# Combinatorial and Analytic Number Theory

Course fall 2007
R. Tijdeman

December 21, 2007

Introduction. This is a new course, however, with some chapters from other courses and some new material. It provides an introduction to combinatorial and analytic number theory giving a survey of the most important results in this area and the most successful methods.

The course will be on Thursdays from 11.15 AM to 1.00 PM, with the exception of 19th and 26th October. Lecture notes will be made by Tünde Kovács, the first chapters after the lectures, later when it becomes more difficult before the lectures. There are exercises, partly as homework. This homework has to made every second week and forms part of the exam. The other part is an oral where I ask the solutions of some exercises and to explain details which are not written out to check whether you have really understood the mathematics. The renumeration is 6 ECTS.

The contents of the course is not fixed yet, but I can give a good impression of the contents.

Chapter 1. Residue classes.
Chapter 2. Sums of squares with the theorem that every positive integer can be written as the sum of four squares.
Chapter 3. The weak prime number theorem.
Chapter 4. Multiplicative functions and Dirichlet series. E.g. Euler products.
Chapter 5. Primes in arithmetic progressions (AP) with the proof that every AP $(a + nd)_{n=1}^{\infty}$ with $\gcd(a, d) = 1$ has infinitely many primes.
Chapter 6. Sieve methods with an upper bound for the number of primes in an AP.
Chapter 7. The circle method with the proof that a sequence without an AP of length 3 has zero density.
Chapter 8. Smooth numbers

Here are some elementary notions and results which we assume without further mention.

The fundamental theorem of arithmetic: Every given natural number can be written uniquely as the product of prime numbers (up to order).

To be more precise, any given natural number $n$ can be represented in the form

$$n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

where $p_1 < p_2 < \ldots < p_r$ are primes and $k_1, k_2, \ldots, k_r$ are positive integers. (In the case $n = 1$ we have $r = 0$.) We shall refer to this representation as writing $n$ in standard form.

Congruences. Let $m \in \mathbf{N}$, $r \in \mathbf{Z}$ such that $0 \leq r < m$. The residue class $r$ modulo $m$ consists of all integers which divide by $m$ have remainder $r$. Let $a$ be an element of the residue class $r$ modulo $m$, then we use the notation $\overline{a}$ for $a + m\mathbf{Z}$ which contains all integers $a \pmod{m}$. Hence $\overline{a} = \overline{b} \Longleftrightarrow m|(a-b)$. We can define the sum and the product of residue classes:

$$\overline{a} + \overline{b} = \overline{a+b}, \tag{1}$$

$$\overline{a} \cdot \overline{b} = \overline{ab}. \tag{2}$$

Let $m \in \mathbf{N}$, $a, b \in \mathbf{Z}$ such that $(a, m) = 1$. Then the equation $\overline{a}x = \overline{b}$ has a unique solution modulo $m$. Let $m \in \mathbf{N}$, $a, b \in \mathbf{Z}$ such that $(a, m) = d$. Then the equation $\overline{a}x = \overline{b}$ has no solution when $d \not| ib$ and there are exactly $d$ solutions when $d|b$.

If we have to deal with a system of linear congruences in a single unknown, each taken to a different modulus, we can use the Chinese Remainder Theorem.

Chinese Remainder Theorem: If $m_1, \ldots, m_r \in \mathbf{N}$ are given moduli relatively prime in pairs, then the system of linear congruences $x \equiv c_1 \pmod{m_1}$, $x \equiv c_2 \pmod{m_2}$,…, $x \equiv c_r \pmod{m_r}$ where $c_1, \ldots, c_r \in \mathbf{Z}$ are given remainders, has a unique solution modulo $m_1 \cdots m_r$.

Let $a \in \mathbf{Z}$, $m \in \mathbf{N}$, $m > 1$ and $(a, m) = 1$. Then the residue class $a$ modulo $m$ is called a prime residue class. Note that the product of two prime residue classes mod $m$ is again a prime residue class mod $m$. We call the uniquely determined residue class $c \pmod{m}$ for which $ac \equiv 1 \pmod{m}$ holds, the inverse residue class of the prime residue class $a \pmod{m}$ and we use the notation $a^{-1} \pmod{m}$. We have $\varphi(m)$ distinct prime residue classes mod $m$ where $\varphi$ is Euler‘s $\varphi$-function. Furthermore, we have:

$$\text{if } (m, n) = 1 \text{ then } \varphi(m, n) = \varphi(m) \cdot \varphi(n), \tag{3}$$

$$\varphi(n) = n \cdot \prod_{p|n,\ p\ prime} \left(1 - \frac{1}{p}\right), \tag{4}$$

$$\sum_{d|n} \varphi(d) = n. \tag{5}$$

Euler's Theorem: Let $a \in \mathbf{Z}$, $m \in \mathbf{N}$ such that $(a, m) = 1$. Then $a^{\varphi(m)} \equiv 1 \pmod{m}$.
If $m$ is a prime, denoted by $p$, then Fermat's Theorem says: Let $a \in \mathbf{Z}$, $p$ be a prime such that $p \not| a$. Then $a^{p-1} \equiv 1 \pmod{p}$.

# Chapter 1

# Residue classes

Literature:
F. Beukers, Getaltheorie voor Beginners, Epsilon Uitgaven, Utrecht, 1998.
K. Chandrasekharan, Introduction to Analytic Number Theory, Springer-Verlag, 1968.
G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers, Oxford at the Clarendom Press, 5th edition, 1979.
L.K. Hua, Introduction to Number Theory, Springer-Verlag, 1982.

**Theorem 1.1.** (Lagrange, 1768) *Let $p$ be a prime, $a_n \in \mathbb{Z}$, $p \not| a_n$. Then the congruence*

$$a_n x^n + \ldots + a_0 \equiv 0 \pmod{p}$$

*has at most $n$ solutions* {mod $p$.

**Proof.** We use induction on $n$. For $n = 1$ we have $a_1 x + a_0 \equiv 0 \pmod{p}$. The fact that $p \not| a_1$ and $a_1 x =^p -a_0$ implies that $x =^p -a_1^{-1} a_0$. So the solution is exactly one residue class. Assume that the theorem holds up to a certain degree $n-1$. Let $x_1$ be a solution. Then $a_n x_1^n + \ldots + a_0 =^p 0$. From this we get $a_n \left(x^n - x_1^n\right) + a_{n-1}\left(x^{n-1} - x_1^{n-1}\right) + \ldots + a_1(x - x_1) =^p 0$. This implies

$$(x - x_1)\left\{a_n\left(x^{n-1} + x^{n-2}x_1 + \ldots + x_1^{n-1}\right) + \ldots + a_1\right\} =^p 0.$$

We get a polynomial with degree $n-1$ and leading coefficient $a_n$, so we can use the induction assumption for this: it can have at most $n-1$ solutions incongruent mod $p$. Recall that $p|ab \Longrightarrow p|a$ or $p|b$. Thus, the solution of our congruence is a solution of $x - x_1 =^p 0$ or $\left\{a_n\left(x^{n-1} + x^{n-2}x_1 + \ldots + x_1^{n-1}\right) + \ldots + a_1\right\} =^p 0$. The first congruence has only one solution, the second one has at most $n-1$ solutions, thus we have at most $n$ solutions. □

**Theorem 1.2.** *Let $p$ be a prime, $d|(p-1)$, $d \in \mathbb{N}$. Then $x^d =^p 1$ has exactly $d$ incongruent solutions mod $p$.*

**Proof.** From Fermat‘s Theorem the congruence $x^{p-1} =^p 1$ has $p-1$ prime residue classes as solutions mod $p$.
Let $p - 1 = k \cdot d$. Then we have

$$0 =^p x^{p-1} - 1 = \left(x^d - 1\right)\left(x^{(k-1)d} + x^{(k-2)d} + \ldots + 1\right).$$

From Theorem 1.1. we know that the first factor on the right-hand side has at most $d$ incongruent solutions and the second factor has at most $(k-1)d$ solutions. Since there are $p-1$ solutions, the first factor has exactly $d$ solutions. □

We call $g$ a primitive root of the positive integer $m$ if $g$ and $m$ are coprime and $g, g^2, \ldots, g^{\varphi(m)}$ are all distinct. It means that $g$ generates all the primitive residue classes of $m$. If $p$ is a prime and $g$ is a primitive root of $p$, then $\{0, g, g^2, \ldots, g^{p-1}\}$ represents the full set of residue classes of $p$.

**Theorem 1.3.** *Let $p$ be a prime, $d \in \mathbb{N}$, $d|(p-1)$. Then there are exactly $\varphi(d)$ distinct residue classes of $p$ of order $d$, and there are exactly $\varphi(p-1)$ primitive roots of $p$.*

**Proof.** We use induction on $d$. Let $d = 1$. The congruence $x^1 =^p 1$ has one solution, namely $\overline{1}$. Let us assume that the theorem is correct for all $d'|(p-1)$ with $1 \leq d' < d$. By Theorem 1.2. we know that $x^d =^p 1$ has $d$ solutions. The order of the residue classes having order less than $d$, is a divisor $d'$ of $d$. According to the induction hypothesis the number of these residue classes is $\sum_{d'|d,d'<d} \varphi(d')$. Then we get $d - \sum_{d'|d,d'<d} \varphi(d')$ solutions of order $d$. By (5) this value equals $\varphi(d)$. The second assertion follows from the first with $d = p-1$. □

**Definition.** Let $p$ be a prime number with $p \not| a$. Then $a$ is said to be a <u>quadratic residue</u> modulo $p$ if there exists some integer $x$ such that $x^2 \equiv a \pmod{p}$. We call $a$ <u>quadratic non-residue</u> modulo $p$ if the congruence $x^2 \equiv b \pmod{p}$ has no solutions.

**Theorem 1.4.** *Let $p$ be an odd prime. Then there exist exactly $\frac{p-1}{2}$ quadratic residues and $\frac{p-1}{2}$ quadratic non-residues modulo $p$.*

**Proof.** Consider the residue classes $1^2$, $2^2$,…,$\left(\frac{p-1}{2}\right)^2$ mod $p$. As $a^2 \equiv (-a)^2$ mod $p$, there cannot be any other quadratic residues. Thus there exist at most $\frac{p-1}{2}$ quadratic residue classes. Let $g$ be a primitive root of $p$. Then $g$, $g^2$, $g^3$,…,$g^{p-1}$ are all distinct mod $p$, and $g^2 = (g)^2$, $g^4 = \left(g^2\right)^2$,…,$g^{p-1} = \left(g^{(p-1)/2}\right)^2$ are all quadratic residue classes. Thus there exist at least $\frac{p-1}{2}$ quadratic residue classes. So we have exactly $\frac{p-1}{2}$ quadratic residues and
$p-1-\frac{p-1}{2} = \frac{p-1}{2}$ quadratic non-residues. □

**Theorem 1.5.** *Let $p$ be a prime, $p > 2$ and $p \not| a$. Then*

$$\begin{aligned} &a \text{ is a quadratic residue mod } p \iff a^{\frac{p-1}{2}} \equiv 1 \ (\text{mod } p) \\ &a \text{ is a quadratic non-residue mod } p \iff a^{\frac{p-1}{2}} \equiv -1 \ (\text{mod } p). \end{aligned}$$

**Proof.** By Euler's Theorem $\left(a^{\frac{p-1}{2}}\right)^2 \equiv a^{p-1} \equiv 1 \pmod{p}$. Thus $a^{\frac{p-1}{2}}$ is a solution of the congruence $x^2 \equiv 1 \pmod{p}$ and then $a^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p}$. Assume that $a$ is a quadratic residue. Then $a \equiv b^2 \pmod{p}$, which implies that $a^{\frac{p-1}{2}} \equiv b^{p-1} \equiv 1 \pmod{p}$. By Theorem 1.2 it follows that the congruence $x^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ has exactly $\frac{p-1}{2}$ solutions. These solutions are exactly the quadratic residues. Then for a quadratic non-residue $a$, we have $a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$. □

**Definition.** Let $p$ be a prime, $p > 2$. The Legendre symbol can be defined as follows:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 \;\; if \;\; a \;\; is \;\; a \;\; quadratic \;\; residue \;\; mod \;\; p, \\ -1 \;\; if \;\; a \;\; is \;\; a \;\; quadratic \;\; non-residue \;\; mod \;\; p, \\ 0 \;\; if \;\; p|a. \end{cases}$$

**Corollary.** (Euler) For a prime $p$, $p > 2$ we have $\left(\frac{a}{p}\right) =^p a^{\frac{p-1}{2}}$.

**Theorem 1.6.** *For an odd prime $p$ and $a, b, k \in \mathbb{Z}$ the following holds:*

$$\left(\tfrac{a}{p}\right) \cdot \left(\tfrac{b}{p}\right) = \left(\tfrac{ab}{p}\right), \;\; \left(\tfrac{a+kp}{p}\right) = \left(\tfrac{a}{p}\right).$$

**Proof.** Exercise!

**Theorem 1.7.** *Let $p$ be an odd prime. Then we have*

$$\left(\tfrac{1}{p}\right) = 1, \;\; \left(\tfrac{-1}{p}\right) = \begin{cases} 1 \;\; if \;\; p \equiv 1 \;\; (mod \;\; 4), \\ -1 \;\; if \;\; p \equiv 3 \;\; (mod \;\; 4). \end{cases} \;\; .$$

**Proof.** The first assertion is trivial. The second assertion follows from

$$\left(\frac{-1}{p}\right) =^p (-1)^{\frac{p-1}{2}} = \begin{cases} 1 \;\; if \;\; p \equiv 1 \;\; (mod \;\; 4), \\ -1 \;\; if \;\; p \equiv 3 \;\; (mod \;\; 4). \end{cases}$$

**Theorem 1.8.** *Let $p$ be an odd prime. Then we have*

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

**Proof.** See Hardy and Wright, Sect. 6.11.

**Theorem 1.9.** (The quadratic reciprocity law.) *Let $p$ and $q$ be distinct odd primes. Then*

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

**Proof.** See Hardy and Wright, Sect. 6.13.

The above theorems enable you to compute Legendre symbols quite fast and hence to determine whether the congruence equation $x^2 \equiv a(\mathrm{mod} p)$ is solvable. E.g.

$$\left(\frac{87}{127}\right) = \left(\frac{3}{127}\right)\left(\frac{29}{127}\right) = -\left(\frac{127}{3}\right)\left(\frac{127}{29}\right) =$$

$$-\left(\frac{1}{3}\right)\left(\frac{11}{29}\right) = -\left(\frac{29}{11}\right) = -\left(\frac{7}{11}\right) = \left(\frac{11}{7}\right) = \left(\frac{4}{7}\right) = 1.$$

## 1.1 Homework for Chapter 1

1. a) Prove $(p-1)! \equiv -1 \pmod{p}$ if and only if $p$ is a prime.
   b) Compute $(p-1)! \pmod{p}$ if $p$ is not prime.

2. Prove
   a) If $g$ is a primitive root of the prime $p$, then the other primitive roots are given by $g^r$ with $1 \leq r < p$, $\gcd(r, p-1) = 1$.
   b) If $m$ has primitive roots, then $m$ has exactly $\varphi(\varphi(m))$ primitive roots.
   c)If $m$ has a primitive root, then $m = 1, 2, 4$ or it is of the form $p^k$ or $2p^k$ with $p$ an odd prime.
   d) All numbers mentioned under c) have primitive roots.

3. Let $p$ be on odd prime. Prove $x^4 \equiv -1 \pmod{p}$ is solvable $\Longleftrightarrow p \equiv 1 \pmod{8}$.

4. Prove that there are infinitely many primes $p$ with $p \equiv 3 \pmod{4}$.

These exercises have to be done for Thursday 27 September.

## 1.2 Further exercises for Chapter 1

1. Let $n \geq 2$ such that $2^n + 1$ is prime. Prove that $n = 2^k$ for some positive integer $k$.

2. Prove: for all positive integer $n$ with $a \neq 1$ we have $n | \varphi(a^n - 1)$.

3. Determine all the quadratic residue classes of 17 and of 19.

4. For which primes is 5 a quadratic residue?

5. a) Let $p$ be an odd prime and $k$ an integer $\geq 2$. Prove that
   $x^2 \equiv d \pmod{p}$ is solvable if and only if $x^2 \equiv d \pmod{p^k}$ is solvable.
   b) Show that the assertion is wrong for $p = 2$.

6. Prove: $p \equiv 3 \pmod{4} \Longleftrightarrow \frac{N-K}{p}$ is odd. (K is the sum of quadratic residues, N is the sum of quadratic non-residues.)

7. Compute $\left(\frac{66}{127}\right)$.

# Chapter 2

# Sums of squares

In this chapter we answer the question which integers can be written as the sum of two, three, four squares. There is a whole theory on in how many ways a number can be represented as the sum of a given number of squares. Here only some elementary proofs are given which are a few hundred years old and proved by the principle of infinite descent.

Note that $2 = 1^2 + 1^2$ and no prime $p \equiv 3 \pmod 4$ can be written as sum of two squares, since every square is $\equiv 0, 1 \pmod 4$.

**Theorem 2.1.** (Euler) *Every prime $p \equiv 1$ (mod 4) can be written as the sum of two squares and in only one way. (Here we consider $(\pm a)^2 + (\pm b)^2$ and $(\pm b)^2 + (\pm a)^2$ as the same representation.)*

**Proof.** Consider the congruence $x^2 \equiv -1 \pmod p$. By Theorem 1.7 we know that this is solvable if $p \equiv 1 \pmod 4$. Choose a solution $x_0$ with $-\frac{p}{2} < x_0 \le \frac{p}{2}$. Thus $x_0^2 + 1 = lp$ with $l < \left(\frac{p^2}{4} + 1\right) \cdot \frac{1}{p} < \frac{1}{2}p < p$.
Let $m$ be the smallest natural number such that $mp$ can be written as the sum of two squares, $mp = x_1^2 + y_1^2$. If $m = 1$ then we are ready. Assume that $m > 1$. We have $1 < m \le l < p$. Choose $x_2$ and $y_2$ from the interval $\left(-\frac{m}{2}, \frac{m}{2}\right]$ such that $x_2 \equiv x_1 \pmod m$, $y_2 \equiv y_1 \pmod m$. Then it follows that $x_2^2 + y_2^2 \equiv x_1^2 + y_1^2 \equiv mp \equiv 0 \pmod m$. Thus if $x_2^2 + y_2^2 = rm$, then $r \le \left(\frac{m^2}{4} + \frac{m^2}{4}\right) \cdot \frac{1}{m} = \frac{m}{2} < m$. Assume $r = 0$, then $x_2 = y_2 = 0$, thus $m | x_1$, $m | y_1$ and $m^2 | x_1^2 + y_1^2 = mp$ from this it follows that $m | p$ in contradiction with $1 < m < p$. Thus $1 \le r < m$. We know that $x_1y_2 - x_2y_1 \equiv x_1y_1 - x_1y_1 \equiv 0$ (mod $m$) and $x_1x_2 + y_1y_2 \equiv x_1^2 + y_1^2 \equiv 0 \pmod m$. Define $x_3 = (x_1x_2 + y_1y_2)/m$, $y_3 = (x_1y_2 - x_2y_1)/m$. Then it follows that $x_3, y_3 \in \mathbb{Z}$ and

$$x_3^2 + y_3^2 = \frac{(x_1x_2 + y_1y_2)^2 + (x_1y_2 - x_2y_1)^2}{m^2} = \frac{(x_1^2 + y_1^2)(x_2^2 + y_2^2)}{m^2} = rp.$$

Since $1 \le r < m$, it is a contradiction with the definition of $m$. Via <u>infinite descent</u> we showed that $p$ is representable as the sum of two squares.

