $B > 0$ such that $\theta(x) < Bx$.

The rest of the proof is left as an exercise.

□

Chebycheff's proof. The key observation is that

$$\prod_{n < p \le 2n} p \,\Big|\, \binom{2n}{n}$$

Therefore

$$\prod_{n < p \le 2n} p \le \binom{2n}{n} \le 2^{2n}$$

and taking logarithms we obtain

$$\theta(2n) - \theta(n) \le 2n \log 2.$$

Similarly, one gets that

$$\theta(2n+1) - \theta(n) \le (2n+1) \log 2.$$

Putting the two previous inequalities together one gets

$$\theta(2x) - \theta(x) \leq 2x \log 2. \tag{6.4}$$

Applying this identity to $\frac{x}{2^k}$ and summing over $k \geq 0$ we find that

$$\theta(x) \leq 2x \log 2$$

and the second inequality is proved with $B = 2\log 2$. Note that we already have $\theta(x) = O(x)$.

For the second part of the proof, see HW4. (But that follows Ramanujan, rather than Chebycheff.) □

Corollary 6.15.

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

Proof. Theorem 6.14 implies that $\theta(x) = O(x)$.

We know that

$$\theta(x) = \sum_{p \leq x} \log p \geq \sum_{\sqrt{x} < p \leq x} \log p \geq \log(\sqrt{x}\,) \sum \sum_{\sqrt{x} < p \leq x} 1 = \frac{\log x}{2} \left(\pi(x) - \pi(\sqrt{x}\,)\right).$$

Therefore,

$$\pi(x) \log x \leq 2\theta(x) + \pi(\sqrt{x}\,) \log x = 2\theta(x) + O(\sqrt{x} \log x) \quad \text{since } \pi(\sqrt{x}) = O(\sqrt{x}\,).$$

Theorem 6.14 implies that $\theta(x) = O(x)$. Plugging this in the inequality above we get

$$\pi(x) \log x = O(x) + O(\sqrt{x} \log x) = O(x).$$

□

Even though Cebycheff could not prove the Prime Number Theorem, he was able to prove another important result. That is, Theorem 6.14 implies that $\theta(Bx/A) > \theta(x)$, so there exists a prime between x and Bx/A. By obtaining explicit constants A and B with $B/A \leq 2$, Chebycheff was able to prove the following result.

Theorem 6.16 (Bertrand's postulate)**.** *There is always a prime between n and $2n$, for $n > 1$.*

Proposition 6.17.

$$\sum_{p \leq n} \frac{\log p}{p} = \log n + O(1).$$

Proof. We start by observing that only primes up to n can divide $n!$ and so the factorization into primes of the factorial has the form

$$n! = \prod_{p|n} p^{e_p}.$$

The number of multiples of p that are $\leq n$ is $\left\lfloor \frac{n}{p} \right\rfloor$. The number of multiples of p^2 that are $\leq n$ is $\left\lfloor \frac{n}{p^2} \right\rfloor$, and so on. Thus

$$e_p = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \dots$$

Note that this is a finite sum, as at some point the powers of p will become larger than n. Taking logs we obtain

$$\log(n!) = \sum_{p\leq n} e_p \log p = \sum_{p \leq n} \log p \left(\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots \right).$$

On the other hand, Proposition 6.12 implies that

$$\log(n!) = \sum_{k\leq n} \log k = n \log n - n + O(\log n)$$

and

$$\sum_{p\leq n} \log p \left(\left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \dots \right) \leq \sum_{p\leq n} \log p \frac{n}{p(p-1)} \leq n \sum_{k=2}^{\infty} \frac{\log k}{k(k-1)} = O(n)$$

since the series is convergent.

Therefore

$$\sum_{p\leq n} \log p \left\lfloor \frac{n}{p} \right\rfloor = n \log n - n + O(n) = n \log n + O(n).$$

Now

$$\sum_{p\leq n} \log p \left\lfloor \frac{n}{p} \right\rfloor \geq \sum_{p\leq n} \log p \left(\frac{n}{p} - 1 \right) = \sum_{p\leq n} \frac{n}{p} \log p - \sum_{p\leq n} \log p \geq n \sum_{p\leq n} \frac{\log p}{p} - \pi(n) \log n,$$

Corollary 6.15 implies that $\pi(n) \log n = O(n)$ and so

$$n \sum_{p\leq n} \frac{\log p}{p} \leq \sum_{p\leq n} \log p \left\lfloor \frac{n}{p} \right\rfloor + (\log n)\pi(n) = n \log n + O(n) + O(n) = n \log n + O(n).$$

Dividing by n yields the desired result. □

Theorem 6.18.

$$\sum_{p\leq n} \frac{1}{p} = \log\log n + O(1).$$

Proof. Set

$$c_n = \begin{cases} \frac{\log p}{p} & n = p \\ 0 & n \text{ is not prime.} \end{cases}$$

and $f(x) = \frac{1}{\log x}$.

Then $S(n) = \sum k \leq nc_k = \sum_{k\leq n} \frac{\log p}{p} = \log n + O(1)$.

$$\sum_{p\leq n} \frac{1}{p} = \sum_{2\leq k\leq n} c_k f(k) = S(n)f(n) + \int_2^n S(t)\frac{1}{t\log^2 t}dt$$

The first term is

$$S(n)f(n) = 1 + O\left(\frac{1}{\log n}\right) = 1 + O\left(\frac{1}{\log n}\right).$$

The second term is

$$\int_2^n \frac{S(t)}{t\log^2 t}dt = \int_2^n \frac{\log t + O(1)}{t\log^2 t}dt = \int_2^n \frac{1}{t\log t}dt + O\left(\int_2^n \frac{1}{t\log^2 t}dt\right)$$

$$= \log\log n - \log\log 2 + O\left(\frac{1}{\log n} - \frac{1}{\log 2}\right) = \log\log n + O(1).$$

Adding them up gives the desired result. □

7 Dirichlet series

Definition. *A Dirichlet series is a function of the form*

$$F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

where (a_n) is some sequence of complex numbers.

Note that in order for this definition to make sense we need to know that the series converges. It will typically not converge everywhere.

Example 7.1. The Riemann zeta function is

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

As you know from calculus, this series is absolutely convergent whenever the exponent is bigger than 1. But in order to see how powerful Riemann's ideas are, one should really think of s as a complex number, $s = \sigma + it$. Then

$$n^s = n^{\sigma+it} = n^\sigma \cdot n^{it}.$$

We know very well what the first exponential, n^σ is, as both numbers are real. For the second, we will write $n = e^{\log n}$ and see that

$$n^{it} = e^{it\log n} = \cos(t\log n) + i\sin(t\log n) \implies \left|n^{it}\right| = \sqrt{\cos^2(t\log n) + \sin^2(t\log n)} = 1.$$

Thus

$$\left|n^{\sigma+it}\right| = n^\sigma$$

and

$$\sum_{n=1}^{\infty}\left|\frac{1}{n^s}\right| = \sum_{n=1}^{\infty}\frac{1}{\left|n^{\sigma+it}\right|} = \sum_{n=1}^{\infty}\frac{1}{n^\sigma}.$$

This implies that the series $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^s}$ is absolutely convergent for $\sigma = \mathrm{Re}(s) > 1$.

