512.81 - R

© *Copyright 1964, by Blaisdell Publishing Company, a Division of Ginn and Company.*

All rights reserved

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 64-10050

Preface

This little book was not planned as a textbook of elementary number theory. Nevertheless it may be regarded as a sort of introduction to number theory, since it does not presuppose any previous knowledge in this field. Using a background of analysis and algebra, the reader is led to the fundamental theorems of number theory: The uniqueness of prime number factorization and the reciprocity law of quadratic residues. In both cases the text does not pursue the direct and shortest path. It seemed to me more attractive to look around in the world of numbers and then to unfold the underlying structure. Thus the uniqueness of prime number factorization comes at the end of a discussion of common fractions, as they appear in the arrangements of the Farey sequences, and the quadratic reciprocity law is attached to investigations in cyclotomy, which start with the Gaussian construction of the regular heptadecagon. Cyclotomy is treated in some detail, because of its own significance and also as framework for the elegant theorems on Gaussian sums. Then there are some discussions of asymptotic laws, as a foretaste of analytic number theory. The presentation climaxes with Dirichlet's theorem about primes in an arithmetic progression and V. Brun's theorem on twin primes.

The book derives its character from its origin: It is the result of a series of lectures given at Haverford College in 1959–1960 under the auspices of the William Pyle Philips Fund. The book does not render these lectures quite faithfully. Besides having rearranged some topics to make them better fit together, I have omitted two lectures on the elements of the theory of partitions, which did not seem to agree too well with the general tenor of the others, and finally I have added the last chapter, which I had planned for my Haverford lectures, but because of lack of time could not deliver there. I did lecture on this last chapter, though at another place, namely, Dartmouth College, in October, 1960.

I wish to thank my old friend Cletus Oakley for encouraging me to collect the lectures for this book. My thanks go also to Louis Solomon for writing out notes of my lectures and to James O. Brooks for his help with the manuscript.

HANS RADEMACHER

Philadelphia, Pa.
February, 1964

Contents

Introduction

Let a and c be positive integers, or "natural numbers" as we call them. We say that a *divides* c and write $a \mid c$ if there exists a natural number b such that $c = ab$. Every natural number a is divisible by 1 and a, divisors which sometimes are singled out as "improper" divisors. Thus, 6 has the proper divisors 2 and 3 and the improper ones 1 and 6. If the number $n > 1$ has no divisors other than the improper ones we say that n is a *prime number*. Otherwise n is *composite*. Evidently every composite number can be factored into primes. The sequence of primes begins with 2, 3, 5, 7, 11, 13, 17, 19, $\cdots$ and has a rather irregular appearance. Many number theoretical problems are concerned with this sequence. First, as we move along the sequence of natural numbers, the prime numbers become more and more scarce. This is quite plausible because a large number has a greater chance than a small one of being composite, as it surpasses more numbers which might be eligible for its divisors. It is even conceivable that all sufficiently large numbers might be composite. This, however, is not so. Euclid (around 300 B.C.) has proved the

THEOREM: *Outside any given finite set of prime numbers there is another one.*

In other words: there is no largest prime number. Here is Euclid's proof:

Let us write down any finite set of prime numbers, e.g., those from 2 to a certain p, say $2, 3, 5, \cdots, p$. We form the product of these primes and then consider the integer

$$N = (2 \cdot 3 \cdot 5 \cdots p) + 1.$$

We write this number as a product of primes $N = q_1 \cdots q_s$, where $s = 1$ would account for the possibility that N might itself already be a prime number. Now no $q_j, j = 1, \cdots, s$ is equal to any of the primes $2, 3, 5, \cdots, p$, since none of these divides N, whereas all the q_j do divide N. Thus there exists a prime distinct from those in the set $2, 3, 5, \cdots, p$, as the theorem states.

We can get a bit more out of the same argument. Let us look at odd primes (i.e., those different from 2). When divided by 4 they leave the odd remainders 1 or 3. We write these cases as $p \equiv 1 \pmod 4$ or $p \equiv 3 \pmod 4$.† For example:

$$5, 13, 17, 29, \cdots \equiv 1 \pmod 4$$

$$3, 7, 11, 19, 23, \cdots \equiv 3 \pmod 4.$$

† Read "modulo 4," that is "by the measure 4."