Assume that $p$ can be written as the sum of two squares in two distinct ways: $x^2 + y^2$ and $X^2 + Y^2$. We know that $(x, y) = 1$ and $(X, Y) = 1$. Choose $y_4$ and $Y_4$ such that $yy_4 \equiv 1 \pmod p$ and $YY_4 \equiv 1 \pmod p$. Then it follows that $(xy_4)^2+1 \equiv x^2y_4^2+y^2y_4^2=(x^2+y^2)y_4^2 \equiv 0 \pmod{p}$ and analogously $(XY_4)^2+1 \equiv 0 \pmod{p}$. Thus $xy_4$ and $XY_4$ are solutions of the congruence $z^2+1 \equiv 0 \pmod{p}$. By Theorem 1.1 we know that it has two solutions, $\pm z_0$ say. Thus $xy_4 \equiv \pm XY_4 \pmod{p}$. This gives that $xY \equiv \pm Xy \pmod{p}$. We get that $p^2=(x^2+y^2)(X^2+Y^2)=(xX \pm yY)^2+(xY \mp Xy)^2 \equiv (xX \pm yY)^2 \pmod{p}$. Thus $p|xX \pm yY$. Let $x_5=\frac{xX \pm yY}{p}$, $y_5=\frac{xY \mp yX}{p}$. Then $x_5, y_5 \in \mathbb{Z}$ and $x_5^2+y_5^2=\frac{p^2}{p^2}=1$. Therefore we get $x_5=0$ or $y_5=0$. Thus $xX=\mp yY$ or $xY=\pm yX$ and as $(x,y)=1$ and $(X,Y)=1$, it follows that $|x|=|Y|$, $|y|=|X|$ or $|x|=|X|$, $|y|=|Y|$. Therefore there cannot exist two distinct representations of $p$ as sum of two squares. □

**Theorem 2.2.** *A natural number $n$ is representable as the sum of two squares if and only if each prime factor of $n$ of the form $4k+3$ occurs to an even power in the prime factorization of $n$.*

(For example $n=2^3 \cdot 3^2 \cdot 11 \cdot 13^2$ is not representable as the sum of two squares, but $n=2^3 \cdot 3^2 \cdot 11^2 \cdot 13$ is.)

**Proof.** If $m$ and $n$ are representable as the sum of two squares, then their product $mn$ is representable, too. Let $m=a^2+b^2$, $n=c^2+d^2$, then $mn=(ac+bd)^2+(ad-bc)^2$.
$\Longleftarrow$ Let us assume that $n$ is of the form $2^k p_1^{k_1} \cdots p_r^{k_r} q_1^{2l_1} \cdots q_s^{2l_s}$ with $p_j \equiv 1 \pmod{4}$ and $q_j \equiv 3 \pmod{4}$ are all primes. The factors $2, p_1, \ldots, p_r, q_1^2, \ldots, q_s^2$ can be written as the sum of two squares. By the statement at the beginning of the proof, we are ready with this direction.
$\Longrightarrow$ Assume that $n=a^2+b^2$. Let $(a,b)=d$, $a_1=a/d$, $b_1=b/d$, $n_1=n/d^2$. Then we have $n_1=a_1^2+b_1^2$ with $(a_1,b_1)=1$. Assume that $q$ is a prime number $\equiv 3 \pmod{4}$ such that it appears to an odd power in the prime factorization of $n$. Then we have $q|n_1$, but $q \not| a_1$ and $q \not| b_1$. Since $-b_1^2 \equiv a_1^2 \pmod{q}$, we get $1=\left(\frac{-b_1^2}{q}\right)=\left(\frac{-1}{q}\right)\left(\frac{b_1^2}{q}\right)=\left(\frac{-1}{q}\right)$. From this it follows that $q \equiv 1 \pmod{4}$, which is a contradiction. □

The problem of the solvability of the equation $n=x^2+y^2+z^2$ is much more difficult than the corresponding question for the sum of either two or four squares. We have the following result.

**Theorem 2.3.** *A natural number $n$ can be written as the sum of three squares if and only if $n$ is not of the form $4^l(8k+7)$.*

**Proof.** $\Longrightarrow$ Assume $n=4^l(8k+7)$. We use induction on $l$. For $l=0$ we observe that $8k+7$ is not representable as the sum of three squares because a square is $\equiv 0, 1$ or $4 \pmod{8}$. Assume that the theorem is correct for $0, 1, \ldots, l-1$ and $n=4^l(8k+7)=a^2+b^2+c^2$. If $a, b, c$ are all even then $\frac{n}{4}=4^{l-1}(8k+7)=\left(\frac{a}{2}\right)^2+\left(\frac{b}{2}\right)^2+\left(\frac{c}{2}\right)^2$, which is a contradiction. Otherwise exactly one of the numbers $a, b, c$ is even and the other two are odd. From this it follows that $a^2+b^2+c^2 \equiv 0+1+1 \pmod{4}$. However $4|n$, which leads to a contradiction.
$\Longleftarrow$ See e.g. Landau, Vorlesungen I, Satz 187, pp. 114-125. □

**Theorem 2.4.** (Girard, Fermat, Lagrange) *Every natural number is representable as the sum of four squares.*

**Proof.** Since $1 = 1^2 + 0^2 + 0^2 + 0^2$ and $2 = 1^2 + 1^2 + 0^2 + 0^2$, we assume $n \geq 3$.

If $m$ and $n$ are representable as the sum of four squares, then their product $mn$ is representable, too. Let $m = x_1^2 + x_2^2 + x_3^2 + x_4^2$, $n = y_1^2 + y_2^2 + y_3^2 + y_4^2$, then $mn = (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + (x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2 + (x_1y_3 - x_3y_1 + x_4y_2 - x_2y_4)^2 + (x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2$. Thus we only need to prove the theorem for prime numbers.

Let $p \equiv 1 \pmod 4$. By Theorem 2.1 we know that $p$ is of the form $a^2 + b^2 + 0^2 + 0^2$. For prime numbers $p \equiv 3 \pmod 4$ we use the method of infinite descent. Let $z$ be the smallest positive quadratic non-residue of $p$. Then $z \geq 2$ and $z - 1$ is a quadratic residue. Further, $\left(\frac{-z}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{z}{p}\right) = -1 \cdot -1 = 1$, thus $-z$ is a quadratic residue of $p$. Let $-z \equiv x^2 \pmod p$ and $z - 1 \equiv y^2 \pmod p$. Then it follows that $x^2 + y^2 + 1 \equiv -z + (z-1) + 1 \equiv 0 \pmod p$. Choose $x_0, y_0$ such that $x_0 \equiv x \pmod p$ and $y_0 \equiv y \pmod p$ with $|x_0| < \frac{p}{2}$ and $|y_0| < \frac{p}{2}$. Then $x_0^2 + y_0^2 + 1 = m_0p$ with $1 \leq m_0p < \frac{p^2}{4} + \frac{p^2}{4} + 1 < \frac{3p^2}{4} < p^2$ and thus $1 \leq m_0 < p$.

Let $m$ be the smallest natural number such that $mp$ can be written as the sum of four squares. If $m = 1$ then we are ready. Assume that $m > 1$ and $mp = a^2 + b^2 + c^2 + d^2$. Choose $A, B, C, D$ from the interval $\left(-\frac{m}{2}, \frac{m}{2}\right]$ such that $a \equiv A \pmod m$, $b \equiv B \pmod m$, $c \equiv C \pmod m$, $d \equiv D \pmod m$. Then we have $A^2 + B^2 + C^2 + D^2 \equiv a^2 + b^2 + c^2 + d^2 \equiv 0 \pmod m$. Thus $A^2 + B^2 + C^2 + D^2 = rm$ for some $r \in \mathbb{Z}$ with $0 \leq r \leq m$. Assume $r = 0$, then we get $A = B = C = D = 0$ and $m^2 | a^2 + b^2 + c^2 + d^2 = mp$, therefore, $m|p$ in contradiction with $1 < m \leq m_0 < p$.

Assume $r = m$, then it follows that $A = B = C = D = m/2$ and $a \equiv b \equiv c \equiv d \equiv \frac{m}{2} \pmod m$, and therefore $a^2 + b^2 + c^2 + d^2 \equiv \frac{m^2}{4} + \frac{m^2}{4} + \frac{m^2}{4} + \frac{m^2}{4} \equiv m^2 \equiv 0 \pmod{m^2}$. Thus, we get $m^2 | mp$, and then $m|p$, in contradiction with $1 < m < p$. We conclude that $1 \leq r < m$.

We can use the property mentioned in the beginning of the proof as follows $(a^2 + b^2 + c^2 + d^2)(A^2 + B^2 + C^2 + D^2) = \alpha^2 + \beta^2 + \gamma^2 + \delta^2$ with

$$\begin{aligned}
\alpha &:= aA + bB + cC + dD \equiv a^2 + b^2 + c^2 + d^2 \equiv 0 \pmod m, \\
\beta &:= aB - bA + cD - dC \equiv ab - ab + cd - cd \equiv 0 \pmod m, \\
\gamma &:= aC - cA + dB - bD \equiv ac - ac + bd - bd \equiv 0 \pmod m, \\
\delta &:= aD - dA + bC - cB \equiv ad - ad + bc - bc \equiv 0 \pmod m.
\end{aligned}$$

Therefore, $\alpha/m$, $\beta/m$, $\gamma/m$ and $\delta/m$ are integers and

$$\left(\frac{\alpha}{m}\right)^2 + \left(\frac{\beta}{m}\right)^2 + \left(\frac{\gamma}{m}\right)^2 + \left(\frac{\delta}{m}\right)^2 = \frac{1}{m^2}(a^2 + b^2 + c^2 + d^2)(A^2 + B^2 + C^2 + D^2) = rp.$$

Thus $rp$ can be written as the sum of four squares, and since $1 \leq r < m$, it is in contradiction with the definition of $m$. □

**Problem of Waring.** We have shown that every positive integer can be written as the sum four squares. Around 1770 the English mathematician Waring posed the following question nowadays known as Waring's problem. Is there for every positive integer $k$ an integer $g(k)$ such that every positive integer can be written as the sum of $g(k)$ $k$th powers of nonnegative integers?

Hilbert showed in 1909 that for every positive integer $k$ there is such a constant $g(k)$. It is now known that, if $g(k)$ denotes the smallest such number, $g(2) = 4$, $g(3) = 9$, $g(4) = 19$, $g(5) = 37$ and, in general, $g(k) = \left[(3/2)^k\right] + 2^k - 2$ for $k \geq 3$.

**Pythagorean triples.** According to the Pythagorean theorem to find all right triangles with integral side lengths, we need to find all triples of positive integers $x, y, z$ satisfying the Diophantine equation $x^2 + y^2 = z^2$. Triples of positive integers satisfying this equation are called Pythagorean triples. A Pythagorean triple $x, y, z$ is called primitive if $(x, y, z) = 1$.

**Theorem 2.5.** *All primitive solutions of $a^2 + b^2 = c^2$ in natural numbers $a, b, c$ with $b$ even are given by*

$$a = r^2 - s^2, \;\; b = 2rs, \;\; c = r^2 + s^2$$

$$\text{with } r > s > 0, \;\; (r, s) = 1 \;\text{ and } 2 \not| (r - s).$$

*and, conversely, every such a pair $(r, s)$ generates a primitive Pythagorean triple $a, b, c$.*

(The smallest values for the parameters $(r, s)$ are $(2, 1)$, $(3, 2)$, $(4, 1)$, $(4, 3)$, $(5, 2)$, $(5, 4)$ which are related to the equations $3^2 + 4^2 = 5^2$, $5^2 + 12^2 = 13^2$, $15^2 + 8^2 = 17^2$, $7^2 + 24^2 = 25^2$, $21^2 + 20^2 = 29^2$, $9^2 + 40^2 = 41^2$.)

**Proof.** Since $a$ is odd, $b$ is even, $c$ must be odd and $2 | c \pm a$. We have $b^2 = c^2 - a^2 = (c - a)(c + a)$. Let $d = (c - a, c + a)$. Then $2 | d$. Further, we have $d | (c + a) - (c - a) = 2a$ and $d^2 | b^2$. Thus, $d | (2a, b)$ and since $(a, b) = 1$, $d | 2$. It follows that $d = 2$ and $\left(\frac{c-a}{2}, \frac{c+a}{2}\right) = 1$. We have $\frac{c-a}{2} \cdot \frac{c+a}{2} = \left(\frac{b}{2}\right)^2$ and since the factors on the left-hand side are coprime, there exist natural numbers $r$ and $s$ such that $r^2 = \frac{c+a}{2}$, $s^2 = \frac{c-a}{2}$ and $(r, s) = 1$. It follows that

$$a = r^2 - s^2, \;\; c = r^2 + s^2, \;\; b = \sqrt{c^2 - a^2} = 2rs$$

and as $c > a > 0$ and $a$ is odd,

$$r > s > 0, \;\; (r, s) = 1 \;\text{ and }\; r^2 - s^2 = a \equiv 1 \pmod 2.$$

Since $r - s | r^2 - s^2$, $r - s$ must be odd. We know that $a^2 + b^2 = c^2$, $b$ is even and if $p | (a, b)$, then $p$ has to be odd, $p | r$ or $p | s$, $p | r^2 - s^2$, hence $p | r$ and $p | s$ which is a contradiction. □

**Remark.** There exists a bijection between the primitive Pythagorean triplets and the pairs $(r, s)$.

**Theorem 2.6.** *The Diophantine equation $a^4 + b^4 = c^2$ has no solutions in nonzero integers $a, b, c$.*

**Proof.** Assume that $c_0$ is the smallest natural number which satisfies the equation for some nonnegative integers $a_0$, $b_0$, i. e. $a_0^4 + b_0^4 = c_0^2$. Then $(a_0, b_0) = 1$ and $\left(a_0^2\right)^2 + \left(b_0^2\right)^2 = c_0^2$. We may assume that $b_0$ is even. By Theorem 2.5 there exist natural numbers $r, s$ such that $a_0^2 = r^2 - s^2$, $b_0^2 = 2rs$, $c_0^2 = r^2 + s^2$, $r > s > 0$, $(r, s) = 1$ and $2 \not| r - s$.
Assume $r$ is even, then $s$ is odd and $a_0^2 \equiv 0 - 1 \equiv 3 \pmod 4$ which is a contradiction. Thus $r$ is odd and since $2 \not| r - s$, then $s$ is even.
From $r > s > 0$, $(r, s) = 1$ and $r\frac{s}{2} = \left(\frac{b_0}{2}\right)^2$, we get that there exist natural numbers $u$ and $v$ with $r = u^2$, $\frac{s}{2} = v^2$. Since $r$ is odd, $u$ is also odd. From $a_0^2 = r^2 - s^2$ it follows that $a_0^2 + \left(2v^2\right)^2 = \left(u^2\right)^2$. We know that $\left(a_0, 2v^2\right) = (a_0, s)|(r^2 - s^2, s) = (r^2, s) = 1$. By Theorem 2.5 there exist natural numbers $\rho$ and $\tau$ such that $a_0 = \rho^2 - \tau^2$, $2v^2 = 2\rho\tau$, $u^2 = \rho^2 + \tau^2$, $\rho > \tau > 0$, $(\rho, \tau) = 1$ and $2 \not| \rho - \tau$. Thus, $v^2 = \rho\tau$ and $(\rho, \tau) = 1$. Let $\alpha$ and $\beta$ natural numbers such that $\rho = \alpha^2$, $\tau = \beta^2$, $(\alpha, \beta) = 1$. We get that $\alpha^4 + \beta^4 = \rho^2 + \tau^2 = u^2$. Thus $\alpha, \beta, u$ are solutions of $a^4 + b^4 = c^2$. Further, $c_0 = r^2 + s^2 = u^4 + 4v^4 > u^4 \geq u > 0$ which implies that $c_0$ cannot be minimal. □

## 2.1 Homework for Chapter 2

1. Construct a solution of $x^2 + y^2 + z^2 + u^2 = 71$ starting from $20^2 + 3^2 + 4^2 + 1^2 = 6 \cdot 71$ and following the proof of Theorem 2.4.
2. Show that for every $k$ at least $\left[(3/2)^k\right] + 2^k - 2$ $k$-th powers are needed to represent every positive integer as sum of $k$-th powers.
3. Prove: the equation $x^4 + 4y^4 = z^2$ has no nontrivial integer solutions.
4. Prove that there exist infinitely many primes $p$ with $p \equiv 1 \pmod 4$.

These exercises have to be done for Thursday 11 October.

## 2.2 Further exercises for Chapter 2

1. Which numbers can be written as the difference of two squares?
2. a) For which primes $p$ is the congruence $x^2 \equiv -2 \pmod p$ solvable?
   b) Which primes $p$ can be written as $2x^2 + y^2$?
   c) Which integers $n$ can be written as $2x^2 + y^2$?
3. Prove: the area of a Pythagorean triangle is not the square of a rational number.
4. Let $r(n)$ be the number of solutions $(a, b) \in \mathbb{Z}^2$ of $n = a^2 + b^2$. (E.g. $r(10) = 8$ in view of $10 = (\pm 1)^2 + (\pm 3)^2 = (\pm 3)^2 + (\pm 1)^2$. Prove that $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} r(n) = \pi$.
   Hint: Represent $a^2 + b^2$ as $a + bi$.

# Chapter 3

# The weak Prime Number Theorem

Literature:
Tom M. Apostel, Introduction to Analytic Number Theory, Springer-Verlag, 1976.
K. Chandrasekharan, Introduction to Analytic Number Theory, Springer-Verlag, 1968.
H. Davenport, Multiplicative Number Theory, Springer Verlag, 1980.
G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers, Oxford at the Clarendom Press, 5th edition, 1979.
L.K. Hua, Introduction to Number Theory, Springer-Verlag, 1982.

Prime numbers play an important role in number theory. In this chapter the central question is how many primes up to $x$ there are. Euclid already proved that there are infinitely many primes. His proof provides a weak lower bound.