Assume $\mathrm{Re}(s) > 1$. (Or if you prefer, think of s as a real number > 1.) Consider the product over all primes

$$\prod_p\left(1-\frac{1}{p^s}\right)^{-1} = \prod_p\frac{1}{1-\frac{1}{p^s}} = \prod_p\left(1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\dots\right)$$

since $\frac{1}{1-x} = 1 + x + x^2 + \dots$ whenever $|x| < 1$. Unique factorization tells us that when we expand the product we get

$$\sum_{n=1}^{\infty}\frac{1}{n^s}.$$

We have proved the following result.

Theorem 7.2 (Euler product expansion)**.** *For* $\mathrm{Re}(s) > 1$,

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s} = \prod_p\left(1-\frac{1}{p^s}\right)^{-1}.$$

The Riemann zeta has many wonderful properties. It can be extended to a (meromorphic) function on all of $\mathbb{C}$, and the Euler product expansion above allows one to relate the zeros of $\zeta(s)$ to the prime numbers. The best possible approximation for $\pi(x)$ depends on the Riemann Hypothesis (see above).

The Euler product expansion is not unique to the Riemann zeta. For any nonzero multiplicative function $f(n)$ we have

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \dots\right).$$

If $f(n)$ happens to be completely multiplicative, we can write this further as

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 - \frac{f(p)}{p^s}\right)^{-1}.$$

7.1 Operations with Dirichlet series

To add two Dirichlet series $F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ and $G(s) = \sum_{n=1}^{\infty} \frac{b_n}{n^s}$, one just adds term by term:

$$F(s) + G(s) = \sum_{n=1}^{\infty} \frac{a_n + b_n}{n^s}.$$

For multiplication by a scalar $c \in \mathbb{C}$ we have

$$cF(s) = \sum_{n=1}^{\infty} \frac{ca_n}{n^s}.$$

Multiplication is a bit more challenging.

$$F(s)G(s) = \left(\sum_{n=1}^{\infty} \frac{a_n}{n^s}\right) \cdot \left(\sum_{n=1}^{\infty} \frac{b_n}{n^s}\right) = \left(\sum_{n=1}^{\infty} \frac{a_m}{m^s}\right) \cdot \left(\sum_{n=1}^{\infty} \frac{b_d}{d^s}\right) = \sum_{m,d\geq 1} \frac{a_m b_d}{(md)^s}.$$

Note that the denominators are again positive integers to the power s and we get a given n^s in the denominator whenever $n = md$. Thus,

$$F(s)G(s) = \sum_{m,d\geq 1} \frac{a_m b_d}{(md)^s} = \sum_{n=1}^{\infty} \frac{c_n}{n^s}, \quad \text{where} \quad c_n = \sum_{md=n} a_m b_d = \sum_{d|n} a_{n/d} b_d = \sum_{d|n} a_d b_{n/d}.$$

Proposition 7.3. *For* $\mathrm{Re}(s) > 1$,

$$\zeta^2(s) = \sum_{n=1}^{\infty} \frac{\tau(n)}{n^s}.$$

7.2 More about Riemann zeta

We want to explore what happens to $\zeta(s)$ as s approaches 1. We will first compute the limit

$$\lim_{s\to 1^+} \zeta(s).$$

Fix $s > 1$ (real). We will apply partial summation to $c_n = 1$ for all n and $f(x) = x^{-s}$. Then

$$S(x) = \sum_{n \leq x} 1 = \lfloor x \rfloor$$

and

$$\sum_{n \leq x} \frac{1}{n^s} = \sum_{n \leq x} c_n f(n) = S(x)f(x) - \int_1^x S(t)f'(t)dt = \frac{\lfloor x \rfloor}{x^s} + s\int_1^x \frac{\lfloor t \rfloor}{t^{s+1}}dt.$$

We denote by $((t)) = t - \lfloor t \rfloor$ the fractional part of t. Then

$$\sum_{n \leq x} \frac{1}{n^s} = \frac{\lfloor x \rfloor}{x^s} + s\int_1^x \frac{t - ((t))}{t^{s+1}}dt = \frac{\lfloor x \rfloor}{x^s} + s\int_1^x t^{-s}dt - s\int_1^x \frac{((t))}{t^{s+1}}dt$$

$$= \frac{\lfloor x \rfloor}{x^s} + \frac{s}{s-1}(1 - x^{1-s}) + \int_1^x st^{-(s+1)}dt.$$

Hence,

$$\zeta(s) = \lim_{x \to \infty} \sum_{n \leq x} \frac{1}{n^s} = \frac{s}{s-1} + \int_1^\infty st^{-(s+1)}dt.$$

On the other hand

$$0 \leq \int_1^\infty s\frac{((t))}{t^{s+1}}dt \leq \int_1^\infty s\frac{1}{t^{s+1}}dt = 1,$$

so the integral converges. Thus

$$\lim_{s \to 1^+}(s-1)\zeta(s) = \lim_{s \to 1^+} s + \lim_{s \to 1^+}(s-1)\int_1^\infty s\frac{((t))}{t^{s+1}}dt = 1 + 0 = 1.$$

This implies the following.

Theorem 7.4.

$$\lim_{s \to 1^+} \zeta(s) = +\infty.$$

In complex analysis terms, $\zeta(s)$ has a simple pole at $s = 1$.

7.3 Other zeta and L-functions

We can think of

$$\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s} = \sum_{\substack{n \in \mathbb{Z}\setminus\{0\} \\ \text{up to sign}}} \frac{1}{|n|^s} = \sum_{n \in (\mathbb{Z}\setminus\{0\})/\mathbb{Z}^\times} \frac{1}{|n|^s}$$

as $\mathbb{Z}^\times = \{n \in \mathbb{Z}; 1/n \in \mathbb{Z}\} = \{\pm 1\}$.

Similarly, we can define for gaussian integers

$$\zeta_{\mathbb{Q}(i)}(s) = \sum_{a+bi\in(\mathbb{Z}[i]\setminus\{0\})/\mathbb{Z}[i]^\times} \frac{1}{N(a+bi)^s} = \sum_{a+bi\in(\mathbb{Z}[i]\setminus\{0\})/\mathbb{Z}[i]^\times} \frac{1}{(a^2+b^2)^s} = \sum_{n=1}^{\infty} \frac{r(n)}{n^s},$$

where $r(n) = \frac{1}{4}$ of the ways that n can be written as sum of two squares. Hence $\zeta_{\mathbb{Q}(i)}(s)$ is also a Dirichlet series.