Now we have just seen that both classes together contain infinitely many primes, but this does not preclude the possibility that one of these classes might contain only finitely many. We can use Euclid's argument to show that there exist infinitely many primes $\equiv 3 \pmod 4$. For let us take any set of such primes $3, 7, 11, \cdots, p$. Consider this time $M = (4 \cdot 3 \cdot 7 \cdot 11 \cdots p) - 1$. Then we see that $M \equiv 3 \pmod 4$. We write now $M = q_1 \cdot q_2 \cdots q_t$ as a product of primes. At least one of these primes q_j is $\equiv 3 \pmod 4$, since the product of two numbers of the form $4k + 1$ is again of this form:

$$(4m + 1)(4n + 1) = 4(4mn + m + n) + 1.$$

But no q_j is among the set $3, 7, 11, \cdots, p$, and so there exists another prime $\equiv 3 \pmod 4$.

It is also true that the class of primes $\equiv 1 \pmod 4$ contains infinitely many elements. But we postpone the proof since it is more involved.†

Exercise: Prove, by the same argument, that there exist infinitely many primes $\equiv 5 \pmod 6$.

The general problem of this type was solved by Lejeune Dirichlet about 125 years ago by quite different methods. Let us look for primes in an arithmetic progression, $a, a + m, a + 2m, \cdots$, or, in brief, $p \equiv a \pmod m$. Here a and m are given natural numbers. Clearly we can find such primes (besides possibly a itself) only if a and m have no common divisor except 1. We let (a, m) denote the greatest common divisor of the two numbers a and m. If $(a, m) = 1$ we say that a and m are *relatively prime* or *coprime*. Following Dirichlet we shall show later that any arithmetic progression with $(a, m) = 1$ contains infinitely many primes (Theorem 57).

Properties of prime numbers will be the main topic of these lectures. Now, the theory of prime numbers leads to problems which have withstood the efforts of the greatest mathematicians through the centuries. However, there are problems wholly or partly accessible by fairly simple means. I mention a few which will occupy us.

(1) If p is a prime then

$$(p - 1)! \equiv -1 \pmod p \text{ or } (p - 1)! + 1 \equiv 0 \pmod p:$$
$$2! + 1 = 3 \equiv 0 \pmod 3,$$
$$4! + 1 = 25 \equiv 0 \pmod 5, \text{ and}$$
$$6! + 1 = 721 \equiv 0 \pmod 7$$

are some examples. The theorem was first published by Waring, but goes under the name of Wilson's theorem. It will appear in our discussion as a simple corollary.

† See Chapter 5, Theorem 19.

(2) Let p be a prime and let a be a positive integer relatively prime to p. Then $a^{p-1} \equiv 1 \pmod p$. This result was proved by Pierre de Fermat (1601–1665), a high judge in Toulouse and one of the greatest number theorists in mathematical history.

(3) If $p \equiv 1 \pmod 4$ then p is the sum of two square numbers,

$$5 = 1^2 + 2^2$$
$$13 = 2^2 + 3^2$$
$$17 = 1^2 + 4^2$$
$$29 = 2^2 + 5^2,$$

and essentially in only one way. Of course, even the longest list of such examples is no proof of the theorem, since the proof must deduce this property out of the general nature of such primes and thus be valid for all (infinitely many) primes. We owe this theorem also to Fermat. Primes $p \equiv 3 \pmod 4$ may not be so decomposed. This is easy to prove. For squares of even numbers are $\equiv 0 \pmod 4$ and squares of odd numbers are $\equiv 1 \pmod 4$. The sum of two squares can therefore only be congruent to 0, 1, or 2 modulo 4, but never $\equiv 3 \pmod 4$.

(4) Let $\pi(x)$ be the number of primes $\leq x$. Carl Friedrich Gauss (1777–1855) conjectured at the age of 15, by inspecting a table of primes, that $\pi(x)$ is approximately $\dfrac{x}{\log x}$ in the sense that the ratio $\pi(x) : \dfrac{x}{\log x}$ converges to 1 if x increases indefinitely. This conjecture was made in 1792, but the theorem was first proved in 1896 independently by Jacques Hadamard and Charles de la Vallée Poussin. We shall prove a much weakened form of this theorem which goes back to Chebyshev.

(5) The prime numbers of the form $p = 2^{2^k} + 1$, e.g., $2^1 + 1 = 3$, $2^2 + 1 = 5$, $2^4 + 1 = 17$, $2^8 + 1 = 257, \cdots$, called Fermat primes, play a fundamental role in the construction of regular polygons, after Gauss. Fermat conjectured that all numbers of the form $2^{2^k} + 1$ are primes. However, this is not so: Euler has shown that $2^{32} + 1$ is composite and contains the prime factor 641. It is not known whether there are infinitely many Fermat primes.