**Theorem 3.1.** *The sequence of prime numbers $p_1 = 2 < p_2 = 3 < p_3 < \ldots < p_n < \ldots$ is infinite and $p_n \leq 2^{2^{n-1}}$ holds for $n = 1, 2, 3, \ldots$.*

**Proof.** $p_1 = 2 \leq 2^{2^0}$, $p_2 = 3 \leq 2^{2^1}$. Assume that the statement is correct for $1, 2, \ldots, n-1$ with $n \geq 3$. Then $p_1 p_2 \cdots p_{n-1} - 1$ is not divisible by any prime number $\leq p_{n-1}$. Thus $p_n \leq p_1 p_2 \cdots p_{n-1} \leq 2^{2^0+2^1+\ldots+2^{n-2}} = 2^{2^{n-1}-1} < 2^{2^{n-1}}$. □

Let $\pi(x) = \sum_{p \leq x} 1$, i.e. the number of primes $\leq x$. By Theorem 3.1 we know that $\pi(2^{2^{n-1}}) \geq n$, which implies $\pi(x) \geq \log\log x$. This approximation can be improved. Gauss (1771-1855) stated

$$\pi(x) \sim \mathrm{li}(x) := \int_2^x \frac{dt}{\log t} \sim \frac{x}{\log x} \quad \text{when } x \to \infty.$$

Here $\sim$ denotes asymptotic equality, i.e. the quotient of the left-hand side and the right-hand side converges to 1 if $x \to \infty$. In fact $\mathrm{li}(x)$ is an excellent approximation for $\pi(x)$, better than $x/\log x$, and it is easy to prove with partial integration that $\mathrm{li}(x) \sim x/\log x$. Around 1850 Riemann sketched a proof of $\pi(x) \sim \mathrm{li}(x)$ using some property of $\zeta(s)$. One of these properties, namely that all of the zeros of $\zeta(s)$ with $0 \leq \mathrm{Re}(s) \leq 1$ lie on the critical line $\mathrm{Re}(s) = \frac{1}{2}$, has not yet been proved and is called the Riemann Hypothesis. In 1896 Hadamard and de le Vallée Poussin proved Gauss' conjecture:

**Theorem 3.2.** (Prime Number Theorem) $\pi(x) \sim \mathrm{li}(x) \sim \dfrac{x}{\log x}$, $x \to \infty$.

In 1947 Erdős and Selberg found a proof of the prime number theorem using only elementary methods. If the Riemann Hypothesis is true then the following holds: for every $\epsilon > 0$ there is an $x_0$ such that $|\pi(x) - \mathrm{li}(x)| \leq x^{\frac{1}{2}+\epsilon}$ for $x \geq x_0(\epsilon)$. Up to now, only the following inequality has been proved (by Korobov and I.M. Vinogradov in 1958):

$$|\pi(x) - \mathrm{li}(x)| \leq x \exp(-c(\log x)^{\frac{3}{5}}) \quad \text{for } x \geq x_0.$$

Here $c$ and $x_0$ are certain constants.

Around 1851, Chebyshev proved that $0.92\frac{x}{\log x} < \pi(x) < 1.11\frac{x}{\log x}$. We give an elementary proof of a slightly weaker theorem:

**Theorem 3.3.** *For all $x \geq 3$ the following holds:* $\dfrac{1}{2}\dfrac{x}{\log x} < \pi(x) < 4\dfrac{x}{\log x}$.

**Lemma 3.1.** *For all natural numbers $n$, we have $\prod_{p \leq n} p < 4^n$.*

**Proof.** For $n = 1, 2$ the theorem is right. Assume that the theorem is correct for $1, 2, \ldots, n-1$. If $n$ is even, then $\prod_{p \leq n} p = \prod_{p \leq n-1} p \leq 4^{n-1} < 4^n$, and we are ready. Assume that $n = 2m+1$. We use $\binom{2m+1}{m} = \dfrac{(2m+1)\cdots(m+2)}{m!}$ is divisible by all prime numbers $p$ with $m+1 < p \leq (2m+1)$. Thus

$$\prod_{m+1 < p \leq 2m+1} p \leq \binom{2m+1}{m} = \frac{1}{2}\left\{\binom{2m+1}{m} + \binom{2m+1}{m+1}\right\}$$

$$< \frac{1}{2}\sum_{r=0}^{2m+1} \binom{2m+1}{r} = \frac{2^{2m+1}}{2} = 2^{2m}.$$

Therefore, we get

$$\prod_{p \leq n} p = \prod_{p \leq 2m+1} p \leq 2^{2m} \prod_{p \leq m+1} p < 2^{2m} 4^{m+1} = 4^{2m+1} = 4^n.$$

□

**Corollary.** For all real numbers $x \geq 1$, we have $\prod_{p \leq x} p < 4^x$.

**Proof of $\pi(x) < 4x/\log x$.**
We have

$$\prod_{p \leq n} p > \prod_{\sqrt{n} < p \leq n} p \geq n^{\frac{1}{2}(\pi(n) - \pi(\sqrt{n}))}.$$

By Lemma 3.1 we find that

$$n^{\frac{1}{2}(\pi(n) - \pi(\sqrt{n}))} < 4^n$$

or, equivalently,

$$\pi(n) - \pi(\sqrt{n}) < \frac{2n \log 4}{\log n} < \frac{3n}{\log n}.$$

Since $\pi(\sqrt{n}) \leq \sqrt{n} < \frac{n}{\log n}$ for $n \geq 3$, it follows that $\pi(n) < \frac{3n}{\log n} + \sqrt{n} < \frac{4n}{\log n}$. Thus, for $x \geq 3$ we have

$$\pi(x) = \pi([x]) < \frac{4[x]}{\log{[x]}} \leq \frac{4x}{\log x}.$$

□

**Lemma 3.2.** *For all natural numbers $n$, we have $\frac{2^{2n}}{2n} \leq \binom{2n}{n} < 2^{2n}$.*

**Proof.** We know that $\binom{2n}{n} < (1+1)^{2n} = 2^{2n}$. On one hand, the largest element in a row of Pascal's Triangle is the middle one, on the other hand

$$\binom{2n}{n} = \binom{2n-1}{n-1} + \binom{2n-1}{n} \geq 2\frac{(1+1)^{2n-1}}{2n} = \frac{2^{2n}}{2n}.$$

□

**Lemma 3.3.** *Assume $p^b | \binom{2n}{n}$ for a certain prime number $p$. Then $p^b \leq 2n$.*

**Proof.** The number of prime factors $p$ in $m!$ equals

$$\left[\frac{m}{p}\right] + \left[\frac{m}{p^2}\right] + \left[\frac{m}{p^3}\right] + \ldots,$$

since among the numbers $1, 2, \ldots, m$ there are $\left[\frac{m}{p}\right]$ which are divisible by $p$, among the numbers $1, 2, \ldots, m$ there are $\left[\frac{m}{p^2}\right]$ which are divisible by $p^2$, etc. The number of prime factors $p$ in $\binom{2n}{n} = \frac{(2n)!}{n!n!}$ is therefore

$$b = \left[\frac{2n}{p}\right] + \left[\frac{2n}{p^2}\right] + \ldots + \left[\frac{2n}{p^t}\right] - 2\left[\frac{n}{p}\right] - 2\left[\frac{n}{p^2}\right] - \ldots - 2\left[\frac{n}{p^t}\right]$$

where $p^t \leq 2n < p^{t+1}$. Thus $b = \sum_{r=1}^{t} \left(\left[\frac{2n}{p^r}\right] - 2\left[\frac{n}{p^r}\right]\right)$.
One can see that $[2x] - 2[x]$ equals 0 or 1 for every $x \in \mathbb{R}$. Thus $b \leq t$ and $p^b \leq p^t \leq 2n$. □

**Proof of $\pi(x) > x/(2\log x)$.**
By Lemma 3.3 we know that

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

By Lemma 5.2 we know that $\binom{2n}{n} \geq 2^{2n}/2n$. Thus $(2n)^{\pi(2n)} \geq 2^{2n}/2n$ or

$$\pi(2n) \geq \frac{2n\log 2}{\log{(2n)}} - 1 \;\text{ for all } n.$$

Let $n = [x]$ and assume that $n \geq 20$. If $n$ is even, say $n = 2m$, then we have

$$\pi(x) = \pi(n) = \pi(2m) \geq \frac{2m\log 2}{\log{(2m)}} - 1 = (\log 2)\frac{n}{\log n} - 1 > \frac{1}{2}\frac{n+1}{\log{(n+1)}} > \frac{1}{2}\frac{x}{\log x}.$$

If $n$ is odd, then we have, by the previous inequality,

$$\pi(x) = \pi(n+1) > \frac{1}{2}\frac{n+2}{\log(n+2)} > \frac{1}{2}\frac{x}{\log x}.$$

For $3 \leq x < 20$ one can verify that the inequality is true by direct calculations. □

**Theorem 3.4.** (Bertrand's Postulate) *For every natural number $n$ there exists at least one prime number $p$ with $n < p \leq 2n$.*

**Lemma 3.4.** *If $n \geq 3$ and $\frac{2}{3}n < p \leq n$ then $p \not| \binom{2n}{n}$.*

**Proof.** From the conditions we know that $p > 2$, $\left[\frac{2n}{p}\right] = 2$ and $p^2 > 2n$. The number of prime factors $p$ in $(2n)!$ is $\left[\frac{2n}{p}\right] + 0 + 0 + \ldots = 2$. The number of prime factors $p$ in $n!$ is $\left[\frac{n}{p}\right] + 0 + 0 + \ldots = 1$. Thus $\binom{2n}{n} = \frac{(2n)!}{n!n!}$ is not divisible by $p$. □

**Proof of Theorem 3.4.**
In the sequence of prime numbers $2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 557$ each number is less then twice the previous one which yields that the theorem is correct for $n \leq 500$. Assume that the theorem is not true for a certain number $n > 500$. By Lemma 3.4, each prime divisor of $\binom{2n}{n}$ at most $2n/3$. Suppose that $p^b || \binom{2n}{n}$. By Lemma 3.3, we know that $p^b \leq 2n$. Thus: if $b \geq 2$, then $p^2 \leq 2n$, hence $p \leq \sqrt{2n}$. Since there exist at most $[\sqrt{2n}]$ such prime numbers, we have

$$\binom{2n}{n} \leq \prod_{p \leq \sqrt{2n}} p^b \prod_{\sqrt{2n} < p \leq \frac{2n}{3}} p \leq (2n)^{[\sqrt{2n}]} \prod_{p \leq \frac{2n}{3}} p.$$

From Lemma 3.2 it follows that $\binom{2n}{n} \geq \frac{2^{2n}}{2n}$ and from Lemma 3.1 we have $\prod_{p \leq 2n/3} p \leq 4^{2n/3}$. Combining these inequalities we obtain that

$$\frac{2^{2n}}{2n} \leq (2n)^{\sqrt{2n}} 4^{2n/3},$$

hence

$$2^{2n/3} \leq (2n)^{\sqrt{2n}+1}$$

and taking logarithms on both sides, we get

$$\frac{2n}{3}\log 2 \leq (\sqrt{2n}+1)\log 2n.$$

However, this is impossible for $n \geq 500$ (check this). Hence, the theorem is correct for $n > 500$, too, and thus for every natural number. □

Theorem 3.4 was formulated by Bertrand in 1845 and it was completely proved by Chebyshev in 1850.

We derive some consequences of Theorem 3.3.

**Theorem 3.5.** *Let $p_r$ be the $r$th prime number ($p_1 = 2, p_2 = 3, p_3 = 5, \ldots$). Then we have*

$$\frac{1}{4} r \log r < p_r < 4r \log r \quad \textit{for } r \geq 2.$$

**Proof.** If $r \geq 2$ then $p_r \geq 3$. By Theorem 3.3 we know that $r = \pi(p_r) < 4\dfrac{p_r}{\log p_r}$. Hence, $p_r > \frac{1}{4} r \log p_r \geq \frac{1}{4} r \log r$. This proves the first inequality.
It is easy to verify the inequality $p_r < 4r \log r$ for $r = 2, 3, \ldots, 8$. Assume that $r \geq 9$. Then $r \geq 4 \log r$. From the first inequality of Theorem 3.3 it follows that $\pi(r^2) > r^2/(4 \log r) \geq r$. Thus $p_r < r^2$, and hence $\log p_r < 2 \log r$. Using the first inequality of Theorem 3.3 we obtain

$$r = \pi(p_r) > \frac{1}{2} \frac{p_r}{\log p_r} > \frac{p_r}{4 \log r}.$$

□

**Theorem 3.6.** *There exist constants* $c_1$ *and* $c_2$ *such that, for* $x \geq 3$*,*

$$\frac{1}{4} \log \log x - c_1 \leq \sum_{p \leq x} \frac{1}{p} \leq 4 \log \log x + c_2.$$

*Hence* $\sum_p \dfrac{1}{p}$ *is divergent.*

**Proof.** We use that $\int_2^x \dfrac{\mathrm{d}t}{t \log t} = [\log \log t]_{t=2}^{t=x} = \log \log x - \log \log 2$. On the one hand,

$$\sum_{p \leq x} \frac{1}{p} \leq \sum_{r=1}^{x} \frac{1}{p_r} < 1 + \sum_{r=3}^{x} \frac{4}{r \log r} \leq 1 + \int_2^x \frac{4\mathrm{d}t}{t \log t} = 4 \log \log x + c_2.$$

On the other hand, by Theorem 3.3 and Theorem 3.5, for $x \geq 20$,

$$\sum_{p \leq x} \frac{1}{p} \geq \sum_{r=1}^{x/(2\log x)} \frac{1}{p_r} > \sum_{r=2}^{x/(2\log x)} \frac{1}{4r \log r} > \frac{1}{4} \int_2^{x/(2\log x)} \frac{\mathrm{d}t}{t \log t} =$$

$$= \frac{1}{4} \log \log \frac{x}{2 \log x} \underbrace{- \frac{1}{4} \log \log 2}_{>0} > \frac{1}{4} \log \log x^{1/3} = \frac{1}{4} \log \log x - \frac{1}{4} \log 3.$$

We can choose a suitable constant such that the inequality is fulfilled also for $2 \leq x < 20$. □

By Theorem 3.2 (PNT) and with the help of other techniques it can be proved that for each $\epsilon > 0$ there exists an $r_0$ such that

$$(1 - \epsilon) r \log r < p_r < (1 + \epsilon) r \log r \quad \text{for } r \geq r_0$$

and

$$\lim_{x \to \infty} \left\{ \sum_{p \leq x} \frac{1}{p} - (\log \log x + \gamma) \right\} = 0$$

where $\gamma$ is the constant of Euler.

Finally, we mention a few classical problems on prime numbers.

The distance of consecutive prime numbers. There is a conjecture which says that for $x > x_0$ each interval $[x, x + 2(\log x)^2]$ contains a prime number. Even under the Riemann Hypothesis the best proved result is only that each interval $[x, x + c(\epsilon)x^{\frac{1}{2}+\epsilon}]$ contains a prime number. Actually, the same statement is proved unconditionally for each interval $[x, x + cx^{11/20}]$.
It is also a conjecture that there are infinitely many twin primes, i.e. pairs of primes $p$, $p + 2$. It is known that there are infinitely many primes $p$ such that $p+2$ has at most two factors. It is not known whether there are infinitely many primes $p$ such that the next prime is less than $p + \frac{1}{10}\log p$.

Goldbach's conjecture. The conjecture states that every even integer greater than 2 can be written as the sum of two primes. In 1937 Vinogradov proved that there exists a number $n_0$ such that every odd number $n > n_0$ can be written as the sum of three primes. In 1966 Chen Ching-Jun showed that every sufficiently large even number can be written as the sum of either two primes or a prime and a product of two primes.

Prime numbers in arithmetic progressions. Let $k, l \in \mathbb{N}$ and $(k, l) = 1$. In 1842 Dirichlet proved that the sequence $l, l + k, l + 2k, \ldots$ contains infinitely many primes. Let $\pi(x, k, l)$ denote the number of primes $p \leq x$ with $p \equiv l \pmod{k}$. Then we have

$$\pi(x, k, l) \sim \frac{1}{\varphi(k)} \frac{x}{\log x}, \ x \to \infty \quad \text{(Prime number theorem for arithmetic progressions.)}$$

Open problem: Are there infinitely many prime numbers in the sequence $\left\{n^2 + 1\right\}_{n=1}^{\infty}$?

## 3.1 Homework for Chapter 3

1. Prove for $N = 1, 2, \ldots$

   a) $\prod_{p \leq N} \dfrac{1}{1 - p^{-1}} \geq \sum_{n=1}^{N} \dfrac{1}{n} > \log N$,

   b) $0 < \log\left(\dfrac{1}{1 - p^{-1}}\right) - \dfrac{1}{p} < \dfrac{1}{2p(p-1)}$,

   c) $\sum_{p \leq N} \left(\log\left(\dfrac{1}{1 - p^{-1}}\right) - \dfrac{1}{p}\right) < \dfrac{1}{2}$,

   d) $\sum_{p \leq N} \dfrac{1}{p} > \log\log N - \dfrac{1}{2}$.

2. Check whether the following sums converge when $x \to \infty$ and if they don't compute the order of growth of the sums.

$$\text{a) } \sum_{p \leq x} \frac{1}{p\log p}, \qquad \text{b) } \sum_{p \leq x} \frac{\log p}{p}, \qquad \text{c) } \sum_{p \leq x} \frac{1}{\sqrt{p}(\log p)^2}.$$

3. a) How many prime factors $p$ divide $\binom{2n}{n}$ if

$$\text{i) } \tfrac{1}{2}n < p \leq \tfrac{2}{3}n, \qquad \text{ii) } \tfrac{2}{5}n < p \leq \tfrac{1}{2}n.$$

   b) Generalize property 3a).

4. Let $\varepsilon > 0$. Prove that the Prime Number Theorem implies that there exists an $x_0$ such that for every $x > x_0$ the interval $[x, (1+\varepsilon)x]$ contains a prime number.

These exercises have to be done for Thursday 1 November.

## 3.2 Further exercises for Chapter 3

1. Prove that $\mathrm{li}(x) \sim \dfrac{x}{\log x}$ as $x \to \infty$.
2. What is the highest power of 10 which divides 1000! ?
3. Prove that there are infinitely many positive integers such that $\pi(n)|n$.
4. Check whether the following sum converges when $x \to \infty$ and if it doesn't compute the order of growth of the sum:
   $\sum_{p \leq x} \dfrac{1}{p \log \log p}$.

# Chapter 4

# Multiplicative functions and Dirichlet series.

## 4.1 Multiplicative functions

An arithmetic function is a map $f : \mathbb{N} \longrightarrow \mathbb{C}$. An arithmetic function is called multiplicative if $f(1) \neq 0$ and $f(mn) = f(m)f(n)$ for every pair $m$, $n$ with $(m, n) = 1$. The function is called completely multiplicative if $f(1) \neq 0$ and $f(mn) = f(m)f(n)$ for every pair $m$, $n$.

**Examples.** The Euler totient function $\varphi$ is multiplicative. The following functions are completely multiplicative: $e$ with $e(1) = 1$ and $e(n) = 0$ for $n \geq 2$, $E$ with $E(n) = 1$ for all $n$ and $N_a$ for $a \in \mathbb{R}$ with $N_a(n) = n^a$ for all $n$.