Moreover, unique factorization in $\mathbb{Z}[i]$ implies that

$$\zeta_{\mathbb{Q}(i)}(s) = \prod_{\pi} \left(1 - \frac{1}{N(\pi)^s}\right)^{-1}$$

where the product runs over gaussian primes π up to units. We do know all gaussian primes:

- $\pi = 1+1 \implies N(\pi) = 2$ and $\left(1 - \frac{1}{N(\pi)^s}\right)^{-1} = \left(1 - \frac{1}{2^s}\right)^{-1}$.

- $\pi = p \equiv 3 (\text{mod } 4) \implies N(\pi) = p^2$ and $\left(1 - \frac{1}{N(\pi)^s}\right)^{-1} = \left(1 - \frac{1}{p^{2s}}\right)^{-1}$.

- π and $\bar{\pi}$ such that $\pi\bar{\pi} = p \equiv 1 (\text{mod } 4) \implies N(\pi) = N(\bar{\pi}) = p$. Hence
$$\left(1 - \frac{1}{N(\pi)^s}\right)^{-1} = \left(1 - \frac{1}{N(\bar{\pi})^s}\right)^{-1} = \left(1 - \frac{1}{p^s}\right)^{-1}.$$

Therefore

$$\begin{aligned}\zeta_{\mathbb{Q}(i)}(s) &= \left(1 - \frac{1}{2^s}\right)^{-1} \prod_{p\equiv 1 \pmod 4} \left(1 - \frac{1}{p^s}\right)^{-2} \prod_{p\equiv 3 \pmod 4} \left(1 - \frac{1}{p^{2s}}\right)^{-1} \\ &= \zeta(s) \prod_{p\equiv 1 \pmod 4} \left(1 - \frac{1}{p^s}\right)^{-1} \prod_{p\equiv 3 \pmod 4} \left(1 + \frac{1}{p^s}\right)^{-1}. \qquad (7.1)\end{aligned}$$

On the other hand, recall the group homomorphism $\chi_4 : (\mathbb{Z}/4\mathbb{Z})^\times \to \mathbb{C}*$, $\chi_4(a(\text{mod } 4)) = \left(\frac{-1}{a}\right) = (-1)^{\frac{a-1}{2}}$. We can extend this to a completely multiplicative function $\chi : \mathbb{Z} \to \mathbb{C}$ given by

$$\chi(n) = \begin{cases} \chi_4(n \pmod 4) & \text{if } \gcd(n,4) = 1 \\ 0 & \text{if } \gcd(n,4) > 1 \end{cases} = \begin{cases} (-1)^{\frac{n-1}{2}} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

Then the Dirichlet series associated to χ is

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_{p} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} = \prod_{p\equiv 1 \pmod 4} \left(1 - \frac{1}{p^s}\right)^{-1} \prod_{p\equiv 3 \pmod 4} \left(1 + \frac{1}{p^s}\right)^{-1} \qquad (7.2)$$

We can replace (7.2) into (7.1) and get

$$\zeta_{\mathbb{Q}(i)}(s) = \zeta(s)L(s,\chi).$$

This relation can be used to show the following result.

Theorem 7.5.

$$\lim_{s\to 1^+} \prod_{p\equiv 1 \pmod 4} \left(1 - \frac{1}{p^s}\right)^{-1} = +\infty.$$

Therefore there are infinitely many primes $p \equiv 1(\text{mod } 4)$.

Proof. We know that

$$\left(1-\frac{1}{2^s}\right)^{-1} \prod_{p\equiv 1 \pmod 4} \left(1-\frac{1}{p^s}\right)^{-2} \prod_{p\equiv 3 \pmod 4} \left(1-\frac{1}{p^{2s}}\right)^{-1} = \zeta_{\mathbb{Q}(i)}(s) = \zeta(s)L(s,\chi).$$

First we look at the product over primes $\equiv 3(\text{mod } 4)$. For $s > 1$ we have

$$\prod_{p\equiv 3 \pmod 4} \left(1-\frac{1}{p^{2s}}\right)^{-1} < \prod_{p\equiv 3 \pmod 4} \left(1-\frac{1}{p^{2}}\right)^{-1} < \prod_{p} \left(1-\frac{1}{p^{2}}\right)^{-1} = \zeta(2)$$

is a finite number. In fact, $\zeta(2) = \pi^2/6$ but we will not prove this. Since the product over primes $\equiv 3(\text{mod } 4)$ is finite, it follows that

$$\begin{aligned}
\lim_{s\to 1^+} \zeta_{\mathbb{Q}(i)}(s) = \infty &\iff \lim_{s\to 1^+} \prod_{p\equiv 1 \pmod 4} \left(1-\frac{1}{p^s}\right)^{-2} = \infty \\
&\iff \lim_{s\to 1^+} \prod_{p\equiv 1 \pmod 4} \left(1-\frac{1}{p^s}\right)^{-1} = \infty.
\end{aligned}$$

Since $\zeta_{\mathbb{Q}(i)}(s) = \zeta(s)L(s,\chi)$ and $\lim_{s\to 1^+}\zeta(s) = \infty$ it is enough to show that $\lim_{s\to 1^+} L(s,\chi)$ exists, is finite and nonzero. However, $L(s,\chi)$ is continuous at $s = 1$ (exercise). Hence all we need to show is that $L(1,\chi) \neq 0$. Also from continuity it follows that

$$L(1,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots < 1$$

and

$$L(1,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots > 1 - \frac{1}{3} = \frac{2}{3}.$$

□

7.4 The average order of $\phi(n)$

We will now discuss an application of Dirichlet series.

Proposition 7.6. *For* $\mathrm{Re}(s) > 1$,

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

Proof. Again, for simplicity, we will think of $s > 1$ real. As $|\mu(n)| \leq 1$ for all positive integers n, the the series on the right hand side in absolutely convergent for $s > 1$. Taking the product with $\zeta(s)$ we obtain

$$\zeta(s) \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \sum_{m=1}^{\infty} \frac{1}{m^s} \sum_{d=1}^{\infty} \frac{\mu(d)}{d^s} = \sum_{n=1}^{\infty} \frac{c_n}{n^s}$$

where

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

Proposition 6.4 implies that

$$c_n = \begin{cases} 1 & n = 1 \\ 0 & n > 1. \end{cases}$$

Therefore

$$\sum_{n=1}^{\infty} \frac{c_n}{n^s} = \frac{1}{1^s} = 1$$

and the result is proved.