(6) The sequence of primes appears to be very irregular, with gaps between consecutive primes varying in size. The gap can be arbitrarily large: indeed all of the $(N - 1)$ consecutive numbers

$$N! + 2, N! + 3, N! + 4, \cdots N! + N$$

are composite, so that the difference between the smallest prime above this set and the greatest prime below this set is at least N.

Leaving aside the pair 2, 3, the difference between two consecutive primes must at least be 2, e.g.,

$$3, 5;\ 5, 7;\ 11, 13;\ 17, 19;\ 29, 31;\ \cdots.$$

Such primes are called "twin primes." It is not known whether there are infinitely many pairs of twin primes, although there are some indications that this is so. The Norwegian mathematician V. Brun modified this problem to make it more accessible. He considered not primes alone but numbers containing only a small number of prime factors. Then, as A. Selberg has shown, among the numbers having at most three prime divisors exist infinitely many pairs of difference 2.

One can generalize this problem in many ways. For example one notices the prime quadruplets

$$\begin{aligned} &11, \quad 13, \quad 17, \quad 19 \\ &101, \quad 103, \quad 107, \quad 109 \\ &191, \quad 193, \quad 197, \quad 199 \\ &821, \quad 823, \quad 827, \quad 829, \end{aligned}$$

each quadruplet lying within a decade. The largest such quadruplet registered in D. N. Lehmer's *Table of Primes* is

$$9933611, \quad 9933613, \quad 9933617, \quad 9933619.$$

The last lectures of this book will discuss twin primes. We shall, however, treat quite the opposite problem: we shall show that there are, in a certain sense, not too many twin primes.

1

Factorization and Farey Fractions

Decomposition of a number; factorization into prime factors. At first glance it seems obvious that the decomposition of a natural number into prime factors is unique up to the order of the factors. We break down a number into factors, each of which, unless it is a prime number, we decompose again until we have reached prime factors throughout. This must take place at some point since the factors become smaller and smaller. For example

$$60 = 4 \cdot 15 = 2 \cdot 2 \cdot 3 \cdot 5$$

or

$$60 = 6 \cdot 10 = 2 \cdot 3 \cdot 2 \cdot 5 .$$

This example shows two different procedures for breaking down the number 60, but both end with the same prime divisors. If not spoiled by too much knowledge you would take this as obvious. But Euclid developed a long and complicated proof for the uniqueness of prime factorization. Why? Why is unique prime factorization not a trivial statement? To answer this let us consider the numbers $n \equiv 1 \pmod 3$:

$$1, 4, 7, 10, 13, 16, 19, 22, 25, \cdots, 100, \cdots .$$

The product of two such numbers is again of the same form, for

$$(3k + 1)(3l + 1) = 3(3kl + k + l) + 1 .$$

Let us call a number in our set "primitive" if it cannot be written as a product of other numbers in the set, which being smaller must precede it in the set. Thus the numbers 4, 7, 10, 13 are primitive, whereas 16 is composite, being $4 \cdot 4$. It is clear that every number which is not primitive itself can be decomposed into a product of primitive numbers. But here is a surprise:

$$100 = 10 \cdot 10 \text{ and } 100 = 4 \cdot 25$$

are two different decompositions into primitive factors. Therefore, the possibility of decomposing a number into factors cannot logically entail the uniqueness of the final decomposition.

Here is another example. We consider the set of numbers of the form $a + b\sqrt{-5}$, where a and b are integers. This set is evidently closed under addition and subtraction, and under multiplication as well:

$$(a + b\sqrt{-5})(c + d\sqrt{-5}) = (ac - 5bd) + (ad + bc)\sqrt{-5}\,.$$

Such closure is necessary, of course, if we are to reasonably call a number $a + b\sqrt{-5}$ an "integer." Again, numbers in this realm are either "primitive," i.e., are not products of others, or are products of primitive numbers. Then $3,\ 7,\ 1 + 2\sqrt{-5},\ 1 - 2\sqrt{-5},\ 4 + \sqrt{-5},\ 4 - \sqrt{-5}$, are all primitive (as can be tested by trial), and the number 21 has 3 distinct factorizations:

$$21 = 3 \cdot 7 = (1 + 2\sqrt{-5})(1 - 2\sqrt{-5}) = (4 + \sqrt{-5})(4 - \sqrt{-5})\,.$$

Common fractions. We are postponing the proof of the uniqueness of prime factorization for the moment and shall obtain it as a result of a study of common fractions. We shall then also give a direct proof of it.