**Theorem 4.1.** *(i) If $f$ is multiplicative then $f(1) = 1$.*
*(ii) There exists exactly one multiplicative function with given values in the prime powers (i.e. $p^k$, $p$ prime, $k \in \mathbb{N}$).*
*(iii) There exists exactly one completely multiplicative function with given values in the prime numbers.*

**Proof.** (i) We know that $f(1) = f(1)f(1)$. Since $f(1) \neq 0$, it follows that $f(1) = 1$.
(ii) Let $f$ be a multiplicative function with given values $f(p^k)$ for all primes $p$ and $k \in \mathbb{N}$. Let $n \in \mathbb{N}$. Let $p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$ be the prime factorization of $n$. Then $f(n) = f(p_1^{k_1}) \cdots f(p_r^{k_r})$. Thus $f(n)$ is uniquely determined.
(iii) Let $f$ be a completely multiplicative function with given values $f(p)$ for all primes $p$. Let $p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$ be the prime factorization of $n$. Then $f(n) = (f(p_1))^{k_1} (f(p_2))^{k_2} \cdots (f(p_r))^{k_r}$. Thus $f(n)$ is uniquely determined. $\Box$

**Examples.** Let $a \in \mathbb{R}$. There exists exactly one multiplicative function $f$ with $f(p^k) = a$ for each prime power $p^k$. This function is denoted by $a^{\omega(n)}$. Notice that $\omega(n) = \sum_{p|n} 1$. There exists exactly one completely multiplicative function $g$ with $g(p) = a$ for each prime $p$. This function is denoted by $a^{\Omega(n)}$. Notice that $\Omega(n) = k_1 + k_2 + \ldots + k_r$ if $p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$ is the prime factorization of $n$.

Assume that $f$ and $g$ are arithmetic functions. The convolution $f * g$ of $f$ and $g$ is defined as follows:

$$f * g(n) = \textstyle\sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{d_1 d_2 = n} f(d_1) g(d_2) \ \text{ for all } n.$$

**Theorem 4.2.**

$$\begin{array}{ll} \text{(i)} \;\; f * g = g * f & \textit{for all } f \textit{ and } g, \\ \text{(ii)} \;\; (f * g) * h = f * (g * h) & \textit{for all } f, g, h, \\ \text{(iii)} \;\; f * e = f & \textit{for all } f. \end{array}$$

**Proof.** Exercise!

**Theorem 4.3.** *The convolution product of two multiplicative functions is again multiplicative.*

**Proof.** Assume that $f$ and $g$ are multiplicative functions and $h := f * g$. Let $(m, n) = 1$. By the fundamental theorem of arithmetic it follows that each divisor $d$ of $mn$ corresponds with a pair $(d_1, d_2)$ such that $d_1 | m$, $d_2 | n$ and $d = d_1 d_2$. Notice that $(d_1, d_2) = 1$. Then

$$h(mn) = \sum_{d|mn} f(d) g\left(\frac{mn}{d}\right) = \sum_{d_1|m} \sum_{d_2|n} f(d_1 d_2) g\left(\frac{mn}{d_1 d_2}\right) =$$

$$= \sum_{d_1|m} \sum_{d_2|n} f(d_1) f(d_2) g\left(\frac{m}{d_1}\right) g\left(\frac{n}{d_2}\right) = \left(\sum_{d_1|m} f(d_1) g\left(\frac{m}{d_1}\right)\right) \left(\sum_{d_2|n} f(d_2) g\left(\frac{n}{d_2}\right)\right) =$$

$$= h(m) h(n).$$

□

**Corollary.** The following functions are multiplicative:
$\tau(n) = \sum_{d|n} 1$ since $\tau = E * E$,
$\sigma_a(n) = \sum_{d|n} d^a$ since $\sigma_a = N_a * E$.

Notice that $\tau = \sigma_0$. We write $\sigma$ instead of $\sigma_1$ for the sum of divisors.

By Theorem 4.1 (i) we have the following in the given notation:

$$\tau(n) = \textstyle\prod_{j=1}^r (k_j + 1) \ \text{ for all natural numbers } n,$$

$$\sigma(n) = \textstyle\prod_{j=1}^r \dfrac{p_j^{k_j+1} - 1}{p_j - 1} \ \text{ for all natural numbers } n.$$

Indeed $\tau(p^k) = k + 1$ and $\sigma(p^k) = 1 + p + \ldots + p^k = \dfrac{p^{k+1} - 1}{p - 1}$.

In ancient times and the middle ages people assigned magic properties to numbers with special divisibility properties. In particular, 6 and 28 are perfect because they equal to the sum of their divisors which are less than the number itself, $6 = 1 + 2 + 3$, $28 = 1 + 2 + 4 + 7 + 14$.

A positive integer $n$ is called perfect if $\sigma(n) = 2n$. Two numbers $m$ and $n$ are called amicable if $\sigma(m) = \sigma(n) = m + n$. An example of amicable pairs is the pair 220, 284.

**Theorem 4.4.** *(i) (*Euclides*) If $2^k - 1$ is prime then $2^{k-1}(2^k - 1)$ is perfect.*
*(ii) (*Euler*) If $n$ is even and perfect, then $n = 2^{k-1}(2^k - 1)$* where $2^k - 1$ is prime.

**Proof.** (i) We know that $(2^{k-1}, 2^k - 1) = 1$. Thus, since $2^k - 1$ is prime, $\sigma(2^{k-1}(2^k - 1)) = \sigma(2^{k-1}) \cdot \sigma(2^k - 1) = \dfrac{2^k - 1}{2 - 1} \cdot 2^k = 2(2^{k-1}(2^k - 1))$.
(ii) Assume that $n$ is even and perfect. Write $n = 2^{k-1}m$ with $k \geq 2$, $m$ is odd. From $\sigma(n) = 2n$, it follows that

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

Since $(2^k, 2^k - 1) = 1$, it follows that $2^k - 1 | m$ and $2^k | \sigma(m)$. Let $\sigma(m) = 2^k c$. Then $2^k m = (2^k - 1)\sigma(m)$ implies that $m = c(2^k - 1)$. Now assume that $c > 1$. Then $1, c, m$ are distinct divisors of $m$. It follows that $\sigma(m) \geq 1 + c + m > c + m = c \cdot 2^k = \sigma(m)$, which is a contradiction. Thus $c = 1$. We get that $\sigma(m) = 2^k$ and $m = 2^k - 1$. Hence $m$ is prime and $n = 2^{k-1}(2^k - 1)$. □

**Remarks.** 1. It is not known if any odd perfect number exists, making the existence of odd perfect numbers appear unlikely.
2. If $2^k - 1$ is prime then $k$ is prime. (Exercise!)
The converse is not true, a counterexample is $2^{11} - 1 = 23 \times 89$. A prime number of the form $2^k - 1$ is called a Mersenne Prime. The largest known Mersenne Prime is the 44th and was found in 2006, it is $2^{32582657} - 1$. From the Mersenne Primes $2^2 - 1$, $2^3 - 1$, $2^5 - 1$, $2^7 - 1$, we get the perfect numbers 6, 28, 496, 8128.

**Theorem 4.5.** *Let $f$ be an arithmetic function with $f(1) \neq 0$. Then there exists exactly one arithmetic function $g$ such that $f * g = e$. If $f$ is multiplicative then $g$ is also multiplicative.*

**Proof.** From $f(1)g(1) = e(1) = 1$ and $f(1) \neq 0$, it follows that $g(1) = 1/f(1)$. Assume that $g(1), g(2), \ldots, g(n-1)$ are already given. Then

$$0 = e(n) = f * g(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{d|n, d>1} f(d) g\left(\frac{n}{d}\right) + f(1)g(n).$$

Since $f(1) \neq 0$ and all values except for $g(n)$ are given, one can compute $g(n)$. With induction it follows that $g$ is uniquely determined.
Assume that $f$ is multiplicative. Let $h$ be a multiplicative function which is equal to $g$ in the prime powers, i. e. $h(p^k) = g(p^k)$. By Theorem 4.1 $h$ is uniquely determined and by Theorem 4.3 $u := f * h$ is multiplicative. Then, for each prime power $p^k$,

$$u(p^k) = \sum_{d|p^k} f(d) h\left(\frac{p^k}{d}\right) = \sum_{j=0}^{k} f(p^j) h(p^{k-j}) =$$

$$= \sum_{j=0}^{k} f(p^j) g(p^{k-j}) = \sum_{d|p^k} f(d) g\left(\frac{p^k}{d}\right) = f * g(p^k) = e(p^k).$$

Since $u$ and $e$ are equal in the prime powers and both of them are multiplicative, by Theorem 4.1 they must be the same. Thus $g = h$ is multiplicative. □

**Remark.** By Theorem 4.2 $e$ is the neutral element of the convolution product. Therefore $g$ is the inverse function of $f$, we denote it by $f^{-1}$. This yields that the multiplicative functions form an abelian group with the convolution product as operator.

**Example.** Let $\mu = E^{-1}$. The function $\mu$ is called Möbius function. Theorem 4.5 implies that $\mu$ is multiplicative. By Theorem 4.1 it suffices to determine $\mu$ in the prime powers. We know that $\mu(1) = 1$, $0 = e(p) = \mu(1)E(p) + \mu(p)E(1) = \mu(1) + \mu(p) = 1 + \mu(p)$, thus $\mu(p) = -1$ for all prime numbers $p$. Further, for $k = 2, 3, \ldots$,

$$0 = e(p^k) = \mu(1)E(p^k) + \mu(p)E(p^{k-1}) = \ldots + \mu(p^k)E(1) =$$

$$= 1 \cdot 1 + (-1) \cdot 1 + \mu(p^2) + \mu(p^3) + \ldots + \mu(p^k).$$

We can see that $\mu(p^k) = 0$ for all $k = 2, 3, \ldots$ and prime numbers $p$.
Hence, by Theorem 4.1, we obtain that

$$\mu(n) = \begin{cases} 1 & \textit{if } n = 1, \\ (-1)^r & \textit{if } n \textit{ is a product of } r \textit{ distinct prime numbers}, \\ 0 & \textit{if } n \textit{ is divisible by a square.} \end{cases}$$

A number $n$ with $\mu(n) \neq 0$ is called a squarefree number. The Möbius function plays an important role in combinatorics, for example in the inclusion-exclusion rule.

**Theorem 4.6.** (Möbius inversion formula) *Let $f$ be an arithmetic function and $F(n) = \sum_{d|n} f(d)$ for all $n$. Then we have*

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

**Proof.** We know that $F = f * E$. Hence, $f = e * f = \mu * E * f = \mu * F$. □

The previous theorems can be used to obtain relations between multiplicative functions.

1. We give another proof for $\sum_{d|n} \varphi(d) = n$.
   Let $p$ be a prime, $k \geq 1$. Then

   $$\mu * N_1(p^k) = \mu(1)p^k + \mu(p)p^{k-1} = p^k - p^{k-1} = p^k \left(1 - \frac{1}{p}\right) = \varphi(p^k).$$

   Since $\mu, N_1$ and $\varphi$ are multiplicative, it follows that $\mu * N_1 = \varphi$. Thus $N_1 = \varphi * E$, and $n = \sum_{d|n} \varphi(d)$ for all $n$.

2. We have $\tau * \varphi = E * E * \mu * N_1 = E * N_1 = \sigma$. Thus

   $$\textstyle\sum_{d|n} \tau(d)\varphi\left(\frac{n}{d}\right) = \sigma(n) \text{ for all } n.$$

## 4.2 Dirichlet series

Multiplicative functions are in correspondence with Dirichlet series. Let $f$ be an arithmetic function and $s \in \mathbb{R}$ (or $\in \mathbb{C}$). The series $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$ is the <u>Dirichlet series</u> corresponding to $f$.

**Theorem 4.7.** *If for certain constants $b$ and $c$, $|f(n)| < bn^c$ holds for all $n \in \mathbb{N}$ then the Dirichlet series converges absolutely for $\Re(s) > c+1$.*

**Proof.** Assume that $\Re(s) > c+1$. Let $N \in \mathbb{N}$. Then we have

$$|\sum_{n=N}^{\infty} \frac{f(n)}{n^s}| \leq \sum_{n=N}^{\infty} \frac{|f(n)|}{n^{\Re(s)}} \leq b \sum_{n=N}^{\infty} n^{c-\Re(s)}.$$

Since $\sum_{n=1}^{\infty} n^{c-\Re(s)}$ is convergent for $\Re(s) > c+1$, by the Theorem of Weierstrass, the Dirichlet series is absolutely convergent. □

**Example.** 1. $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ is convergent for $\Re(s) > 1$.

2. $\Phi(s) = \sum_{n=1}^{\infty} \frac{\varphi(n)}{n^s}$ is convergent for $\Re(s) > 2$ because $\varphi(n) \leq n$.

The following theorem says that the convolution product of arithmetic functions corresponds to the ordinary product of Dirichlet series.

**Theorem 4.8.** *If $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$ and $G(s) = \sum_{n=1}^{\infty} g(n)n^{-s}$ are absolutely convergent, then*

$$F(s)G(s) = \sum_{n=1}^{\infty} (f * g(n))n^{-s}.$$

**Proof.** $F(s)G(s) = \sum_{d_1=1}^{\infty} f(d_1)d_1^{-s} \sum_{d_2=1}^{\infty} g(d_2)d_2^{-s} = \sum_{k=1}^{\infty} \underbrace{\left( \sum_{d_1 d_2 = k} f(d_1)g(d_2) \right)}_{f*g(k)} k^{-s}.$

The summation is correct since both series are absolutely convergent. □

The next theorem shows the significance of (completely) multiplicative functions.

**Theorem 4.9.** *(i) If $f$ is multiplicative and $\sum_{n=1}^{\infty} f(n)n^{-s}$ is absolutely convergent then we have*

$$\sum_{n=1}^{\infty} f(n)n^{-s} = \prod_p \left\{1 + f(p)p^{-s} + f(p^2)p^{-2s} + \ldots\right\}. \quad \textit{(\underline{Euler product})}$$

*(ii) If $f$ is completely multiplicative and $\sum_{n=1}^{\infty} f(n)n^{-s}$ is absolutely convergent then we have*

$$\sum_{n=1}^{\infty} f(n)n^{-s} = \prod_p \left\{1 - f(p)p^{-s}\right\}^{-1}. \quad \textit{(\underline{Euler product})}$$

**Proof.** (i) By the fundamental theorem of arithmetic we get that each term $f(n)n^{-s} = f(p_1^{k_1})p_1^{-k_1 s} \cdots f(p_r^{k_r})p_r^{-k_r s}$ is counted exactly once on both sides.

Because of the absolute convergence both sides are equal.
(ii) We have $\sum_{j=0}^{\infty} f(p^j)p^{-js} = \sum_{j=0}^{\infty}(f(p))^j(p^{-s})^j = (1-f(p)p^{-s})^{-1}$. □

**Example.** 1. $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p (1-p^{-s})^{-1}$ for $\Re(s) > 1$.
2. $M(s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)} = \prod_p (1-p^{-s})$ for $\Re(s) > 1$.
3. $\Phi(s)\zeta(s) = \sum_{n=1}^{\infty} \frac{\varphi * E(n)}{n^s} = \sum_{n=1}^{\infty} \frac{n}{n^s} = \zeta(s-1)$, thus $\Phi(s) = \frac{\zeta(s-1)}{\zeta(s)}$ for $\Re(s) > 2$.

There exists a number $\sigma_a \in [-\infty, \infty]$, called the abscissa of absolute convergence, such that the series $\sum_{n=1}^{\infty} f(n)n^{-s}$ converges absolutely if $\Re(s) > \sigma_a$ and does not converge absolutely if $\Re(s) < \sigma_a$.

It is possible that a Dirichlet series is conditionally convergent in a point $s$ with $\Re(s) < \sigma_a$. However, for any such $s$ we have $\Re(s) \geq \sigma_a - 1$. A Dirichlet series converges in a half plane $\Re(s) > \sigma_c$. Thus $\sigma_c \geq \sigma_a - 1$. If a Dirichlet series has abscissa of convergence $\sigma_c$ then it is uniformly convergent on any compact set contained in the half plane $\Re(s) > \sigma_c$ (and not convergent for $\Re(s) < \sigma_c$). In the halfplane $\Re(s) > \sigma_c$ it represents an analytic function whose successive derivatives are obtained by termswise differentation of the series. If $F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}$ with $f(n) \geq 0$ for all $n \geq 1$, and $\sigma_c$ is finite, then the point of intersection of the real axis with the line of convergence is a singularity of $F(s)$.

## 4.3 Homework for Chapter 4

1. Prove without taking care of convergence:
   a) $(\zeta(s))^2 = \sum_{n=1}^{\infty} \frac{\tau(n)}{n^s}$, b) $\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$,
   c) $\frac{\zeta(2s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{(-1)^{\Omega(n)}}{n^s} = \prod_p \frac{1}{1+p^{-s}}$.
2. Prove that $\prod_p = \frac{p^2+1}{p^2-1} = \frac{5}{2}$. (Hint: $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(4) = \frac{\pi^4}{90}$.)

These exercises (and two from Chapter 5) have to be done by Thursday 15 November.

## 4.4 Further exercises for Chapter 4

1. Prove that if $2^k - 1$ is prime, then $k$ is prime.
2. Compute a) $\prod_p = \left(1 + \frac{1}{p^2}\right)$, b) $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2}$.
3. Compute $(2^\omega * \mu)(n)$ for every positive integer $n$.
4. Determine the function $f$ for which $\zeta(s)\zeta(2s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ $(s > 1)$.

# Chapter 5

# Primes in arithmetic progressions.

## 5.1 Dirichlet characters

Literature:
H.L. Montgomery and R.C. Vaughan, Multiplicative Number Theory I, Classical Theory, Cambridge University Press, 2007. (New)
K. Chandrasekharan, Introduction to Analytic Number Theory, Springer-Verlag, 1968.
H. Davenport, Multiplicative Number Theory, Springer Verlag, 1980.

A Dirichlet character is a completely multiplicative function $\chi$ which is periodic mod $k$ for some given $k$ and for which $\chi(n) = 0$ if $(n, k) > 1$. If $k$ is a prime $p$, then it has a primitive root $g$, that is $g, g^2, \ldots, g^{p-1} = 1$ are just representing the residue classes not $\equiv 0 \pmod{p}$. Then $\chi$ is completely determined by $\chi(g)$. Furthermore, $(\chi(g))^{p-1} = \chi(g^{p-1}) = \chi(1) = 1$. Hence $\chi(g)$ is a $(p-1)$th root of unity. For each choice of $\chi(g)$ we find another character. The number of Dirichlet characters mod $p$ is therefore $p - 1$.

If $\chi$ assumes only real values, we call it a real character. If $\chi(g) = 1$, then we have the principal character $\chi_0$ defined by

$$\chi_0(n) = \begin{cases} 1 & if \ \ p \not| n, \\ 0 & otherwise. \end{cases}$$

If $\chi(g) \neq 1$, then $\chi(g) = -1$. The Dirichlet character in this case is the Legendre symbol.