□

Theorem 7.7.

$$\sum_{m=1}^{n} \phi(m) = \frac{3n^2}{\pi^2} + O(n \log n).$$

Proof. We know from Proposition 6.6 that

$$\phi(m) = m \sum_{d|m} \frac{\mu(d)}{d}.$$

Plugging into the sum we want to estimate we get

$$\sum_{m=1}^{n} \phi(m) = \sum_{m=1}^{n} m \sum_{d|m} \frac{\mu(d)}{d} = \sum_{1 \leq m \leq n, d|m} \frac{m}{d} \mu(d) = \sum_{dd' \leq n} d' \mu(d) = \sum_{d=1}^{n} \mu(d) \sum_{d' \leq \frac{n}{d}} d'.$$

But

$$\sum_{d' \leq \frac{n}{d}} d' = \frac{1}{2} \left\lfloor \frac{n}{d} \right\rfloor \left(\left\lfloor \frac{n}{d} \right\rfloor + 1 \right),$$

therefore

$$\sum_{m=1}^{n} \phi(m) = \sum_{d=1}^{n} \mu(d) \frac{\left\lfloor \frac{n}{d} \right\rfloor^2 + \left\lfloor \frac{n}{d} \right\rfloor}{2}.$$

We also know that $\mu(m) = O(1)$ and that $\lfloor x \rfloor = x + O(1) \implies \lfloor x \rfloor^2 = x^2 + O(x)$. Using these estimates in the above equation yields

$$\sum_{m=1}^{n} \phi(m) = \frac{1}{2} \sum_{d=1}^{n} \mu(d) \left(\frac{n}{d} \right)^2 + O\left(\sum_{d=1}^{n} \frac{n}{d} \right).$$

We now look at the error term. We know from Theorem 6.10 that

$$\sum_{d=1}^{n} \frac{1}{d} = \log n + O(1) \implies \sum_{d=1}^{n} \frac{n}{d} = n \log n + O(n) \implies O\left(\sum_{d=1}^{n} \frac{n}{d} \right) = O(n \log n).$$

Therefore

$$\sum_{m=1}^{n} \phi(m) = \frac{1}{2} \sum_{d=1}^{n} \mu(d) \left(\frac{n}{d} \right)^2 + O(n \log n) = \frac{n^2}{2} \sum_{d=1}^{n} \frac{\mu(d)}{d^2}.$$

We now examine the main term.

$$\sum_{d=1}^{n} \frac{\mu(d)}{d^2} = \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - \sum_{d=n+1}^{\infty} \frac{\mu(d)}{d^2} \overset{\text{Prop 7.6}}{=} \frac{1}{\zeta(2)} - \sum_{d=n+1}^{\infty} \frac{\mu(d)}{d^2}.$$

However,

$$\sum_{d=n+1}^{\infty} \frac{\mu(d)}{d^2} = O\left(\sum_{d=n+1}^{\infty} \frac{1}{d^2} \right)$$

and

$$\sum_{d=n+1}^{\infty} \frac{1}{d^2} \leq \int_{n}^{\infty} \frac{dx}{x^2} = \frac{1}{n}$$

Going back to the sum of interest, we get

$$\sum_{m=1}^{n} \phi(m) = \frac{n^2}{2\zeta(2)} + n^2 O\left(\frac{1}{n} \right) + O(n \log n) = \frac{n^2}{2\zeta(2)} + O(n) + O(n \log n) = \frac{n^2}{2\zeta(2)} + O(n \log n).$$

Plugging in $\zeta(2) = \pi^2/6$ we obtain the desired result.

□

8 Sieving

8.1 The sieve of Eratosthenes

One method for estimating the size of $\pi(x)$ is based upon the observation that if an integer $1 < n \leq x$ is not divisible by any prime $p \leq \sqrt{x}$, then n is prime. Thus if we list all the numbers up to x and discard all the multiples of 2, then the multiples of 3 and so on until all the multiples of primes up to $\sqrt{x}$ have been discarded, then the remaining numbers are all prime. This is called the sieve of Eratosthenes.

We can modify the process described above by discarding the multiples of the first r primes, $p_1, p_2, \ldots, p_r$. We retain r as an independent variable until the best choice for r becomes apparent. If there are any primes between p_r and $\sqrt{x}$ ($p_r < p \leq \sqrt{x}$) then it is no longer the case that all the remaining numbers are prime. However no primes other than $p_1, p_2, \ldots, p_r$ have been removed. We will denote by $A(x, r)$ the number of integers up to x that are no multiples of $p_1, \ldots, p_r$. Then

$$\pi(x) \leq r + A(x, r).$$

Next we would like to estimate $A(x, r)$. For that we will use Theorem 6.4 from the textbook with $S = \{n \in \mathbb{Z}_{>0}; n \leq x\}$ and $S_k = \{n \in S; p_k \mid n\}$ for each $1 \leq k \leq r$. Then Theorem 6.4 from the textbook tells us that

$$A(x,r) = \#(S \setminus \bigcup_{k=1}^{r} S_k) = \#S - \sum_{k=1}^{r} \#S_k + \sum_{k \neq j} \#(S_k \cap S_j) - \ldots$$

But $\#S = \lfloor x \rfloor, \#S_k = \left\lfloor \frac{x}{p_k} \right\rfloor, \#(S_k \cap S_j) = \left\lfloor \frac{x}{p_k p_j} \right\rfloor \ldots$ Hence

$$A(x,r) = \lfloor x \rfloor - \sum_{k=1}^{r} \left\lfloor \frac{x}{p_k} \right\rfloor + \sum_{k \neq j} \left\lfloor \frac{x}{p_k p_j} \right\rfloor - \cdots + (-1)^r \left\lfloor \frac{x}{p_1 \ldots p_r} \right\rfloor. \qquad (8.1)$$

The difference between each term and the same term without the floor function is at most 1. Therefore

$$\begin{aligned} A(x,r) &\leq x - \sum_{k=1}^{r} \frac{x}{p_k} + \cdots + (-1)^r \frac{x}{p_1 \ldots p_r} + 1 + \binom{r}{1} + \cdots + \binom{r}{r} \\ &= x \prod_{k=1}^{r} \left(1 - \frac{1}{p_k}\right) + 2^r \end{aligned}$$

We obtain that

$$\pi(x) \leq r + x \prod_{k=1}^{r} \left(1 - \frac{1}{p_k}\right) + 2^r \qquad (8.2)$$

Note that if we choose r such that p_r is the largest prime $\leq \sqrt{x}$, since $r > x^{\frac{1}{2} - \epsilon}$ (a fact Chebycheff proved), we have $2^r > x$. But we already know that $\pi(x) < x$, so (8.2) tells us nothing new. The underlying reason for getting such a bad estimate from (8.2) is the fact that the expression (8.1) has many terms equal to 0, and for each of them we make an error of almost 1 when we drop the floor function.

Our next goal is to see how we can do better with a different choice of r. First, a preliminary result that estimated the product of the primes in (8.2).

Proposition 8.1. *If $x \leq 2$, then*

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) < \frac{1}{\log x}$$

Proof.

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} = \prod_{p \leq x} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \dots\right)$$

The product is equal to $\sum n^{-1}$ taken over all n whose prime factors are all $\leq x$. All the integers up to x have this property, so we can write

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} > \sum_{n \leq x} \frac{1}{n} = 1 + \frac{1}{2} + \dots + \frac{1}{\lfloor x \rfloor} > \sum_{k=1}^{\lfloor x \rfloor} \int_k^{k+1} \frac{1}{k} dx > \int_1^{\lfloor x \rfloor + 1} \frac{dx}{x} = \log(\lfloor x \rfloor + 1) > \log x.$$

□

Now we can use the sieve of Eratothenes to find an upper bound for $\pi(x)$.