A common fraction is denoted by the symbol $\frac{x}{y}$ where x and y are integers and where the denominator $y > 0$. We assume the rules of calculation with integers as given and reduce all statements about fractions to statements about integers. For the moment we are interested only in the *ordering* of the fractions. First we define *equality*.

DEFINITION: $\frac{a}{b} = \frac{c}{d}$ *is equivalent to* $ad = bc$.

The relation of equality is an "equivalence relation" and as such has the required properties:

Reflexivity: $\frac{a}{b} = \frac{a}{b}$ since $ab = ba$.

Symmetry: If $\frac{a}{b} = \frac{c}{d}$ then $\frac{c}{d} = \frac{a}{b}$. Indeed $ad = bc$ implies $cb = da$.

Transitivity: If $\frac{a}{b} = \frac{c}{d}$ and $\frac{c}{d} = \frac{e}{f}$, then $\frac{a}{b} = \frac{e}{f}$ because $ad = bc$, $cf = de$,

and thus $adf = bcf = bde$. Therefore $af = be$ since $d \neq 0$.

Secondly we define *inequality*.

DEFINITION: $\frac{a}{b} < \frac{c}{d}$ *is equivalent to* $ad < bc$.

It is easily seen that this relation, as it should be, is

Irreflexive: It is *not* true that $\frac{a}{b} < \frac{a}{b}$.

Asymmetric: If $\frac{a}{b} < \frac{c}{d}$, then it is *not* true that $\frac{c}{d} < \frac{a}{b}$.

Transitive: If $\frac{a}{b} < \frac{c}{d}$ and $\frac{c}{d} < \frac{e}{f}$, then $\frac{a}{b} < \frac{e}{f}$.

These properties follow immediately from the definition of inequality and from the fact that the denominators are always taken as positive.

We also stipulate that $\frac{a}{b} < \frac{c}{d}$ means the same as $\frac{c}{d} > \frac{a}{b}$. We then have among common fractions the *trichotomy*: For two fractions $\frac{a}{b}, \frac{c}{d}$ one and only one of the statements

$$\frac{a}{b} < \frac{c}{d}, \qquad \frac{a}{b} = \frac{c}{d}, \qquad \frac{a}{b} > \frac{c}{d}$$

is true.

Indeed, among integers we have either $ad < bc$ or $ad = bc$ or $ad > bc$.

These definitions, in particular the transitivity of equality and inequality, permit an *ordering* of the common fractions according to $\leqq$.

Now we usually take it for granted that, among the infinitely many fractions which are *equal*, for example

$$\frac{15}{20} = \frac{12}{16} = \frac{3}{4} = \frac{9}{12} = \cdots,$$

there is exactly one which is reduced, i.e., in which numerator and denominator have no common divisor except 1. This is, however, by no means trivial. If we think of our example of the numbers $\equiv 1 \pmod 3$, we have all the properties of ordering which we just discussed present there also. However, we have

$$\frac{4}{10} = \frac{10}{25}$$

which in that domain are two reduced fractions, since there the numbers 4, 10, 25 are primitive numbers.

If, therefore, we start a theory of common fractions *ab ovo*, we first have to admit the possibility that two reduced fractions that are *not identical* might nevertheless be *equal*.

Farey's scheme of common fractions. In order to gain insight into the properties of common fractions, we investigate some observations which the mineralogist Farey made and published without proof in 1816. Cauchy immediately furnished the proofs.

Farey wrote down the ordered sequence of all nonnegative reduced fractions between 0 and 1 whose denominators are limited by a number N,

called the *order* of the *Farey sequence*. The Farey sequences of orders 1, 2, 3, 4, 5, respectively, are

$$\frac{0}{1}, \frac{1}{1}$$

$$\frac{0}{1}, \frac{1}{2}, \frac{1}{1}$$

$$\frac{0}{1}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{1}{1}$$

$$\frac{0}{1}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{1}{1}$$

$$\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}$$

We find in each of these five sequences fractions ordered only by the relation $<$. Since, in general, we have not yet proved that two *equal* reduced fractions are *identical*, we have to envisage the possibility of equalities in our series of order N and would then agree to order such fractions according to the sizes of their numerators. As we see, this possibility does not occur in the orders 1 to 5. As a matter of fact we shall show *that it will never appear*. In other words, two consecutive fractions $\frac{h}{k}, \frac{l}{m}$ in a Farey sequence of order N will always obey the relation

$$\frac{h}{k} < \frac{l}{m} \tag{1.1}$$

or, according to definition, $hm < kl$. More precisely, Farey found the following about the difference of these numbers.