Next we consider characters to prime power moduli. If $q$ is an odd power or $q = 4$, then there exists a primitive root $g$ of $q$ and we can proceed as above, the only difference being that the modulus to which it is defined is now $\varphi(q) = p^{\alpha-1}(p-1)$ when $q = p^\alpha$. We define the characters to the modulus $q$ by taking any $\varphi(q)$th root of unity $\chi(g)$ and putting

$$\chi(n) = \begin{cases} (\chi(g))^m & for \quad (n,p)=1, \quad n \equiv g^m \quad (mod \quad q), \\ 0 & otherwise. \end{cases}$$

The number of characters is $\varphi(q) = \varphi(p^\alpha)$. If $q = 2^\alpha$ with $\alpha \geq 3$, then there is no primitive root $g$, but every primitive residue class is representable uniquely as $(-1)^\nu 5^{\nu'}$ where $\nu$ is defined to the modulus 2 and $\nu'$ to the modulus $2^{\alpha-2}$. We define the characters in this case by

$$\chi(n) = \omega^\nu (\omega')^{\nu'}$$

where $\omega^2 = 1$ and $(\omega')^{2^{\alpha-2}} = 1$. The number of characters is $2^{\alpha-1} = \varphi(2^\alpha)$.

We shall now define the Dirichlet characters to the general modulus $k$. Let

$$k = 2^\alpha p_1^{\alpha_1} p_2^{\alpha_2} \cdots \tag{5.1}$$

be the standard factorization of $k$ into prime powers. We define the characters to the modulus $k$ as products of arbitrary characters to the various prime power moduli. The total number of characters is

$$\varphi(2^\alpha)\varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2}) \cdots = \varphi(k),$$

since $\varphi$ is multiplicative. One of the characters is the principal character $\chi_0(n)$ defined by

$$\chi_0(n) = \begin{cases} 1 & if \quad (n,k)=1, \\ 0 & otherwise. \end{cases}$$

The characters have an important property that can be expressed in two forms. In the first form it states that

$$\sum_{n=0}^{k-1} \chi(n) = \begin{cases} \varphi(k) & if \quad \chi = \chi_0, \\ 0 & otherwise. \end{cases} \tag{5.2}$$

The truth of (5.2) is an immediate deduction from the representation of the general character. For, if $k$ is given by (5.1),

$$\chi(n) = e^{2\pi i \left[ \frac{m_0\nu_0}{2} + \frac{m_0'\nu_0'}{2^{\alpha-2}} + \frac{m_1\nu_1}{\varphi(p_1^{\alpha_1})} + \frac{m_2\nu_2}{\varphi(p_2^{\alpha_2})} + \ldots \right]} \tag{5.3}$$

for $(n,k)=1$, where $m_0$, $m_0'$, $m_1$, $m_2$, $\ldots$ are integers which take all the values modulo the corresponding denominators, each an equal number of times. (If $\alpha \leq 2$, the second term is to be omitted; if $\alpha = 1$, then $\nu_0$ assumes only the value 0.) Hence summation over $n$ is equivalent to a summation over $\nu_0, \nu_0', \nu_1, \nu_2, \ldots$, each to its respective modulus, and this gives 0 unless each of $m_0, m_0', m_1, m_2, \ldots$ is congruent to 0 with respect to its corresponding modulus in view of

$$\sum_{l=0}^{m-1} e^{2\pi i \frac{l}{m}} = \begin{cases} m & if \quad l \equiv 0 \quad (mod \quad m), \\ 0 & otherwise. \end{cases}$$

In the 'unless' case $\chi = \chi_0$, the value of the sum is $\varphi(k)$.

The second form of the property is that

$$\sum_{\chi} \chi(n) = \begin{cases} \varphi(k) & if \quad n \equiv 1 \pmod{k}, \\ 0 & otherwise, \end{cases} \tag{5.4}$$

where the summation is over all the $\varphi(k)$ characters. The same proof applies, but with $m$'s and $\nu$'s interchanged. The only case in which the sum does not vanish is that in which all the $\nu$'s are 0, and then $n \equiv 1 \pmod{k}$. We use (5.4) to select those integers which are in a given residue class. If $(a, k) = 1$, then

$$\frac{1}{\varphi(k)} \sum_{\chi} \overline{\chi(a)}\chi(n) = \begin{cases} 1 & if \quad n \equiv a \quad (mod \quad k), \\ 0 & otherwise, \end{cases} \tag{5.5}$$

for we have $\overline{\chi(a)}\chi(n) = \chi(n')$ where $n' \equiv 1 \pmod{k}$ if and only if $n \equiv a$ (mod $k$).

## 5.2 L-functions and the proof of Dirichlet's theorem

The aim of this section is to prove Dirichlet's theorem on arithmetic progressions:

**Theorem 5.1.**(Dirichlet, 1842) *Every arithmetic progression* $\{kn + l\}_{n=1}^{\infty}$ *with* $(k, l) = 1$ *contains infinitely many primes.*

For the proof we need a special kind of Dirichlet series.

An <u>L-function</u> $L(s, \chi)$ is a Dirichlet series

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

where $\chi$ is a Dirichlet character. Since $|\chi(n)| \leq 1$ for all $n$, but $|\chi(n)| = 1$ in an arithmetic progression, we have $\sigma_a = 1$. Since $\chi$ is completely multiplicative, we have an Euler product

$$L(s, \chi) = \prod_{p} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

It follows that

$$\log L(s, \chi) = \textstyle\sum_p \sum_{m=1}^{\infty} \frac{1}{m} \displaystyle\frac{(\chi(p))^m}{p^{ms}} \qquad \text{for } \sigma > 1.$$

On using (5.5) we obtain, for $(k,l)=1$,

$$\frac{1}{\varphi(k)}\sum_{\chi}\overline{\chi}(l)\log L(s,\chi)= \tag{5.6}$$

$$=\sum_p\sum_{m=1}^{\infty}\frac{1}{mp^{ms}}\frac{1}{\varphi(k)}\sum_{\chi}\overline{\chi}(l)\chi(p^m)=\sum_{m=1}^{\infty}\sum_{p:p^m\equiv l(modk)}\frac{1}{mp^{ms}}\quad\text{for }\sigma>1.$$

Let $s=\sigma>1$. The terms with $m\geq 2$ contribute at most $\sum_p\sum_{m=2}^{\infty}\frac{1}{mp^{m\sigma}}<\sum_p\sum_{m=2}^{\infty}\frac{1}{p^m}<\sum_{n=2}^{\infty}\sum_{m=2}^{\infty}\frac{1}{n^m}=\sum_{n=2}^{\infty}\frac{1}{n(n-1)}=1$ to this sum. Hence the right-hand side of (5.6) is $\sum_{p:p\equiv l(modk)}\frac{1}{p^{\sigma}}+O(1)$. Our goal is to prove that this tends to $\infty$ as $\sigma\downarrow 1$, as this implies that there exist infinitely many primes $p$ with $p\equiv l(\text{mod}k)$.

The contribution of $\chi_0$ to the left-hand side of (5.6) is

$$\frac{1}{\varphi(k)}\log L(\sigma,\chi_0)=\frac{1}{\varphi(k)}\sum_p\sum_{m=1}^{\infty}\frac{1}{mp^{m\sigma}}-\frac{1}{\varphi(k)}\sum_{p|k}\sum_{m=1}^{\infty}\frac{1}{mp^{m\sigma}}.$$

Since the first term on the right-hand side tends to $\infty$ as $\sigma\downarrow 1$, whereas the second term is bounded, it suffices to prove that, for $\chi\neq\chi_0$, $\log L(\sigma,\chi)$ is bounded as $\sigma\downarrow 1$.

First we note that for $\chi\neq\chi_0$ the series $L(s,\chi)=\sum_{n=1}^{\infty}\frac{\chi(n)}{n}$ converges for $\sigma>0$. Because of (5.2) the partial sums $\left\{\sum_{n=1}^{N}\chi(n)\right\}_{N=1}^{\infty}$ are bounded. The factors $\frac{1}{n^s}$ tend in absolute value monotonically to 0. Hence, by Dirichlet's criterion (see Exercise 3), $\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}$ converges for $\sigma>0$ and, by the properties of Dirichlet series, $L(s,\chi)$ is an analytic function on $\sigma>0$ for $\chi\neq\chi_0$.

Consider the product

$$P(s):=\prod_{\chi}L(s,\chi)$$

where $\chi$ runs through all characters mod $k$. If $L(1,\chi)=0$ for at least one character $\chi\neq\chi_0$, then the simple zero at $s=1$ of $L(s,\chi_0)=\zeta(s)\prod_{p|k}\left(1-\frac{1}{p^s}\right)$ would be cancelled by the zero of $L(s,\chi)$ at $s=1$ and $P(s)$ would be regular for $\sigma>0$. Otherwise $L(1,\chi)\neq 0$ and $\log L(s,\chi)\to\log L(1,\chi)$ for all $\chi\neq\chi_0$. It is therefore sufficient to prove that

$$P(s)\text{ is not regular for }\sigma>0, \tag{5.7}$$

and the proof of an arithmetic theorem has been reduced to a result on analytic functions.

Suppose $P(s)$ is regular for $s=\sigma>0$. Then (5.6) holds for $\sigma>0$. On applying it with $l=1$ we obtain

$$\log P(\sigma)=\sum_{\chi}\log L(\sigma,\chi)=\varphi(k)\sum_{m=1}^{\infty}\sum_{p:p^m\equiv 1(modk)}\frac{1}{mp^{m\sigma}}\quad(\sigma>0).$$

Note that if $\varphi(k) = h$, then $p^h \equiv 1 \pmod{k}$ for every prime $p$ with $p \not| k$. Hence, on taking $\sigma = \frac{1}{h}$, and only considering $m = h$, we obtain

$$\log P(\frac{1}{h}) \geq \varphi(k) \sum_{p \not| k} \frac{1}{hp} = \frac{\varphi(k)}{h} \left( \sum_p \frac{1}{p} - \sum_{p|k} \frac{1}{p} \right) = \infty.$$

Thus $P(s)$ is not regular for $\sigma > 0$, and therefore Dirichlet's theorem has been proved.

## 5.3 Siegel zeros

It is rather simple to prove that $\zeta(s)$ has no real zeros $\sigma$ with $0 < \sigma < 1$. However, the generalization to $L$-functions is not known. Since remainder terms are closely related to zero-free regions, estimates for real zeros of $L$-functions are important in this context. Since $L'(\sigma, \chi)$ can be well bounded, estimates for the zeros follow by the mean value theorem from lower bounds for $|L(1, \chi)|$.

We call the real zeros of $L$-functions Siegel zeros. The characters assuming non-real values do not give the worst problems: there exists a positive absolute constant $c$ such that if $\chi$ is a complex character to the modulus $k$, any zero $\beta + i\gamma$ of $L(s, \chi)$ satisfies

$$\beta < 1 - \frac{c}{\log k + \log(|\gamma| + 2)}.$$

When dealing with real characters $\chi$, we may assume $\chi \neq \chi_0$. Actually, there is at most one zero close to 1: there exists an absolute constant $c > 0$ such that the only possible zero $\beta + i\gamma$ of $L(s, \chi)$ satisfying $|\gamma| < \frac{c}{\log k}$, $\beta > 1 - \frac{c}{\log k}$ is a single simple real zero $\beta_1$.

Landau showed that values of $k$ with such Siegel zeros are rare: if $\chi_1$ and $\chi_2$ are distinct real primitive characters to the moduli $k_1, k_2$ respectively with real Siegel zeros $\beta_1, \beta_2$, then

$$\min(\beta_1, \beta_2) < 1 - \frac{c}{\log k_1 k_2} \quad (c \text{ positive constant}).$$

The special Siegel zero with $\beta > 1 - \frac{c}{\log k}$ occurs in formulas for the remainder. For example, denote by $\pi(x; k, l)$ the number of primes at most $x$ which is $\equiv l \pmod{k}$. Then, for $k \leq \exp[(\log x)^{1/2}]$

$$\pi(x; k, l) = \frac{x}{\varphi(k) \log x} - \frac{\overline{\chi_1}(l) x^{\beta_1}}{\varphi(k) \beta_1 \log x} + O\left\{ x \exp[-C(\log x)^{1/2}] \right\}$$

where $\chi_1$ is the single real character mod $k$ which has a Siegel zero $\beta_1 > 1 - \frac{c}{\log k}$.

The best effective estimate for $\beta_1$ is $\beta_1 < 1 - \frac{c}{k^{1/2}(\log k)^2}$. This implies

$$\pi(x; k, l) = \frac{x}{\varphi(k) \log x} + O\left\{ x \exp\left(-c(\log x)^{1/2}\right) \right\} \tag{5.8}$$

only for $k \leq (\log x)^{1-\delta}$ for some fixed $\delta > 0$. This shows that the Prime Number Theorem with error term can be extended to arithmetic progressions and that the primes are equally distributed among the primitive residue classes mod $k$.

## 5.4 Homework for Chapter 5

1. a) Determine all characters mod 12.
   b) Determine all characters mod 40.
   c) Prove that $\sum_{l=0}^{m-1} e^{2\pi i \frac{la}{m}} = \begin{cases} m & if \quad a \equiv 0 \pmod{m}, \\ 0 & otherwise. \end{cases}$
2. a) Prove that if $\chi$ is a character mod $k$, then $\overline{\chi}$ is also a character mod $k$.
   b) Prove that $\prod_\chi |L(\sigma, \chi)| > 1$ for $\sigma > 1$.

These exercises together with those of Chapter 4 have to be done for Thursday 15 November.

## 5.5 Further exercises for Chapter 5

1. a) Prove that $L(s, \chi_0) = \zeta(s) \prod_{p|k} \left(1 - \dfrac{1}{p^s}\right)$ $(\sigma > 1)$.
   b) Compute $\mathrm{res}_{s=1} L(s, \chi_0)$.
2. Prove Dirichlet's convergence criterion by partial summation: if $(a_n)$ is a complex sequence such that its partial sums are bounded and $(b_n)$ is a sequence of positive real numbers the absolute values of which tend monotonically to infinity, then the series $\sum_{n=1}^{\infty} \frac{a_n}{b_n}$ converges.
3. Prove that if $P(s)$ is regular, then (5.6) holds for $\sigma > 0$.
4. Prove that $L(s, \chi) \neq 0$ for $\sigma > 1$.
5. Let $\chi$ be the character mod 4 with $\chi \neq \chi_0$. Check whether $L(s, \chi)$ has a Siegel zero.

# Chapter 6

# Sieve methods.

Literature:

H. Halberstam, H.-E. Richert, Sieve Methods, Academic Press, 1974.

C. Hooley, Applications of Sieve Methods to the Theory of Numbers, Cambridge Tract 70, Cambridge University Press, 1976.

W. Schwarz, Einführung in Siebmethoden der analytischen Zahlentheorie, Bibliographisches Institut Mannheim, 1974.

G. Greaves, Sieves in Number Theory, Springer, 2001.

## 6.1 Some classical additive problems

Sieve methods and the Hardy-Littlewood circle method belong to the most successful methods when dealing with problems of the following types. We deal with them in Chapter 6 and 7, respectively. The simplest sieve method is the sieve of Eratosthenes by which all the composite numbres are sieved out so that the primes are left.

Goldbach's problem. In two letters to Euler in 1742, Goldbach conjectured that every even integer $> 2$ is the sum of two primes and that every odd integer $> 5$ is the sum of three primes. The latter statement was asymptotically solved by Vinogradov in 1937. He proved by the Hardy-Littlewood circle method that every odd integer $> N_0$ is the sum of three primes, and even that the number of such presentations of odd $N$ is $>> N^2$. In 1922 Hardy and Littlewood had applied their method to prove such results under assumption of the GRH. The attacks to the still open binary problem for even integers are closely related to attacks to the Twin prime problem. Chen has proved that the number of primes $p$ such that $N-p$ is composed of at most two primes is $>> \frac{N}{(\log N)^2}$. Chen used a sieve method and multiplicative number theory. By the Circle method it can be shown that the number $E(N)$ of even numbers $n \leq N$ for which $n$ is not the sum of two primes satisfies $E(N) << \frac{N}{(\log N)^A}$ for some constant $A$.

Another classical problem is as follows. Let $P(x)$ denote any irreducible polynomial which represents integer values at integer points which have no common nontrivial divisor. Does $P(x)$ represent a prime for infinitely many integers $x$? The classical example is $x^2 + 1$. Revealed posthumously as little more than a fragment in one of Chebyshev's manuscripts, it was first published and proved by Markov in 1895 that the greatest prime factor $P_x$ of $\prod_{n \le x}(n^2+1)$ satisfies $P_x/x \to \infty$ as $x \to \infty$. Hooley proved by combining Chebyshev's method with a sieve method in 1967 that $P_x > x^{11/10}$ for any irreducible quadratic polynomial $x^2 - D$ with $D \in \mathbb{Z}$.

By another sieve Hooley obtained a conditional result for Artin's conjecture on primitive roots. Artin conjectured that for any given integer $a$ other than $1, -1$ or a perfect square, the number $N_a(x)$ of primes $p \le x$ for which $a$ is a primitive root modulo $p$ satisfies

$$N_a(x) \sim \frac{A(a)x}{\log x} \quad (x \to \infty)$$

with $A(a) > 0$. Hooley proved this (with a revised value of $A(a)$ proposed by Heilbronn) subject to the assumption that the Riemann hypothesis holds for Dedekind zeta functions over certain Kummerian fields.

## 6.2 Sieve methods

Let $A$ be a finite set of integers and $P$ a set of primes. A sieve method deals with estimates for the number $S$ of elements from $A$ that are not divisible by any prime from $P$. To have nontrivial estimates for $S$, the set $A$ should be regular in some way. We require that the number $|A_d|$ of elements of

$$A_d := \{a \in A | a \equiv 0 \pmod{d}\}$$

should satisfy

$$|A_d| = \frac{\omega_0(d)}{d}|A| + r_d \qquad \text{for } d \text{ squarefree}$$

where $\omega_0(d)$ is a multiplicative function and $|r_d|$ is small. Hence, for every $d$ the fraction of elements of $A$ divisible by $d$ should be almost a multiple of $1/d$.

For example, if $A = (x-y, x]$, then

$$|A_d| = \frac{1}{d}y + r_d \qquad \text{with } |r_d| \le 1$$

and if $A = \{n(n+2)|1 \le n \le x\}$ then $|A_p| = \frac{2}{p}x + r_p$ for $p > 2$ prime and

$$|A_d| = \frac{2^{\omega(d)}}{d}x + r_d \qquad \text{with } |r_d| \le \omega_0(d) \tag{6.1}$$

for odd $d$. Recall that $\omega(d) = \sum_{p|d} 1$.

We are interested in $S(A,P,z) := |\Big\{a \in A|\gcd\left(a, \prod_{p \in P, p \le z} p\right) = 1\Big\}|$.