Proposition 8.2.

$$\pi(x) < \frac{x}{\log \log x} + o\left(\frac{x}{\log \log x}\right) \ll \frac{x}{\log \log x}.$$

Proof. From (8.2) we know that

$$\pi(x) < r + 2^r + x \prod_{k=1}^{r} \left(1 - \frac{1}{p_k}\right) \leq 2^{r+1} + x \prod_{k=1}^{r} \left(1 - \frac{1}{p_k}\right).$$

By Proposition 8.1 we then have

$$\pi(x) < 2^{r+1} + \frac{x}{\log p_r}.$$

Since $p_r > r$, the above inequality yields

$$\pi(x) < 2^{r+1} + \frac{x}{\log r}.$$

The moment has come for us to choose an appropriate r. Namely, for $r = 1 + \lfloor \log x \rfloor$, we get

$$\pi(x) < 2^{\lfloor \log x \rfloor + 2} + \frac{x}{\log(\lfloor \log x \rfloor + 1)} = \frac{x}{\log(\lfloor \log x \rfloor + 1)} + 4 \cdot 2^{\lfloor \log x \rfloor} < \frac{x}{\log \log x} + 4 \cdot 2^{\log x}.$$

Notice that $2^{\log x} = x^{\log 2} = o\left(\frac{x}{\log\log x}\right)$ since $\log 2 < 1$. Therefore

$$\pi(x) < \frac{x}{\log\log x} + o\left(\frac{x}{\log\log x}\right) = O\left(\frac{x}{\log\log x}\right).$$

□

Proposition 8.1 tells us that the primes are not too rare. If, for instance, they were rarer than squares (i.e. $p_k > k^2$ for some $k \geq a$) then we would have

$$\prod_{k=1}^{r}\left(1-\frac{1}{p_k}\right) > A\prod_{k=a}^{\infty}\left(1-\frac{1}{k^2}\right),$$

where

$$A = \prod_{k=1}^{a-1}\left(1-\frac{1}{p_k}\right).$$

Note that A is a positive constant. The other product is

$$\prod_{k=a}^{r}\left(1-\frac{1}{k^2}\right) = \prod_{k=a}^{r}\frac{k-1}{k}\frac{k+1}{k} = \left(\prod_{k=a}^{r}\frac{k-1}{k}\right)\left(\prod_{k=a}^{r}\frac{k+1}{k}\right) = \frac{a-1}{r}\cdot\frac{r+1}{a} > \frac{1}{2}.$$

This would tell us that

$$\prod_{k=1}^{r}\left(1-\frac{1}{p_k}\right) > \frac{A}{2}$$

for all $r > a$, which contradicts Proposition 8.1.

On the other hand, Proposition 8.2 tells us that primes are not too frequent either. But in the course of the proof of Proposition 8.2 we had to use Proposition 8.1. That is, in order to prove that there are not too many primes, we had to use the fact that they are not too rare either.

8.2 The sieve problem

Let $\mathcal{P}$ be a set of primes (could be finite, or infinite, or even all primes). Assume we have a finite set $\mathcal{A}$ of objects, and a way of associating a subset $\mathcal{A}_p \subset \mathcal{A}$ to each prime $p \in \mathcal{P}$. The *sieve problem* is to estimate (from above and below) the size of the set

$$\mathcal{S}(\mathcal{A},\mathcal{P}) = \mathcal{A} \setminus \left(\bigcup_{p\in\mathcal{P}}\mathcal{A}_p\right) = \bigcap_{p\in\mathcal{P}}\left(\mathcal{A}\setminus\mathcal{A}_p\right).$$

This is the formulation of the problem in the most general context. The explicit answer will usually come from the inclusion-exclusion principle. That is, from Theorem 6.4 in the textbook which can be reformulated in this context as follows.

Theorem 8.3 (Inclusion-exclusion principle)**.** *With the notation above we have*

$$\#\mathcal{S}(\mathcal{A},\mathcal{P}) = \sum_{\mathcal{I}\subseteq\mathcal{P}} (-1)^{\#\mathcal{I}} \#\mathcal{A}_{\mathcal{I}},$$

where $\mathcal{A}_{\mathcal{I}} = \bigcap_{p\in\mathcal{I}} \mathcal{A}_p$ *for each nonempty subset* $\mathcal{I} \subseteq \mathcal{P}$ *and* $\mathcal{A}_\emptyset = \mathcal{A}$.

This formula is the basis of many questions in probability theory. In number theory, we often take $\mathcal{A}$ to be a finite set of positive integers and $\mathcal{A}_p$ to be the subset of $\mathcal{A}$ consisting of elements that lye in certain congruence classes modulo p.

Example 8.4. For the sieve of Eratosthenes, $\mathcal{A} = \{n \in \mathbb{Z}_{>0}; n \leq x\}, \mathcal{P} = \{p_1, \ldots, p_r\}$ the set of the first r primes, and for each $p \in \mathcal{P}$, the set corresponding subset of $\mathcal{A}$ is $\mathcal{A}_p = \{n \in \mathcal{A}; p \mid n\}$.

We could also reverse the perspective and think of $\mathcal{S}(\mathcal{A},\mathcal{P})$ as a given set whose size we want to estimate. We seek to do this by looking at its image modulo primes $p \in \mathcal{P}$ for some set of primes $\mathcal{P}$. This point of view is the one adopted in the sieving technique called the *large sieve.*

We could even enlarge this reversed perspective by looking modulo prime powers, not just primes. Namely, let $\mathcal{B}$ be a finite set of positive integers and $\mathcal{T}$ a set of prime powers. Suppose that we know the size of $\mathcal{B}(\text{mod } t)$ for any $t \in \mathcal{T}$. We then seek to estimate the size of $\mathcal{B}$ itself. This is the approach of the *larger sieve.*

8.3 Selberg sieve

First, let us rewrite the results we obtained from the sieve of Eratosthenes in Section 8.1 using more compact notation. For a given $z > 0$ we will write

$$\mathcal{P}_z = \{p \text{ prime}; p \leq z\} \text{ the set of primes up to } z$$

and

$$\pi(x,z) = \#\{n \in \mathbb{Z}_{>0}; n \leq x, p \nmid n \text{ for any prime } p \in \mathcal{P}_z\}.$$

Note that $\pi(x,z)$ is the number of positive integers up to x that are not divisible by any prime in $\mathcal{P}_z$ and $\pi(x,z) = \#A(x,r)$ in the notation of Section 8.1 where r is chosen such that p_r is the larger prime $\leq z$. As in Example 8.4, $\mathcal{A} = \{n \in \mathbb{Z}_{>0}; n \leq x\}, \mathcal{P} = \mathcal{P}_z$ and $\mathcal{A}_p = \{n \in \mathcal{A}; p \mid n\}$ for any $p \in \mathcal{P}_z$.

We will also write

$$P_z = \prod_{p\in\mathcal{P}_z} p.$$

With this notation (8.1) becomes

$$\pi(x,z) = \sum_{d\mid P_z} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor = \sum_{d\mid P_z} \mu(d) \left(\sum_{n\leq x, d\mid n} 1 \right) = \sum_{n\leq x} \left(\sum_{d\mid \gcd(n,P_z)} \mu(d) \right). \qquad (8.3)$$

Selberg noticed that one can replace the internal sum above by a quadratic form. His observation is the following.

Remark 8.5. For any sequence $(\lambda_d)_{d\geq 1}$ of real numbers with $\lambda_1 = 1$, we have

$$0 \leq \sum_{d|k} \mu(d) \leq \left(\sum_{d|k} \lambda_d \right)^2 . \tag{8.4}$$

Indeed, we know that

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

For $k = 1$ the inequality becomes

$$0 \leq 1 \leq (\lambda_1)^2 = 1.$$

For $k > 1$ the inequality is

$$0 \leq 0 \leq \left(\sum_{d|k} \lambda_d \right)^2 .$$

From (8.3) and (8.4) we obtain

$$\begin{aligned} \pi(x, z) &\leq \sum_{n\leq x} \left(\sum_{d|\gcd(n,P_z)} \lambda_d \right)^2 = \sum_{n\leq x} \left(\sum_{d_1,d_2|\gcd(n,P_z)} \lambda_{d_1}\lambda_{d_2} \right) \\ &= \sum_{d_1,d_2|P_z} \lambda_{d_1}\lambda_{d_2} \sum_{n\leq x, \mathrm{lcm}(d_1,d_2)|n} 1 = \sum_{d_1,d_2|P_z} \lambda_{d_1}\lambda_{d_2} \left\lfloor \frac{x}{\mathrm{lcm}(d_1, d_2)} \right\rfloor . \end{aligned}$$

Since $\lfloor x \rfloor = x + O(1)$, we get

$$\pi(x, z) \leq x \sum_{d_1,d_2\leq z} \frac{\lambda_{d_1}\lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)} + O\left(\sum_{d_1,d_2|P_z} |\lambda_{d_1}| \, |\lambda_{d_2}| \right) .$$

For simplicity, let us assume that $\lambda_d = 0$ for all $d > z$. The equation above becomes

$$\pi(x, z) \leq x \sum_{d_1,d_2\leq z} \frac{\lambda_{d_1}\lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)} + O\left(\sum_{d_1,d_2\leq z} |\lambda_{d_1}| \, |\lambda_{d_2}| \right) . \tag{8.5}$$

Note that if we knew that $|\lambda_d| \leq 1$ for all d, the last term in (8.5) would become $O(z^2)$. The key to using (8.5) is to note that the sum that appears in the “main term”

$$\sum_{d_1,d_2\leq z} \frac{\lambda_{d_1}\lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)}$$

is a quadratic form in $\lambda_d, 1 \leq d \leq z$ and try to minimize it.

We start by replacing $\mathrm{lcm}(d_1, d_2) = \dfrac{d_1 d_2}{\gcd(d_1, d_2)}$. This yields

$$\sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)} = \sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} \gcd(d_1, d_2)$$

Using the fact that $\sum_{n|k} \phi(n) = k$ we can write

$$\begin{aligned}\sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)} &= \sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} \sum_{n|\gcd(d_1, d_2)} \phi(n) = \sum_{n \leq z} \phi(n) \sum_{\substack{d_1, d_2 \leq z \\ n|d_1, n|d_2}} \frac{\lambda_{d_1} \lambda_{d_2}}{d_1 d_2} \\ &= \sum_{n \leq z} \phi(n) \left(\sum_{d \leq z, n|d} \frac{\lambda_d}{d} \right)^2 = \sum_{n \leq z} \phi(n) u_n^2 \qquad (8.6)\end{aligned}$$

for

$$u_n = \sum_{d \leq z, n|d} \frac{\lambda_d}{d} \qquad (8.7)$$

We will use the following result to extract λ_d as a function of the u_n's.

Theorem 8.6 (Dual Möbius inversion)**.** *Let $\mathcal{D}$ be a set of positive integers that is divisor closed (i.e. $d \in \mathcal{D}$ and $d' \mid d \implies d' \in \mathcal{D}$). Let $f(n), g(n)$ be complex-valued functions defined on the positive integers. Then*

$$f(n) = \sum_{n|d} g(d) \iff g(n) = \sum_{n|d} \mu\left(\frac{d}{n}\right) f(d)$$

provided the sums are absolutely convergent.

Proof. Exercise. □

First note that (8.7) implies that $u_n = 0$ for all $n > z$. Thus the dual Möbius inversion implies that

$$\frac{\lambda_n}{n} = \sum_{d \leq z, n|d} \mu\left(\frac{d}{n}\right) u_d. \qquad (8.8)$$

(Since both the sum above and the one in(8.7) have finitely many terms, they are both absolutely convergent.) Taking $n = 1$ in (8.8) we see that

$$\sum_{d \leq z} \mu(d) u_d = 1. \qquad (8.9)$$

Lemma 8.7. *With the notations in this section we have*

$$\sum_{n\leq z}\phi(n)u_n^2=\sum_{n\leq z}\phi(n)\left(u_n-\frac{\mu(n)}{\phi(n)V(z)}\right)^2+\frac{1}{V(z)}$$

where

$$V(z)=\sum_{d\leq z}\frac{\mu(d)^2}{\phi(d)}.$$

Proof. Exercise. □

Corollary 8.8.

$$\sum_{n\leq z}\phi(n)u_n^2\geq\frac{1}{V(z)}$$

and equality is obtained exactly when

$$u_n=\frac{\mu(n)}{\phi(n)V(z)}$$

for all $1\leq n\leq z$. *In this case*

$$\lambda_n=\frac{n}{V(z)}\sum_{d\leq z,n|d}\frac{\mu(d/n)\mu(d)}{\phi(d)}$$

for all $1\leq n\leq z$.