We expect that, under regularity conditions,

$$S(A,P,z) \approx |A| \prod_{p\in P, p\leq z} \left(1 - \frac{\omega_0(p)}{p}\right) =: |A|V(z).$$

Sieve results provide upper and lower bounds for $S(A,P,z)$ in terms of $|A|V(z)$. The method is usually elementary, combinatorial and complicated. An estimate which is sufficient for several applications is referred to as Brun's sieve (HR, p.68). It states that if there is a constant $c_0$ such that

$$\omega_0(p) \leq c_0 \tag{6.2}$$

for all $p \in P, p \leq z$, and moreover

$$|R_d| \leq \omega_0(d) \qquad \text{if } \mu(d) \neq 0, \gcd(d, \textstyle\prod_{p\in P, p\leq z} p) = 1, \tag{6.3}$$

then

$$S(A,P,z) << \ _{c_0} \ | \ A \ | \prod_{p<z} \left(1 - \frac{\omega_0(p)}{p}\right) \ \ if \ \ z \leq |A|^{c_0}. \tag{6.4}$$

As an application we have the following general result on arithmetic progressions:

Let $a_i, b_i$ $(i = 1, \ldots, g)$ be pairs of integers satisfying $\gcd(a_i, b_i) = 1$ for $i = 1, \ldots, g$. Suppose

$$E := \prod_{i=1}^{g} a_i \prod_{1\leq r<s\leq g} (a_r b_s - a_s b_r) \neq 0.$$

Then, for $x \geq 2$,

$$| \ \{m \leq x : |a_i m + b_i| \ \text{ is prime for } \ i = 1, \ldots, g\} \ |$$

$$<<_g \frac{x}{(\log x)^g} \prod_{p|E} \left(1 - \frac{1}{p}\right)^{\omega_0(p) - g}$$

where $\omega_0(p)$ denotes the number of solutions of $\prod_{i=1}^{g}(a_i m + b_i) \equiv 0 \pmod{p}$.

The latter statement is proved by applying the former statement with $P$ is the set of all primes, $x = |A|, z \approx \sqrt{|A|}$. The proof involves some estimations too complicated to be presented here.

We give some straightforward applications.

1. Take $a_1 = a_2 = 1$, $b_1 = 0$, $b_2 = 2$, $g = 2$.
   It follows that the number of primes $p \leq x$ such that $p, p+2$ is a twin prime is $<< \dfrac{x}{(\log x)^2}$, whence $\sum' \frac{1}{p} < \infty$.
   This has been proved by Brun.

2. Take $g = 2$, $a_1 = 1$, $b_1 = 0$, $a_2 = 2$, $b_2 = -1$.
   It follows that the number of primes $p \leq x$ such that $2p - 1$ is a prime is $<< \dfrac{x}{(\log x)^2}$.

3. Take $g = 1$, $a_1 = k$, $b_1 = l$. We assume $\gcd(k, l) = 1$.
It follows that $\omega_0(p) = 0$ for $p|k$ whence

$$\pi(x, k, l) << \frac{x/k}{\log(x/k)} \prod_{p|k} \left(1 - \frac{1}{p}\right)^{-1} = \frac{x}{\varphi(k) \log(x/k)}.$$

This is a weak form of the so-called Brun-Titchmarsh theorem.

4. Take $g = 2$, $a_1 = 1$, $b_1 = 0$, $a_2 = 1$, $b_2 = -2N$.
It follows that $\omega_0(p) = 1$ for $p|N$ whence the number of prime pairs $(p, q)$ such that $p + q = 2N$ is

$$<< \frac{2N}{(\log 2N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right).$$

Selberg introduced a sieve method which in some ways simplifies and improves upon Brun's sieve. Selberg presented both an upper bound and a lower bound sieve. The problem is that the summation $\sum_{d|n}^{*}$ extends over so many terms that an estimation of the remainder exceeds the main term for this reason. The basic idea is to find upper and lower bounds for the remainder terms such that the remainder consists of much fewer non-zero terms. There exists now a great variety of sieve methods. See for example Halberstam-Richert or Hooley.

## 6.3 The large sieve

In the sieve methods treated up to now a number from $A$ was only sifted out if it was divisible by some prime $p \in P$. In the large sieve for each prime $p \in P$ all numbers of $A$ are sifted out which belong to one of $\omega_0(p)$ residue classes mod $p$.

One can use the large sieve to measure the uniform distribution of the integers from some set $A$ over the various residue classes. Let $A$ be a set of positive integers from the interval $(X - Y, X]$. For any positive integer $q$ define

$$A(q, h) = \sum_{a \in A, a \equiv h (\mathrm{mod} q)} 1,$$

so that $\sum_{h=1}^{q} A(q, h) = |A|$. The expression

$$D_p := \sum_{h=1}^{p} \left\{ A(p, h) - \frac{|A|}{p} \right\}^2$$

is a measure for the uniform distribution of $A$ among the residue classes mod $p$. Define

$$S(x) = \sum_{a \in A} e^{2\pi i a x}.$$

Then

$$\sum_{n=1}^{p-1} |S(\frac{n}{p})|^2 = \sum_{a \in A} \sum_{a' \in A} \sum_{n=1}^{p-1} e^{2\pi i (a - a') \frac{n}{p}}.$$

The inner sum is $p-1$ if $a \equiv a' \pmod{p}$ and $-1$ otherwise. Hence the sum is equal to

$$p\sum_{a\in A}\sum_{a'\in A, a\equiv a'(\text{mod} p)} 1 - |A|^2 = p\sum_{h=1}^{p}\left(\sum_{a\in A, a\equiv h(\text{mod} p)} 1\right)^2 - |A|^2$$

so that

$$pD_p = p\sum_{h=1}^{p} A^2(p,h) - |A|^2 = \sum_{n=1}^{p-1} |S(\frac{n}{p})|^2. \tag{6.5}$$

We shall be concerned with nontrivial estimates of the sum $\sum_{p\le Y_0} pD_p$. Note that this is a sum of type $\sum_{r=1}^{R} |S(x_r)|^2$ where the real numbers $x_r$ are well-separated in $[0,1]$ in the sense that $|x_r - x_{r'}| \ge \frac{1}{pp'} \ge \frac{1}{Y_0^2}$ for $r \ne r'$. Put $\delta = \dfrac{1}{Y_0^2}$. Put $S_0(x) = S(x)e^{-2\pi i(X-\frac{1}{2}Y)x}$. Then $|S_0(x)| = |S(x)|$ for all $x$. Since

$$S_0^2(x) - S_0^2(x_r) = 2\int_{x_r}^{x} S_0(y)S_0^{'}(y)dy,$$

we have

$$|S_0(x_r)|^2 \le |S_0(x)|^2 + 2|\int_{x_r}^{x} |S_0(y)S_0^{'}(y)|dy|.$$

Integrate over the interval $\left(x_r - \frac{1}{2}\delta, x_r + \frac{1}{2}\delta\right)$ to arrive at

$$\delta|S_0(x_r)|^2 \le \int_{x_r-\frac{1}{2}\delta}^{x_r+\frac{1}{2}\delta} |S_0(x)|^2 dx + \delta\int_{x_r-\frac{1}{2}\delta}^{x_r+\frac{1}{2}\delta} |S_0(y)S_0^{'}(y)|dy.$$

Summing over $r$ and using the disjointness of the intervals we obtain

$$\sum_{r=1}^{R} |S_0(x_r)|^2 \le \delta^{-1}\int_0^1 |S_0(x)|^2 dx + \int_0^1 |S_0(y)S_0^{'}(y)|dy.$$

Note that $\int_0^1 |S_0(x)|^2 dx = |A|$. Similarly $\int_0^1 |S_0^{'}(x)|^2 dx \le \pi^2 Y^2 |A|$, since $|a - (X - \frac{1}{2}Y)| \le \frac{1}{2}Y$ for every $a \in A$. By Cauchy's inequality we obtain

$$\sum_{p\le Y_0} pD_p = \sum_{r=1}^{R} |S(x_r)|^2 = \sum_{r=1}^{R} |S_0(x_r)|^2 \le \delta^{-1}|A| + |A|^{1/2}\cdot\pi Y|A|^{1/2} = (Y_0^2 + \pi Y)|A|.$$

This proof is due to Gallagher. Bombieri and Davenport obtained the sharper estimates $2\max(Y, Y_0^2)|A|$ and $(Y^{1/2} + Y_0)^2|A|$, respectively. The saving compared with the trivial estimate $2YY_0|A|$ is very considerable if $Y_0 \le Y^{1/2}$.

## 6.4 Homework for Chapter 6

1. Let $f$ be an arithmetic function with $f(1) = 1$ and $m, n, p$ positive integers with $p$ squarefree.

a) Prove that $\dfrac{1}{\mathrm{lcm}(m,n)} = \dfrac{1}{mn}\sum_{k|\ \gcd(m,n)} \varphi(k)$.

b) Prove that $\sum_{m|p,n|p} \dfrac{f(m)f(n)}{\mathrm{lcm}(m,n)} = \sum_{k|p} \varphi(k)y_k^2$, where $y_k = \sum_{n \text{ with } k|n,n|p} \dfrac{f(n)}{n}$.

c) Prove that $f(n) = n\sum_{k \text{ with } n|k,k|p} \ y_k\mu\left(\dfrac{k}{n}\right)$.

2. Compute a lower bound for $\prod_{p\le x}\left(1-\dfrac{1}{p}\right)^{-1}$ which tends to $\infty$ when $x \to \infty$.

These exercises also have to be done for Thursday 29 November.

## 6.5 Further exercises for Chapter 6

1. Prove formula (1.1).

2. a) Prove that $\sum_{d|n} \mu(d)\dfrac{\omega_0(d)}{d} = \prod_{p|n}\left(1-\dfrac{\omega_0(p)}{p}\right)$.

   b) Prove that $\sum_{d|n} \mu(d)|A_d| = |A|\prod_{p|n}\left(1-\dfrac{\omega_0(p)}{p}\right) + \sum_{d|n}\mu(d)r_d$.

   c) Check that $S(A,P,z) = \sum_d^* \mu(d)|A_d|$ where the summation extends over all $d$ which are composed of primes $p \in P$ with $p \le z$.

3. Give an upper bound for the number of primes $p \le x$ such that both $kp+1$ and $lp+1$ are primes for given integers $k, l$ with $1 < k < l$.

# Chapter 7

# The Hardy-Littlewood circle method.

Literature:

R. C. Vaughan, The Hardy-Littlewood Method, Cambridge University Press, 1981.

The Hardy-Littlewood method or circle method has been applied to several classical number theoretic problems of an additive type. We mention the most famous ones.

A classical problem which has been settled is Waring's problem. In 1770 Waring asserted that every positive integer is a sum of at most nine positive integral cubes, also a sum of at most 19 biquadrates, and so on. The first proof that every integer is the sum of four squares was given by Lagrange in the same year 1770. Let $g(k)$ be the least $s$ such that every positive integer is a sum of at most $s$ $k$th powers of positive integers. In 1909 Hilbert showed by a combinatorial argument that $g(k)$ is finite for all $k$. It is easy to show that

$$g(k) \geq 2^k + \left[ \left( \frac{3}{2} \right)^k \right] - 2 \quad \text{for all} \quad k \geq 3. \tag{7.1}$$

It is very plausible that this always holds with equality, and the current state is as follows. Suppose that $k \neq 4$. It has been shown that when

$$2^k \left\{ \left( \frac{3}{2} \right)^k \right\} + \left[ \left( \frac{3}{2} \right)^k \right] \leq 2^k \tag{7.2}$$

one has equality sign in (7.1), but when (7.2) is false one has either

$$g(k) = 2^k + \left[ \left( \frac{3}{2} \right)^k \right] + \left[ \left( \frac{4}{3} \right)^k \right] - 2$$

or

$$g(k) = 2^k + \left[ \left( \frac{3}{2} \right)^k \right] + \left[ \left( \frac{4}{3} \right)^k \right] - 3$$

according as

$$\left[\left(\frac{4}{3}\right)^k\right]\left[\left(\frac{3}{2}\right)^k\right]+\left[\left(\frac{4}{3}\right)^k\right]+\left[\left(\frac{3}{2}\right)^k\right]$$

is equal to $2^k$ or larger than $2^k$.

In 1964 Stemmerl verified on a computer that (7.2) holds whenever $k \leq 200000$ and in 1957 Mahler showed that there are only finitely many $k$ for which (7.2) does not hold. No exceptions are known, and unfortunately the method (Thue-Siegel-Roth method) will not give an upper bound for the solutions. In 1986 Balasubramanian, Deshouillers and Dress proved that every integer $> 10^{367}$ is a sum of nineteen fourth powers.
The more fundamental problem is that of estimating the number $G(k)$, defined for $k \geq 2$ to be the least $s$ such that every sufficiently large natural number is the sum of at most $s$ $k$th powers. It turns out that $G(k)$ is much smaller than $g(k)$ when $k$ is large. The only known values of $G(k)$ are $G(2) = 4$ and $G(4) = 16$, the latter result due to Davenport in 1939. In 1943 Linnik showed that $G(3) \leq 7$, Davenport proved $G(5) \leq 23, G(6) \leq 36$ in 1942. Vinogradov has shown that $G(k) \leq (2 + o(1))k \log k$ as $k \to \infty$. It is conjectured that $G(3) = 4$.

Another classical problem set starts with van der Waerden who proved in 1927 that given natural numbers $l, r$ there exists an $n_0(l, r)$ such that if $n \geq n_0(l, r)$ and $\{1, 2, \ldots, n\}$ is partitioned into $r$ sets, then at least one set contains $l$ terms in arithmetic progression.

For an arbitrary set $\mathcal{A}$ of natural numbers, let

$$A(n) = A(n, \mathcal{A}) = \sum_{a \leq n, a \in \mathcal{A}} 1, \quad D(n) = D(n, \mathcal{A}) = \frac{1}{n}A(n) \tag{7.3}$$

and write $\underline{d}$ and $\overline{d}$ for the lower and upper asymptotic densities of $\mathcal{A}$,

$$\underline{d} = \underline{d}(\mathcal{A}) = \text{liminf}_{n\to\infty} D(n) \quad \text{and} \quad \overline{d} = \overline{d}(\mathcal{A}) = \text{limsup}_{n\to\infty} D(n) \tag{7.4}$$

respectively. When $\underline{d} = \overline{d}$ let $d = d(\mathcal{A})$ denote their common value, the asymptotic density of $\mathcal{A}$. Erdős and Turán (1936), in discussing the nature of the known proofs of van der Waerden's theorem, conjectured that every set $\mathcal{A}$ with $\overline{d}(\mathcal{A}) > 0$ contains arbitrarily long arithmetic progressions. An equivalent assertion is that if there is an $l$ such that $\mathcal{A}$ contains no arithmetic progression of $l$ terms, then $d(\mathcal{A}) = 0$.

The first non-trivial case is $l = 3$. The initial breakthrough was made by Roth (1952, 1953, 1954) in establishing this case by an ingenious adaptation of the Hardy-Littlewood method. We shall give his proof later in this chapter.

By a different method, Szemerédi (1969) proved the conjecture for $l = 4$, and Roth (1972) gave an alternative proof by an approach related to that of his earlier method.

In 1975, Szemerédi established the general case. Unfortunately Szemerédi's proof uses van der Waerden's theorem. More recently Furstenburg (1977) has given a proof of Szemerédi's theorem based on ideas from ergodic theory. (This is one of the subjects of a course which is given by Frank Redig.) In 2002 Gowers gave still another proof of Szemerédi's theorem giving better quantitative estimates.

It follows that sequences which are sufficiently dense contain arbitrarily long arithmetic progressions. It was a old conjecture that the primes are such a sequence. Quite recently this was proved by Green and Tao. Erdős conjectured that $(a_n)_{n=1}^{\infty}$ is such a sequence when $\sum a_n$ diverges.

Let $M(n)$ denote the largest number of elements which can be taken from $\{1, 2, \dots, n\}$ with no three of them in progression. Let

$$\mu(n) = n^{-1}M(n).$$

Then Roth's theorem is the assertion $\lim_{n\to\infty} \mu(n) = 0$. As the following lemma shows, it is quite easy to prove that the limit exists. Its value is another matter.

**Lemma 7.1.** *The limit* $\lim_{n\to\infty} \mu(n)$ *exists. Also, for* $m \geq n$ *one has* $\mu(m) \leq 2\mu(n)$.

**Proof.** It is a trivial consequence of the definition of $M$ that

$$M(m+n) \leq M(m) + M(n).$$

Hence

$$M(m) \leq \left[\frac{m}{n}\right] M(n) + M\left(m - n\left[\frac{m}{n}\right]\right) \leq \frac{m}{n}M(n) + n.$$

Therefore $\mu(m) \leq \mu(n) + n/m$, so that

$$\limsup_{m\to\infty} \mu(m) \leq \mu(n)$$

whence

$$\limsup_{m\to\infty} \mu(m) \leq \liminf_{n\to\infty} \mu(n).$$

Also, when $m \geq n$, $M(m) \leq (\frac{m}{n} + 1)M(n) \leq 2M(n)\frac{m}{n}$. □

The following theorem not only shows that the limit is 0, but gives a bound for the size of $M(n)$.

**Theorem 7.1.** (Roth) *Let* $n \geq 3$. *Then* $\mu(n) \ll (\log\log n)^{-1}$.

Choose $\mathcal{M} \subset \{1, 2, \dots, n\}$ so that $\sharp\mathcal{M} = M(n)$ and no three elements of $\mathcal{M}$ are in arithmetic progression. Let

$$f(\alpha) = \sum_{m\in\mathcal{M}} e(\alpha m).$$

Then

$$M(n) = \int_0^1 f(\alpha)^2 f(-2\alpha)d\alpha \tag{7.5}$$

since the right-hand side is the number of solutions of $m_1 + m_2 = 2m_3$ with $m_1, m_2, m_3 \in \mathcal{M}$ and, by the construction of $\mathcal{M}$, such solutions can only occur when $m_1 = m_2 = m_3$.

Let $\kappa$ denote the characteristic function of $\mathcal{M}$, so that

$$f(\alpha) = \sum_{x=1}^{n} \kappa(x)e(\alpha x). \tag{7.6}$$

Suppose that

$$m < n, \tag{7.7}$$

and consider

$$v(\alpha) = \mu(m) \sum_{x=1}^{n} e(\alpha x) \tag{7.8}$$

and

$$E(\alpha) = v(\alpha) - f(\alpha).$$

Then

$$E(\alpha) = \sum_{x=1}^{n} c(x)e(\alpha x) \tag{7.9}$$

with

$$c(x) = \mu(m) - \kappa(x). \tag{7.10}$$

The idea of the proof is that, if $M(n)$ is close to $n$, then

$$\int_0^1 f(\alpha)^2 f(-2\alpha)d\alpha$$

ought to be closer to $M(n)^2$ than to $M(n)$. To show this, one first of all uses the disorderly arithmetical structure of $\mathcal{M}$ to replace $f$ by $v$ with a relatively small error. It is a fairly general principle that sums of the form

$$\sum_{x \le n, x \in \mathcal{A}} e(\alpha x)$$

tend to have large peaks at $a/q$ when the elements of $\mathcal{A}$ are regularly distributed in residue classes modulo $q$. Note that $v(\alpha)$ has its peaks at the integers.

Let

$$F(\alpha) = \sum_{z=0}^{m-1} e(\alpha z). \tag{7.11}$$

**Lemma 7.2.** *Let $q$ be a natural number with $q < n/m$, and for $y = 1, 2, \ldots, n - mq$ let*

$$\sigma(y) = \sigma(y; m, q) = \sum_{x=0}^{m-1} c(y + xq). \tag{7.12}$$

*Then*

$$\sigma(y) \geq 0 \quad (y = 1, 2, \ldots, n - mq) \tag{7.13}$$

*and*

$$F(\alpha q)E(\alpha) = \sum_{y=1}^{n-mq} \sigma(y)e(\alpha(y + mq - q)) + R(\alpha) \tag{7.14}$$

*where $R(\alpha)$ satisfies*

$$|R(\alpha)| < 2m^2 q. \tag{7.15}$$

**Proof.** By collecting together the terms in the product $FE$ for which $x + zq = h + mq - q$ one obtains

$$F(\alpha q)E(\alpha) = \sum_{h=1+q-mq}^{n} e(\alpha(h + mq - q))$$

$$\times \sum_{z=0 \text{ with } 1 \leq h+q(m-1-z) \leq n}^{m-1} c(h + q(m - 1 - z)).$$

The innermost sum is at most $m$ in absolute value, and so the total contribution from the terms with $h \leq 0$ and $h > n - mq$ does not exceed, in absolute value, $m(mq + (m - 1)q) < 2m^2 q$. For the remaining values of $h$ one has $1 \leq h + q(m - 1 - z) \leq n$ for all $z$ in the interval $[0, m - 1]$. This gives (7.14) and (7.15).

By (7.10) and (7.12),

$$\sigma(y) = M(m) - \sum_{x=0}^{m-1} \kappa(y + xq).$$

Let

$$r = \sum_{x=0}^{m-1} \kappa(y + xq).$$

Then $r$ is the number of elements of $\mathcal{M}$ among $y, y + q, \ldots, y + (m - 1)q$. Let these elements be $y + x_1 q, \ldots, y + x_r q$. Then no three are in arithmetic progression. Hence no three of $x_1, \ldots, x_r$ are in arithmetic progression. Likewise for $1 + x_1, \ldots, 1 + x_r$. Moreover $1 + x_j \leq m$. Hence $r \leq M(m)$, which gives (7.13). □

**Lemma 7.3.** *Suppose that $2m^2 < n$. Then, for every real number $\alpha$,*

$$|E(\alpha)| < 2n(\mu(m) - \mu(n)) + 16m^2.$$

**Proof.** By a theorem of Dirichlet, there exist $a, q$ such that $(a, q) = 1$, $1 \leq q \leq 2m$ and $|\alpha - a/q| \leq 1/(2qm)$. Then

$$F(\alpha q) = F(\alpha q - a) = F(\beta)$$

where $|\beta| \leq 1/(2m)$. Hence, by (7.11),

$$|F(\alpha q)| =| \frac{\sin \pi m\beta}{\sin \pi\beta} |\geq \frac{2m}{\pi}.$$

Thus, by Lemma 7.2,

$$\frac{1}{2}m|E(\alpha)| \leq \frac{2}{\pi}m|E(\alpha)| \leq |F(\alpha q)E(\alpha)|$$

$$< \sum_{y=1}^{n-mq} \sigma(y) + 2m^2q < mE(0) + 8m^3.$$

Moreover, by (7.9) and (7.10),

$$E(0) = \sum_{x=1}^{n} (\mu(m) - \kappa(x)) = n(\mu(m) - \mu(n)).$$

The lemma follows at once. □

**Proof of Theorem 7.1.** Let

$$I = \int_0^1 f(\alpha)^2 v(-2\alpha)d\alpha. \tag{7.16}$$

Then, by (7.6) and (7.8),

$$I = \sum_{a\in\mathcal{M}} \sum_{b\in\mathcal{M},2|a+b} \mu(m).$$

Thus, if $M_1$ is the number of odd elements of $\mathcal{M}$ and $M_2$ the number of even elements, so that $M_1 + M_2 = M(n)$, then

$$I = \mu(m)(M_1^2 + M_2^2) \geq \frac{1}{2}\mu(m)M(n)^2. \tag{7.17}$$

By (7.5) and (7.16),

$$|M(n) - I| \leq \left(\max_{\alpha} |E(\alpha)|\right) \int_0^1 |f(\alpha)|^2 d\alpha.$$

Therefore, by Lemma 7.3 and Parseval's identity, when $2m^2 < n$ one has

$$|M(n) - I| \leq (2n(\mu(m) - \mu(n)) + 16m^2)M(n).$$

Hence, by (7.17),

$$\mu(m)\mu(n) \leq 4(\mu(m) - \mu(n)) + 34m^2n^{-1} \quad (2m^2 < n). \tag{7.18}$$

Letting $n \to \infty$ and then $m \to \infty$ shows that $\tau = \lim_{n\to\infty} \mu(n)$ satisfies $\tau^2 \leq 0$. To establish the quantitive version of this, let

$$\lambda(x) = \mu\left(2^{3^x}\right).$$

By Lemma 7.1, it suffices to show that $\lambda(2x) \ll x^{-1}$. By (7.18),

$$\lambda(y)\lambda(y+1) \leq 4(\lambda(y) - \lambda(y+1)) + 34 \times 2^{-3^y}.$$

Dividing by $\lambda(y)\lambda(y+1)$, summing over $y = x, x+1, \dots, 2x-1$ and appealing to Lemma 7.1 gives one

$$x \leq 4\lambda(2x)^{-1} + 200x\lambda(2x)^{-2}2^{-3^x}.$$

When $\lambda(2x) > 1/x$ the second term on the right is $< \frac{1}{2}x$ for $x$ sufficiently large, so that $\lambda(2x) < 8/x$, which gives the desired conclusion. □

## 7.1 Homework for Chapter 7

1. Let $R_s(m,n)$ be the number of representations of $m$ as the sum of $s$ non-negative $k$th powers, none of which exceed $n$. Put $f(\alpha) = \sum_{m=1}^{N} e(\alpha m^k)$, where $N = [n^{1/k}]$.
   a) Prove that
   $$(f(\alpha))^s = \sum_{m=s}^{sN^k} R_s(m,n)e(\alpha m).$$
   b) Prove that
   $$R_s(m,n) = R_s(m,m) \quad \text{for} \quad m \leq n.$$
   c) Prove that
   $$R_s(n,n) = \int_0^1 (f(\alpha))^s e(-\alpha n)d\alpha.$$

2. Let $P, Q$ denote real numbers with $P > 1,\ Q \geq 2P$. Show that the intervals
   $$\left\{\alpha : |\alpha - a/q| \leq \frac{1}{qQ}\right\}$$
   with $q \leq P$ and $(a,q) = 1$ are pairwise disjoint.

These exercises also have to be done for Thursday 29 November.

## 7.2 Further exercises for Chapter 7

1. Prove Dirichlet's theorem that for given $m$ there exist $a, q$ such that $(a,q) = 1$, $1 \leq q \leq m$ and $|\alpha - a/q| \leq 1/(qm)$.

2. Generalize Lemma 7.1 to arithmetic progressions larger than 3.

3. Compute the lower and upper asymptotic densities of
   a) the primes b) the squares c) the arithmetic progression $\{a + nd\}_{n=1}^{\infty}$.

# Chapter 8

# Smooth numbers.

Smooth numbers are numbers composed of small primes. These numbers play an important role in some algorithms, e.g. factorization algorithms. The study of smooth numbers is relevant for the complexity analysis of such algorithms. Surveys of results on these numbers can be found in

K. K. Norton, Numbers with small prime factors and the least $k$th power non-residue, Mem. Amer. Math. Soc. No. 106 (1971), AMS, Providence RI, 1971.

A. Hildebrand and G. Tenenbaum, Integers without large prime factors, J. Théor. Nombres Bordeaux 5 (1993), 411-484.

## 8.1 Introduction

Let $\psi(x,y)$ denote the number of positive integers $\leq x$ which are free of prime factors $> y$. The behaviour of $\psi(x,y)$ depends on the range of $y$. Ennola proved in 1969 that for $2 \leq y \leq \sqrt{\log x}$ one has

$$\psi(x,y) = \frac{1}{\pi(y)!} \prod_{p \leq y} \left( \frac{\log x}{\log p} \right) \left( 1 + O\left( \frac{y^2}{\log x \log p} \right) \right). \tag{8.1}$$

See further Section 8.3.

On the other hand, if $y$ is large, then $\psi(x,y) \sim x\rho(u)$ where $u = \dfrac{\log x}{\log y}$ and $\rho(u)$ is defined by

$$\rho(u) = 1 \quad (0 \leq u \leq 1), \quad u\rho'(u) + \rho(u-1) = 0 \quad (u > 1). \tag{8.2}$$

For $u \geq 3$ one has (De Bruijn, 1951, see also Cameron, Erdös, Pomerance

(1983))

$$\rho(u) = \exp\left\{-u\left(\log u + \log\log u - 1 + O\left(\frac{\log\log u}{\log u}\right)\right)\right\}.$$

De Bruijn has proved that $\psi(x,y) \sim x\rho(u)$ holds for $u \le (\log x)^{3/8-\varepsilon}$ $(\varepsilon > 0)$. See further Section 8.2.

The behaviour of $\psi(x,y)$ changes dramatically around $y = \log x$. This change is reflected in a result of De Bruijn published in 1966, which will be mentioned in Section 8.4.
Around 1985 Hildebrand described the asymptotic behaviour of $\psi(x,y)$ itself. It involves implicitly defined functions. The result is mainly of theoretical interest. Some of the older estimates are based on the Buchstab functional equation:

$$\psi(x,y) = \psi(x,z) - \textstyle\sum_{y<p\le z} \psi\left(\frac{x}{p}, p\right) \qquad (1 \le y < z \le x). \tag{8.3}$$

Most of the modern papers on $\psi(x,y)$ are based on the Hildebrand functional equation:

$$\psi(x,y)\log x = \textstyle\int_1^x \dfrac{\psi(t,y)}{t}dt + \sum_{p^m \le x, p\le y} (\log p)\psi\left(\frac{x}{p^m}, y\right) \qquad (x \ge 1, y \ge 1). \tag{8.4}$$

## 8.2 $y$ is large compared to $x$

We shall show that $\psi(x, x^{1/u}) \sim x\rho(u)$ for $0 < u \le 2$. We use the so-called Buchstab functional equation. To see that this formula is correct, observe that the difference $\psi(x,z) - \psi(x,y)$ equals the cardinality of the set of positive integers $\le x$ having greatest prime factor in $(y,z]$. Let $p$ be a prime with $y < p \le z$. Then the cardinality of the set of positive integers $\le x$ having greatest prime factor $p$ equals $\psi(\frac{x}{p}, p)$. Thus

$$\psi(x,z) - \psi(x,y) = \sum_{y<p\le z} \psi\left(\frac{x}{p}, p\right)$$

which is equivalent with (8.3). If we choose $z = x$, we obtain

$$\psi(x,y) = [x] - \sum_{y<p\le x} \left[\frac{x}{p}\right]. \tag{8.5}$$

For $0 < u \le 1$ we have $\psi(x,y) = [x] \sim x = x\rho(u)$ for $0 < u \le 1$.

Let $1 < u \le 2$. Recall that

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

Hence

$$\sum_{x^{1/u}<p\le x} \frac{1}{p} = \log u + O\left(\frac{1}{\log x}\right).$$

Substituting this into (8.5) yields

$$\psi(x, x^{1/u}) = x + O(1) - x \sum_{x^{1/u}<p\le x} \frac{1}{p} + O\left(\frac{x}{\log x}\right) =$$

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

On the other hand, for $1 < u \le 2$, by (8.2),

$$\rho(u) = \int_1^u \rho'(t)dt + 1 = 1 - \int_1^u \frac{\rho(t-1)}{t}dt = 1 - \int_1^u \frac{1}{t}dt = 1 - \log u.$$

Thus $\psi(x, x^{1/u}) = x\rho(u) + O\left(\frac{x}{\log x}\right)$ $(1 < u \le 2)$.

The result for general $u$ is achieved by induction reducing the interval $(m, m+1]$ to the preceding $(m-1, m]$ by applying the Buchstab functional equation. Let $u \in (m, m+1]$. Then, by $\psi\left(\frac{x}{p}, p\right) = \psi\left(\frac{x}{p}, \left(\frac{x}{p}\right)^{\frac{\log p}{\log (x/p)}}\right)$,

$$\psi(x, x^{1/u}) \approx \psi(x, x^{1/m}) - \sum_{x^{1/u}<p\le x^{1/m}} \psi\left(\frac{x}{p}, p\right)$$

$$\approx x\rho(m) - \sum_{x^{1/u}<p\le x^{1/m}} \frac{x}{p}\rho\left(\frac{\log(x/p)}{\log p}\right)$$

$$\approx x\rho(m) - \int_{x^{1/u}}^{x^{1/m}} \frac{x}{t}\rho\left(\frac{\log(x/t)}{\log t}\right) d\pi(t)$$

$$\approx x\rho(m) - \int_{x^{1/u}}^{x^{1/m}} \frac{x}{t}\rho\left(\frac{\log x}{\log t} - 1\right) d\frac{t}{\log t}$$

$$\approx x\rho(m) - x\int_m^u \frac{\rho(\omega-1)}{\omega}d\omega = \rho(u).$$

## 8.3 $y$ is small compared to $x$

We prove formula (8.1) in a weakened form and for fixed $y$. So we want to estimate the number of positive integers $n \le x$ composed of prime factors $\le y$. Then

$$n = p_1^{k_1} \cdots p_r^{k_r}$$

with $p_1 = 2, p_2 = 3, p_3 = 5, \ldots$, $p_r$ is the largest prime $\le x$ and

$$\log n = k_1 \log p_1 + \ldots + k_r \log p_r \le \log x.$$

Hence $r = \pi(y)$. This is the number of integer points $(k_1, \ldots, k_r)$ in the generalized triangle with corner points

$$(0,0,\ldots,0), \left(\frac{\log x}{\log p_1}, 0, \ldots, 0\right), \ldots, \left(0,0,\ldots,0,\frac{\log x}{\log p_r}\right).$$

For large $x$ this number equals the area

$$\frac{1}{r!}\left(\frac{\log x}{\log p_1}\right)\left(\frac{\log x}{\log p_2}\right)\cdots\left(\frac{\log x}{\log p_r}\right) + O(\log x)^{r-1}.$$

Thus,

$$\psi(x,y) = \frac{1}{(\pi(y))!}\prod_{p\le y}\left(\frac{\log x}{\log p}\right)\left(1 + O\left(\frac{1}{\log x}\right)\right).$$

## 8.4 A result of De Bruijn

We state a result of De Bruijn which illustrates the transition of the situation of small $y$ to large $y$ which occurs around $y = \log x$.

**Theorem 8.1.** (De Bruijn, 1966) *Let $x \ge 3, 2 \le y \le x$ and write*

$$Z = \frac{\log x}{\log y}\log\left(1 + \frac{y}{\log x}\right) + \frac{y}{\log y}\log\left(1 + \frac{\log x}{y}\right).$$

*Then*

$$\log\psi(x,y) = Z\left\{1 + O\left(\frac{1}{\log y}\right) + O\left(\frac{1}{\log\log x}\right) + O\left(\frac{\log y}{\log x}\right)\right\}. \tag{8.6}$$

(Here we assume that $y \to \infty$ and $\log x/\log y \to \infty$.)

For example, if $y \approx \log x$, then

$$Z = 2\frac{\log x}{\log\log x}\log 2,$$

hence $\log\psi(x, \log x) \sim 2\dfrac{\log x}{\log\log x}\log 2$.

If $y \approx \log\log x$, then

$$Z = \frac{\log x}{\log\log\log x}\log\left(1 + \frac{\log\log x}{\log x}\right) + \frac{\log\log x}{\log\log\log x}\left(1 + \frac{\log x}{\log\log x}\right),$$

hence $\log\psi(x, \log\log x) \sim \dfrac{\log x}{\log\log\log x}$.

## 8.5 Homework for Chapter 8

1. Prove
   a) $\rho(u) > 0$ for $u > 0$.
   b) $\rho(u)$ is monotone decreasing for $u > 1$.
   c) $\rho(u) \leq \dfrac{1}{\Gamma(u+1)}$ for $u \geq 0$ where $\Gamma(u)$ is the Gamma-function defined by $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$ which satisfies $\Gamma(u+1) = u\Gamma(u)$ and, moreover, $\Gamma(n) = (n-1)!$ for all positive integers $n$.

2. An important function for the analysis of the $\rho$-function is the function $\xi(u)$ defined as the positive function such that

$$\frac{e^{\xi(u)}-1}{\xi(u)} = u \qquad \text{for } u > 1.$$

   a) Prove that if $f(x) = (e^x - 1)/x$, then $f(x) = 1 + \dfrac{x}{2!} + \dfrac{x^2}{3!} + \dfrac{x^3}{4!} + \dots$.
   b) Prove that $\xi(u)$ is well defined.
   c) Prove that $\xi = \log\xi + \log u + O\left(\dfrac{1}{\xi u}\right)$ as $\xi u \to \infty$.
   d) Prove that $\xi = \log u + \log\log u + O\left(\dfrac{\log\log u}{\log u}\right)$ as $u \to \infty$.

These exercises have to be done before Tuesday 9 January 2008.

## 8.6 Further exercises for Chapter 8

1. Prove that $u\rho(u) = \int_{u-1}^{u} \rho(t)dt$ for $u \geq 1$.

2. Prove (8.3).

3. Prove (8.4).

4. What does (8.6) say if

$$\text{a) } \frac{y}{\log x} \to 0, \qquad \text{b) } \frac{\log x}{y} \to 0, \qquad \text{c) } y = c\log x$$

   where $c$ is come positive constant?

5. Calculate the asymptotic behaviour of $\psi\left(x, e^{\sqrt{\log x}}\right)$ by using Theorem 8.1.

# Chapter 9

# Simultaneous Diophantine approximations.

A Diophantine approximation is an approximation of a real number by a rational number. The aim of simultaneous Diophantine approximation is to approximate distinct real numbers by rational numbers that have the same denominator. The first two theorems concern the former case, first homogeneous, then inhomogeneous.

**Theorem 9.1.** *Let $\alpha, Q \in \mathbb{R}, Q > 1$. Then there exist integers $p, q$ with $1 \leq q < Q$ such that*

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

**Proof.** We denote the fractional part of $x$ as $\{x\} = x - [x]$. Choose a positive integer $n$ such that $n < Q \leq n+1$ and consider the $n+2$ numbers $0, 1, \{\alpha\}, \{2\alpha\}, \ldots, \{n\alpha\}$. All of these numbers are in the interval $[0,1]$. By the box principle, one of the intervals $\left[0, \frac{1}{n+1}\right], \left(\frac{1}{n+1}, \frac{2}{n+1}\right], \ldots, \left(\frac{n-1}{n+1}, \frac{n}{n+1}\right], \left(\frac{n}{n+1}, 1\right]$ contains at least two of the numbers. But 0 and 1 cannot be numbers which are in the same interval. Thus there exist numbers $r_1, r_2, s_1, s_2$ with $0 \leq r_i \leq n$, $r_1 > r_2$ such that

$$|(r_1\alpha - s_1) - (r_2\alpha - s_2)| \leq \frac{1}{n+1} \leq \frac{1}{Q}.$$

Let $q = r_1 - r_2, p = s_1 - s_2$. Then $1 \leq q \leq n < Q$ and $|q\alpha - p| \leq \dfrac{1}{Q}$. □

**Corollary 9.1.** Let $\alpha \notin \mathbb{Q}$. Then the inequality $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$ has infinitely many solutions in rational number $p/q$.