Proof. The first statement follows immediately from the previous lemma. The second statement follows from plugging the expression for u_d in(8.8). □

Therefore, with the choice of λ_d from the Lemma 8.7 we have from (8.5)

$$\pi(x,z)\leq\frac{x}{V(z)}+O\left(\sum_{d_1,d_2\leq z}|\lambda_{d_1}|\,|\lambda_{d_2}|\right). \tag{8.10}$$

We would like to have $|\lambda_d|\leq 1$ for all $d\leq z$. This is equivalent to$|\lambda_nV(z)|\leq V(z)$. Indeed, Corollary 8.8 implies that

$$\lambda_nV(z)=n\sum_{d\leq z,n|d}\frac{\mu(d/n)\mu(d)}{\phi(d)}=n\sum_{t\leq z/n}\frac{\mu(t)\mu(nt)}{\phi(nt)}$$

Since $\mu(nt)=0$ whenever n and t are not coprime we obtain that

$$\lambda_nV(z)=n\sum_{t\leq z/n,(t,n)=1}\frac{\mu(t)^2\mu(n)}{\phi(n)\phi(t)}=\frac{n\mu(n)}{\phi(n)}\sum_{t\leq z/n,(t,n)=1}\frac{\mu(t)^2}{\phi(t)}.$$

Therefore

$$\begin{aligned}|\lambda_n V(z)| &\le \frac{n}{\phi(n)} \sum_{t\le z/n,(t,n)=1} \frac{\mu(t)^2}{\phi(t)} = \prod_{p|n} \frac{p}{p-1} \sum_{t\le z/n,(t,n)=1} \frac{\mu(t)^2}{\phi(t)} \\ &= \prod_{\substack{d|n \\ d\ \square\text{-free}}} \frac{1}{\phi(d)} \sum_{t\le z/n,(t,n)=1} \frac{\mu(t)^2}{\phi(t)} \le \sum_{t\le z} \frac{\mu(t)^2}{\phi(t)} = V(z).\end{aligned}$$

Thus we do obtain $|\lambda_n| \le 1$ for all $n \le z$ as desired. Thus, (8.10) implies the following result.

Theorem 8.9. *As $x, z \to \infty$,*

$$\pi(x,z) \le \frac{x}{V(z)} + O(z^2)$$

where

$$V(z) = \sum_{d\le z} \frac{\mu(d)^2}{\phi(d)}.$$

We can now deduce Chebycheff's upper bound for $\pi(x)$.

Corollary 8.10.

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

Proof. We know that $\pi(x) \le z + \pi(x,z)$. Thus, Theorem 8.9 implies that

$$\pi(x) \le \frac{x}{V(z)} + O(z^2).$$

Now we need a lower bound for $V(z)$ and to choose an appropriate z. We have

$$V(z) = \sum_{d\le z} \frac{\mu(d)^2}{\phi(d)} \ge \sum_{d\le z} \frac{\mu(d)^2}{d} = \sum_{\substack{d\le z \\ d\ \square\text{-free}}} \frac{1}{d} = \sum_{d\le z} \frac{1}{d} - \sum_{\substack{d\le z \\ d\text{ not }\square\text{-free}}} \frac{1}{d}.$$

Recall that

$$\sum_{d\le z} \frac{1}{d} = \log z + O(1).$$

Moreover,

$$\sum_{\substack{d\le z \\ d\text{ not }\square\text{-free}}} \frac{1}{d} \le \frac{1}{4} \sum_{d\le z/4} \frac{1}{d} + \frac{1}{9} \sum_{d\le z/9} \frac{1}{d} + \cdots \le [\zeta(2)-1] \sum_{d\le z} \frac{1}{d}.$$

Therefore,

$$V(z) \ge \log z\,(1 - [\zeta(2)-1]) \log z + O(1).$$

Since

$$1 - [\zeta(2) - 1] = 2 - \zeta(2) = 2 - \frac{\pi^2}{6} > 0,$$

we obtain

$$V(z) \gg \log z.$$

Hence

$$\pi(x) \ll \frac{x}{\log z} + z^2.$$

Choosing $z = \left(\frac{x}{\log x}\right)^2$ gives the desired upper bound. □

9 Twin prime conjecture

9.1 A brief history

The twin prime conjecture states that there are infinitely many primes p with the property that $p + 2$ is also prime. In other words it states that there are infinitely many pairs of primes p, q such that $|p - q| = 2$. Two such primes are called twin primes, hence the name of the conjecture.

The approach that so far has proved the strongest result towards the this conjecture has been to consider

$$\liminf_{n\to\infty} (p_{n+1} - p_n)$$

where p_n denotes the nth prime and try to prove an upper bound for it.

Recall that for a sequence $(a_n)_{n\geq 1}$ we say that $b = \liminf a_n$ is the smallest number such that there is a subsequence $(a_{n_k})_{k\geq 1}$ of $(a_n)_{n\geq 1}$ whose limit is b.

The twin prime conjecture can be restated as follows.

Conjecture 9.1 (Twin prime conjecture)**.**

$$\liminf_{n\to\infty} (p_{n+1} - p_n) = 2.$$

In general, if one manages to prove that $\liminf(p_{n+1} - p_n) \leq A$, it means that there are infinitely many pairs of consecutive primes p_n, p_{n+1} such that the distance between the two primes is at most A.

In 2009 Goldston, Pintz and Yildirim introduced a new method for counting tuples of primes, and this allowed them to show that

$$\liminf \frac{p_{n+1} - p_n}{\log p_n} = 0.$$

This implies that the gaps between primes are much smaller that the logarithm of the primes infinitely often. But it does not tell us anything about bounded gaps. Then in April 2013, Zhang managed to prove that

$$\liminf_{n\to\infty}(p_{n+1}-p_n) < 70,000,000,$$

thereby establishing for the first time the existence of infinitely many bounded gaps between primes. Later the same year, in November 2013, Maynard used a substantially easier method (with an extension of the Selberg sieve as a key ingredient) to improve this bound to 600. In fact, he proved that for each m there exist a bound B_m (depending only on m) such that there are infinitely many intervals of length B_m that contain at least m primes. In other words, that

$$\liminf_{n\to\infty}(p_{n+m-1}-p_n) \leq B_m.$$

If Zhang's method is combined with the Maynard's, then that the bound can be further reduced to 200. If all of these techniques could be pushed to their limit then we would obtain $B(=B_1)=6$, (as the Tao's Polymath project has shown). New ideas are still needed to have a feasible plan for proving the twin prime conjecture.

9.2 Outline of the method

The basic idea goes back to Hardy and Littlewood in the early 1900's.

Definition. *The finite set $\mathcal{H}$ of non-negative integers is* admissible *if they do not cover all the residue classes modulo any prime. That is, for each prime p there exists $a_p \in \mathbb{Z}$ such that $a_p \not\equiv h (\mathrm{mod}\ p)$ for any $h \in \mathcal{H}$.*

Fix an admissible set $\mathcal{H} = \{h_1, \ldots, h_k\}$.

Conjecture 9.2 (Hardy–Littlewood)**.** *There are infinitely many positive integers n for which $n+h_1, \ldots, n+h_k$ are all prime.*

In fact, Hardy and Littlewood not only conjectured that there are infinitely many such n's, but they gave a guess for their frequency.

Remark 9.3. The twin prime conjecture is a special case of Conjecture 9.2. Indeed it would follow if $\mathcal{H} = \{0, 2\}$.