**Theorem 9.2.** *If $\alpha \notin \mathbb{Q}$, $\theta$ and $N > 0$ are real numbers, then there exist integers $p$ and $q$ with $q > N$ such that $|q\alpha - p - \theta| \leq \frac{3}{q}$.*

**Proof.** By the previous corollary, there exist numbers $r, s \in \mathbb{Z}$ with $s > 2N$ such that $|\alpha - \frac{r}{s}| < \frac{1}{s^2}$. We may assume that $(r, s) = 1$. Choose $m \in \mathbb{Z}$ such that $|\theta - \frac{m}{s}| \leq \frac{1}{2s}$. We know that there exist integers $u, v$ such that $m = vr - us$. Moreover, we can choose $v$ such that $|v| \leq s/2$. We obtain

$$|q\alpha - p - \theta| \leq |q\alpha - p - \frac{m}{s}| + |\frac{m}{s} - \theta| \leq$$

$$\leq |q\alpha - p - \frac{vr}{s} + u| + \frac{1}{2s} \leq |(q-v)\alpha - (p-u)| + |v||\alpha - \frac{r}{s}| + \frac{1}{2s} \leq$$

$$\leq |(q-v)\alpha - (p-u)| + \frac{1}{2s} + \frac{1}{2s}.$$

Choose $q = s + v$ and $p = r + u$. Then we have

$$|q\alpha - p - \theta| \leq |s\alpha - r| + \frac{1}{s} < \frac{2}{s} \leq \frac{3}{s+v} = \frac{3}{q}.$$

Furthermore, $q \geq s - |v| \geq \dfrac{s}{2} > N$. □

**Corollary 9.2.** If $\alpha \notin \mathbb{Q}$ and $\theta \in \mathbb{R}$ then the inequality $|\alpha - \frac{p+\theta}{q}| \leq \frac{3}{q^2}$ has infinitely many solutions with $p \in \mathbb{Z}, q \in \mathbb{N}$.

The following theorem is a generalization of Theorem 9.1 to the simultaneous case. (By using the point $(1, 1, \dots, 1)$ it is possible even to require $q < Q^n$.)

**Theorem 9.3.** (Dirichlet) *Let $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$, $Q \in \mathbb{N}$, $Q > 1$. Then there exist integers $p_1, p_2, \dots, p_n, q$ with $1 \leq q \leq Q^n$ such that*

$$|q\alpha_j - p_j| \leq \frac{1}{Q} \qquad \textit{for } j = 1, \dots, n.$$

**Proof.** Consider the points of $\mathbb{R}^n$ of the form

$$(\{\alpha_1 x\}, \{\alpha_2 x\}, \dots, \{\alpha_n x\}) \qquad \text{with } x = 0, 1, \dots, Q^n.$$

So we get $Q^n{+}1$ points in the unit cube $T = \{(t_1, \dots, t_n) \mid 0 \leq t_j \leq 1 \quad \text{for} \quad j = 1, \dots, n\}$. Divide each edge of $T$ into $Q$ equal parts of length $1/Q$. So $T$ is split into $Q^n$ subcubes. Thus there is a subcube that contains at least two of the $Q^n{+}1$ points. Therefore, there are $x^{(1)}, x^{(2)}$ with $0 \leq x^{(2)} < x^{(1)} < Q^n$ such that $|\left\{\alpha_j x^{(1)}\right\} - \left\{\alpha_j x^{(2)}\right\}| \leq \frac{1}{Q}$ for $j = 1, \dots, n$. We define integers $y_1^{(1)}, y_2^{(1)}, \dots, y_n^{(1)}$ and $y_1^{(2)}, y_2^{(2)}, \dots, y_n^{(2)}$ such that

$$\left\{\alpha_j x^{(1)}\right\} = \alpha_j x^{(1)} - y_j^{(1)}, \left\{\alpha_j x^{(2)}\right\} = \alpha_j x^{(2)} - y_j^{(2)}.$$

It follows that

$$|(\alpha_j x^{(1)} - y_j^{(1)}) - (\alpha_j x^{(2)} - y_j^{(2)})| \leq \frac{1}{Q} \qquad \text{for } j = 1, \dots, n.$$

Choose $q = x^{(1)} - x^{(2)}$, $p_j = y_j^{(1)} - y_j^{(2)}$ for $j = 1, \ldots, n$. Then we have $1 \le q < Q^n$ and

$$|q\alpha_j - p_j| \le \frac{1}{Q} \qquad \text{for } j = 1, \ldots, n.$$

□

**Corollary 9.3.** Assume that at least one of the numbers $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ is irrational. Then there are infinitely many tuples of integers $p_1, p_2, \ldots, p_n, q$ with $(p_1, p_2, \ldots, p_n, q) = 1$ such that

$$|\alpha_j - \frac{p_j}{q}| < \frac{1}{q^{1+1/n}} \qquad \text{for } j = 1, \ldots, n.$$

**Proof.** By Theorem 9.3, for each $Q$ we can choose integers $p_1, \ldots, p_n, q$ with

$$|\alpha_j - \frac{p_j}{q}| \le \frac{1}{qQ} < \frac{1}{q^{1+1/n}} \qquad \text{for } j = 1, \ldots, n.$$

Choose $Q_0 = 2$. Then we find a solution $p_1^{(0)}, \ldots, p_n^{(0)}, q^{(0)}$. Dividing these numbers by their greatest common divisor, we get a coprime solution.
Assume $\alpha_j \notin \mathbb{Q}$. Then $|\alpha_j - \frac{p_j^{(0)}}{q^{(0)}}| \ne 0$. Choose $Q_1 \in \mathbb{N}$ such that $|q^{(0)}\alpha_j - p_j^{(0)}| > 1/Q_1$. For this $Q_1$ we find a new solution $p_1^{(1)}, \ldots, p_n^{(1)}, q^{(1)}$ by Theorem 9.3. Dividing the numbers by their greatest common divisor, we get a new coprime solution since $|q^{(1)}\alpha_j - p_j^{(1)}| \le \frac{1}{Q_1}$.
Now $|\alpha_j - \frac{p_j^{(1)}}{q^{(1)}}| \ne 0$. Choose $Q_2 \in \mathbb{N}$ such that $|q^{(1)}\alpha_j - p_j^{(1)}| > 1/Q_2$.
We proceed by induction. In this way we find infinitely many distinct coprime solutions $p_1, p_2, \ldots, p_n, q$. □

The generalization of Theorem 9.2 to simultaneous approximation is deeper. We say that $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent over $\mathbb{Z}$, if

$$m_1\alpha_1 + m_2\alpha_2 + \ldots + m_n\alpha_n = 0, \qquad m_1, \ldots, m_n \in \mathbb{Z}$$

implies that $m_1 = m_2 = \ldots = m_n = 0$.

**Theorem 9.4.** (Kronecker,1884) *Let $1, \alpha_1, \alpha_2, \ldots, \alpha_n$ be linearly independent over $\mathbb{Z}$. Assume that $N > 0$, $\varepsilon > 0$ and $\theta_1, \theta_2, \ldots, \theta_n$ are real numbers. Then there exist integers $p_1, p_2, \ldots, p_n, q$ with $q > N$ such that*

$$|q\alpha_j - p_j - \theta_j| < \varepsilon \qquad \textit{for } j = 1, \ldots, n.$$

**Remark.** The condition that $1, \alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent should not be omitted. For $n = 1$ it means that $\alpha$ is irrational. For $\alpha_1 = 1$ the inequality $|q\alpha_1 - p_1 - \frac{1}{2}| < \frac{1}{4}$ is not soluble. Likewise the system $|q\sqrt{2} - p_1 - \frac{1}{4}| < \frac{1}{20}$, $|q\sqrt{8} - p_2 - \frac{1}{4}| < \frac{1}{20}$ is not soluble.

Our proof of Theorem 9.4 is derived from the following theorem which is proved by complex analysis. A proof using only real analysis can be found in Hardy and Wright, The Theory of Numbers, Ch. 23.

**Theorem 9.5.** *Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be linearly independent over $\mathbb{Z}$. Assume that $N > 0$, $\varepsilon > 0$ and $\theta_1, \theta_2, \ldots, \theta_n$ are real numbers. Then there are numbers $p_1, p_2, \ldots, p_n \in \mathbb{Z}$ and $t \in \mathbb{R}$, $t > N$ such that*

$$|t\alpha_j - p_j - \theta_j| < \varepsilon \qquad \textit{for } j = 1, \ldots, n.$$

For the proof we need the following lemmas:

**Lemma 9.1.** *Assume that $c_1, \ldots, c_n$ are distinct real numbers and $b_1, \ldots, b_n$ are arbitrary complex numbers. Put $f(t) = \sum_{\nu=1}^{r} b_\nu e^{c_\nu it}$. Then*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(t) e^{-c_\mu it} dt = b_\mu \qquad \textit{for } \mu = 1, \ldots, r.$$

**Proof.** For $c = 0$ we obtain $\lim_{T\to\infty} \dfrac{1}{T} \int_0^T e^{cit} dt = \lim_{T\to\infty} \dfrac{T}{T} = 1$.
For $c \neq 0$ we get

$$\frac{1}{T} \int_0^T e^{cit} dt = \frac{1}{ciT} \left[e^{cit}\right]_{t=0}^{t=T} = \frac{e^{ciT} - 1}{ciT} \to 0 \qquad \text{if } T \to \infty.$$

Thus

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T e^{cit} dt = \begin{cases} 1 & \text{if} \quad c = 0, \\ 0 & \text{if} \quad c \neq 0. \end{cases}$$

It follows that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(t) e^{-c_\mu it} dt = \lim_{T\to\infty} \sum_{\nu=1}^{r} b_\nu \frac{1}{T} \int_0^T e^{(c_\nu - c_\mu)it} dt = b_\mu.$$

□

**Lemma 9.2.** *Let $\psi(x_1, \ldots, x_n) = 1 + x_1 + \ldots + x_n$. Let $k \in \mathbb{N}$. Then*

$$(\psi(x_1, \ldots, x_n))^k = \sum_{k_1 + \ldots + k_n \leq k, k_j \geq 0, k_j \in \mathbb{Z}} a_{k_1 \ldots k_n} x_1^{k_1} \cdots x_n^{k_n}$$

*with $a_{k_1 \ldots k_n}$ positive integers such that*

$$\sum_{k_1, \ldots, k_n} a_{k_1 \ldots k_n} = (n+1)^k.$$

*The number of terms in the summation is at most $(k+1)^n$.*

**Proof.** The first assertion follows directly, the second follows by taking $x_1 = \ldots = x_n = 1$. For each $k_j$ there are $k+1$ choices, so the total number of terms does not exceed $(k+1)^n$.

**Proof of Theorem 9.5.** (H. Bohr, 1914) Let

$$F(t) = 1 + \sum_{j=1}^{n} e^{2\pi i(\alpha_j t - \theta_j)}, \qquad t \in \mathbb{R}.$$

Then $|F(t)| \leq n+1$ for all $t \in \mathbb{R}$. By Lemma 9.2 we have

$$(F(t))^k = \sum_{k_1+\ldots+k_n \leq k, k_j \geq 0, k_j \in \mathbb{Z}} a_{k_1\ldots k_n} e(\sum_{j=1}^{n} k_j(\alpha_j t - \theta_j)),$$

with at most $(k+1)^n$ terms and the sum of the coefficients is $(n+1)^k$. Define

$$b_{k_1\ldots k_n} = a_{k_1\ldots k_n} e(-\sum_{j=1}^{n} k_j \theta_j).$$

Then we have $|b_{k_1\ldots k_n}| = a_{k_1\ldots k_n}$ and

$$(F(t))^k = \sum_{k_1+\ldots+k_n \leq k, k_j \geq 0, k_j \in \mathbb{Z}} b_{k_1\ldots k_n} e(\sum_{j=1}^{n} k_j \alpha_j t).$$

Since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent, all terms $\sum_{j=1}^{n} k_j \alpha_j$ are distinct. By Lemma 9.1 we get

$$b_{k_1\ldots k_n} = \lim_{T\to\infty} \frac{1}{T} \int_0^T (F(t))^k e(-\sum_{j=1}^{n} k_j \alpha_j t) dt.$$

Assume that there exist real numbers $\lambda < n+1$ and $t_0 > 0$ such that

$$|F(t)| \leq \lambda \qquad \text{for } t \geq t_0.$$

Then

$$|b_{k_1\ldots k_n}| \leq \lim_{T\to\infty} \frac{1}{T} \int_0^T |(F(t))^k| dt$$

$$\leq \lim_{T\to\infty} \frac{1}{T} \int_0^{t_0} (n+1)^k dt + \lim_{T\to\infty} \frac{1}{T} \int_{t_0}^{T} \lambda^k dt = \lambda^k.$$

Thus

$$(n+1)^k = \sum |b_{k_1\ldots k_n}| \leq \lambda^k (k+1)^n.$$

Therefore

$$\left(\frac{n+1}{\lambda}\right)^k \leq (k+1)^n.$$

The left-hand side grows exponentially in $k$ to $\infty$, the right-hand side polynomially. Thus, if $k$ is sufficiently large then we get a contradiction. Therefore for each $t_0 > 0$ and $\lambda < n+1$ there exists a $t > t_0$ with $|F(t)| > \lambda$.

Let $N > 0$ and $0 < \varepsilon < \frac{1}{2}$. For $t_0 = N$ and $\lambda = n + 1 - \frac{\varepsilon^2}{2}$ we can find a $t > N$ such that for $j = 1, \ldots, n$

$$n + 1 - \frac{\varepsilon^2}{2} < |F(t)| \leq |1 + e((\alpha_j t - \theta_j))| + n - 1.$$

Thus

$$|e^{\pi i(\alpha_j t - \theta_j)} + e^{-\pi i(\alpha_j t - \theta_j)}| > 2 - \frac{\varepsilon^2}{2}$$

which implies

$$\cos\left(\pi(\alpha_j t - \theta_j)\right) > 1 - \frac{\varepsilon^2}{4} \qquad \text{for } j = 1, \ldots, n.$$

Observe that $|\cos \pi x| > 1 - \frac{\delta^2}{4} \Longrightarrow ||x|| < \delta$, where $||x||$ denotes the distance of $x$ from the nearest integer number. (Exercise.) Thus

$$||\alpha_j t - \theta_j|| < \varepsilon \qquad \text{for } j = 1, \ldots, n.$$

Let $p_j$ be the integer number nearest to $\alpha_j t - \theta_j$. Then it follows that

$$|\alpha_j t - p_j - \theta_j| < \varepsilon \qquad \text{for } j = 1, \ldots, n.$$

□

**Proof of Theorem 9.4.** Put $A = \max\left(1, |\alpha_1|, |\alpha_2|, \ldots, |\alpha_n|\right)$. We use Theorem 9.5 for $1, \alpha_1, \ldots, \alpha_n$ with $0, \theta_1, \ldots, \theta_n$ as the corresponding $\theta's$ and with $\varepsilon/2A$ in the place of $\varepsilon$ and $N + \varepsilon$ instead of $N$. Hence there exist integers $q = p_0, p_1, \ldots, p_n \in \mathbb{Z}$ and $t \in \mathbb{R}, t > N + \varepsilon$ such that

$$|t\alpha_j - p_j - \theta_j| < \frac{\varepsilon}{2A} \qquad \text{for } j = 1, \ldots, n$$

and

$$|t - q| < \frac{\varepsilon}{2A}.$$

It follows that $q > t - \varepsilon > N$ and

$$|q\alpha_j - p_j - \theta_j| \leq |(q - t)\alpha_j| + |t\alpha_j - p_j - \theta_j| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2A} \leq \varepsilon \qquad \text{for } j = 1, \ldots, n.$$

□

**Remark.** The various proofs of theorems of Dirichlet and Kronecker do not lead to an algorithm which computes approximating numbers. In 1983 A. K. Lenstra, H. W. Lenstra and L. Lovász developed a basis reduction algorithm that can find approximations as in Theorem 9.3 in an efficient way, but the bound is somewhat worse than the one in Dirichlet's theorem.

## 9.1 Homework for Chapter 9

1. Which of the following systems are solvable in $x, y, z \in \mathbb{Z}$ with $(x, y, z) \neq (0, 0, 0)$?
   a) $|x \log 2 - y| \leq \frac{1}{100}$, $|x \log 3 - z| \leq \frac{1}{100}$,
   b) $|x \log 2 - y| \leq \frac{1}{100}$, $|x \log 4 - z| \leq \frac{1}{100}$,
   c) $|x \log 2 - y| \leq \frac{1}{100}$, $|x \log 3 - z - 1/2| \leq \frac{1}{100}$,
   d) $|x \log 2 - y| \leq \frac{1}{100}$, $|x \log 4 - z - 1/2| \leq \frac{1}{100}$.

2. a) Prove Corollary 9.2.
   b) Prove that the equation $x^2 - 2y^2 = \pm 1$ has infinitely many solutions in positive integers $x, y$.

The homework of Chapters 8 and 9 can be put in my mailbox until 9 January, 2008.

## 9.2 Further exercises for Chapter 9

1. Check whether the following inequalities have infinitely many solutions:
   a) $|q\sqrt{2} - p| < \frac{1}{2q}$ in $p, q \in \mathbb{N}$,
   b) $|q\sqrt{2} - p - \sqrt{3}| \leq \frac{3}{q}$ in $p, q \in \mathbb{N}$,
   c) the system $|q\sqrt{2} - p_1| \leq \frac{1}{\sqrt{q}}$, $|q\sqrt{3} - p_2| \leq \frac{1}{\sqrt{q}}$ in $p_1, p_2, q \in \mathbb{N}$.

2. Prove:
   a) $\alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbb{Z} \iff \alpha_1, \ldots, \alpha_n$ are linearly independent over $\mathbb{Q}$.
   b) $1, \alpha$ are linearly independent over $\mathbb{Z} \iff \alpha \notin \mathbb{Q}$.
   c) $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ are linearly independent over $\mathbb{Z}$.

3. Prove: $|\cos \pi x| \geq 1 - \frac{1}{2}\delta^2 \Longrightarrow ||x|| < \delta$.

4. A star has $n$ planets which run with constant angular velocities around the star and all lie in the same plane. Prove that, seen from the star, the planets are almost in the same direction (within an angle $\varepsilon > 0$) infinitely often if
   a) the planets have once be seen in the same direction,
   b) their angular velocities are linearly independent over $\mathbb{Z}$.