How can one attack such a question? One could take $\mathcal{P}$ to the be set of all primes and consider its indicator functiion

$$\chi(n) = \chi_{\mathcal{P}}(n) = \begin{cases} 1 & n \text{ is prime;} \\ 0 & \text{otherwise.} \end{cases}$$

Then one considers

$$\chi(n+h_1)\chi(n+h_2)\ldots\chi(n+h_k).$$

If we could prove that this product is nonzero infinitely often, Conjecture 9.2 would follow. But unfortunately the product is not understood well enough to permit this. Instead, it has been profitable to look at

$$\sum_{N<n<2N} \chi(n+h_1)\chi(n+h_2)\dots\chi(n+h_k)$$

and tries to prove that it grows without bound as $N \to \infty$. In fact, Hardy and Littlewood conjectured an asymptotic formula for this function.

The basic idea of the GPY method is to consider the sum

$$S(N,\rho) = \sum_{N\le n<2N} \Big(\sum_{i=1}^{k} \chi(n+h_i) - \rho\Big) w_n. \tag{9.1}$$

where $\rho > 0$ and w_n are non-negative weights. If we can show that $S(N,\rho) > 0$ then at least one term in the sum over n must have a positive contribution. By the non-negativity of w_n, this means that there must be some integer $n \in [N, 2N]$ such that at least $\lfloor \rho+1 \rfloor$ of the $n+h_i$ are prime. Thus if $S(N,\rho) > 0$ for all large N, there are infinitely many integers n for which at least $\lfloor \rho+1 \rfloor$ of the $n+h_i$ are prime (and so there are infinitely many bounded length intervals containing $\lfloor \rho+1 \rfloor$ primes).

The weights w_n are typically chosen to mimic Selberg sieve weights. Estimating (9.1) can be interpreted as a 'k-dimensional' sieve problem. The standard Selberg k-dimensional weights (which can be shown to be essentially optimal in other contexts) are

$$w_n = \Big(\sum_{\substack{d|\prod_{i=1}^{k}(n+h_i)\\ d<z}} \lambda_d\Big)^2, \qquad \lambda_d = \mu(d)(\log z/d)^k. \tag{9.2}$$

With this choice we find that we just fail to prove the existence of bounded gaps between primes if we assume the Elliott-Halberstam conjecture. The key new idea in the paper of Goldston, Pintz and Yıldırım was to consider more general sieve weights of the form

$$\lambda_d = \mu(d)F(\log z/d), \tag{9.3}$$

for a suitable smooth function F. Goldston, Pintz and Yıldırım chose $F(x) = x^{k+l}$ for suitable $l \in \mathbb{N}$, which has been shown to be essentially optimal when k is large. This allows us to gain a factor of approximately 2 for large k over the previous choice of sieve weights.

As a result we just fail to prove bounded gaps because they have to let the length $k \to \infty$. But it gives us enough control to prove the result with logarithms in the denominator. For GPY, this improvement in the weights is the difference between complete failure (as Selberg himself and many others had failed) to and success.

The GPY method relies heavily on the distribution of primes in arithmetic progressions. Given $\theta > 0$, we say the primes have *level of distribution* θ if, for any $A > 0$, we have

$$\sum_{q\le x^\theta} \max_{(a,q)=1} \Big|\pi(x;q,a) - \frac{\pi(x)}{\phi(q)}\Big| \ll_A \frac{x}{(\log x)^A}. \tag{9.4}$$

The Bombieri-Vinogradov theorem (which is based on the *large sieve*) establishes that the primes have level of distribution θ for any $\theta < 1/2$, and Elliott and Halberstam conjectured that this could be extended to any $\theta < 1$.

The original work of Goldston, Pintz and Yıldırım showed the existence of bounded gaps between primes if (9.4) holds for some $\theta > 1/2$. Moreover, under the Elliott-Halberstam conjecture one had $\liminf_n (p_{n+1} - p_n) \leq 16$. The key breakthrough of Zhang's work was in establishing a suitably weakened form of (9.4) holds for some $\theta > 1/2$.

If one looks for bounded length intervals containing two or more primes, then the GPY method fails to prove such strong results. Unconditionally we are only able to improve upon the trivial bound from the prime number theorem by a constant factor, and even assuming the Elliott-Halberstam conjecture, the best available result is

$$\liminf_n \frac{p_{n+2} - p_n}{\log p_n} = 0. \tag{9.5}$$

The aim of this paper is to introduce a refinement of the GPY method which removes the barrier of $\theta = 1/2$ to establishing bounded gaps between primes, and allows us to show the existence of arbitrarily many primes in bounded length intervals. Maynard's new method also has the benefit that it produces numerically superior results to previous approaches.

As a result we just fail to prove bounded gaps using the fact that the primes have exponent of distribution θ for any $\theta < 1/2$, but succeed in doing so if we assume they have level of distribution $\theta > 1/2$.

The new ingredient in Maynard's method is to consider a more general form of the sieve weights

$$w_n = \Big(\sum_{d_i | n + h_i \forall i} \lambda_{d_1, \ldots, d_k} \Big)^2. \tag{9.6}$$

This allows us to improve on the previous choice of sieve weights by an arbitrarily large factor, provided that k is sufficiently large. It is the extra flexibility gained by allowing the weights to depend on the divisors of each factor individually which gives this improvement. It is true that it leads to a harder optimization problem than the ones of from GPY (we have seen a baby version in Corollary 8.8), but it is one that Maynard can still solve.

As a result, he manages to prove there are infinitely many n such that the k-tuple $(n + h_1, \ldots, n + h_{105})$ contains at least two primes for $h_1, \ldots, h_{1-5}$ the elements of the admissible set

$$\begin{aligned}\mathcal{H} = \{&0, 10, 12, 24, 28, 30, 34, 42, 48, 52, 54, 64, 70, 72, 78, 82, 90, 94, 100, 112, 114, 118, 120, 124,\\ &132, 138, 148, 154, 168, 174, 178, 180, 184, 190, 192, 202, 204, 208, 220, 222, 232, 234, 250,\\ &252, 258, 262, 264, 268, 280, 288, 294, 300, 310, 322, 324, 328, 330, 334, 342, 352, 358, 360,\\ &364, 372, 378, 384, 390, 394, 400, 402, 408, 412, 418, 420, 430, 432, 442, 444, 450, 454, 462,\\ &468, 472, 478, 484, 490, 492, 498, 504, 510, 528, 532, 534, 538, 544, 558, 562, 570, 574, 580,\\ &582, 588, 594, 598, 600\}.\end{aligned}$$

This implies that

$$\liminf(p_{n+1} - p_n) \leq 600.$$

If he assumes the Elliott-Halberstam conjecture, the he can reduce the gap proving the same result with $\mathcal{H} = \{0, 2, 6, 8, 12\}$. This implies that

$$\liminf(p_{n+1} - p_n) \leq 12.$$

The interesting fact here is that Maynard does not make use of Zhang's results at all. Combining Zhang's method with Maynard's, and improving some of their estimates, a group of mathematicians led by Terence Tao proved that there are infinitely many primes that are situated at distance ≤ 246. This is an unconditional result, that does not assume any unproven conjectures.