THEOREM 1: *For two consecutive fractions* $\frac{h}{k}, \frac{l}{m}$ *of the Farey sequence of order* N *we have*

$$hm = kl - 1 ,$$

or in another formulation

$$\begin{vmatrix} h & l \\ k & m \end{vmatrix} = -1 . \tag{1.2}$$

Proof: We prove this theorem by induction. It is true, as inspection shows, for $N = 1, 2, 3, 4, 5$. Assume now that it is true for order N. Then let $\frac{a}{b}$ be a *reduced* fraction not in the Farey sequence of order N, i.e., for which $b > N$. This $\frac{a}{b}$ will have its place somewhere in the interval [0, 1]

between some two consecutive fractions $\frac{h}{k}$ and $\frac{l}{m}$ of the Farey sequence of order N:

$$\frac{h}{k} \leqq \frac{a}{b} \leqq \frac{l}{m} . \tag{1.3}$$

We assume hypothesis (1.2) concerning $\frac{h}{k}, \frac{l}{m}$. The "=" signs cannot be discarded *a priori*. We can only say, in view of (1.1), that not both of them can be valid. Now we set

$$\begin{aligned} \lambda &= \begin{vmatrix} a & h \\ b & k \end{vmatrix} = ak - bh \geqq 0 \\ \mu &= \begin{vmatrix} l & a \\ m & b \end{vmatrix} = -am + bl \geqq 0 \end{aligned} \tag{1.4}$$

where λ, μ are nonnegative integers, by the definition of equality and inequality of fractions. We can solve these equations for a and b and obtain through elementary rules

$$\begin{aligned} a &= \frac{\begin{vmatrix} \lambda & -h \\ \mu & l \end{vmatrix}}{\begin{vmatrix} k & -h \\ -m & l \end{vmatrix}} = \mu h + \lambda l \\ b &= \frac{\begin{vmatrix} k & \lambda \\ -m & \mu \end{vmatrix}}{\begin{vmatrix} k & -h \\ -m & l \end{vmatrix}} = \mu k + \lambda m , \end{aligned} \tag{1.5}$$

where we have used (1.2) for the determination of the denominator. *All proper reduced fractions* $\frac{a}{b}$ in (1.3) can be expressed through (1.5) by means of suitable nonnegative integers λ, μ as our argument shows.

Conversely, for all $\lambda, \mu \geqq 0$, $\lambda + \mu > 0$, we have

$$\frac{h}{k} \leqq \frac{\mu h + \lambda l}{\mu k + \lambda m} \leqq \frac{l}{m} ,$$

since $h(\mu k + \lambda m) \leqq k(\mu h + \lambda l)$ and $(\mu h + \lambda l)m \leqq (\mu k + \lambda m)l$ because of (1.2).

Now $\lambda = 0$ or $\mu = 0$, admitted through (1.4) and (1.5), are both impossible. If $\lambda = 0$, we would have

$$\frac{a}{b} = \frac{\mu h}{\mu k}$$

which can be a *reduced* fraction only for $\mu = 1$. We would then have

$$a = h, \qquad b = k$$

which contradicts the fact $b > N$. Similarly $\mu \neq 0$. *All* the proper fractions $\frac{a}{b}$ in (1.3) therefore appear in the form

$$\frac{a}{b} = \frac{\mu h + \lambda l}{\mu k + \lambda m}, \lambda \geqq 1, \mu \geqq 1 .$$

Now, the *smallest* value that the denominator b can attain is $N + 1$, which would occur if $\frac{a}{b}$ belonged to the Farey sequence of order $N + 1$, but not of order N. Then λ, μ must be as small as possible, i.e., $\lambda = 1$, $\mu = 1$, so that we have

$$a = h + l \qquad b = k + m = N + 1 .$$

This new fraction $\frac{a}{b} = \frac{h + l}{k + m}$ satisfies Farey's theorem with respect to its neighbors: Indeed we have

$$\begin{vmatrix} h & a \\ k & b \end{vmatrix} = \begin{vmatrix} h & h + l \\ k & k + m \end{vmatrix} = \begin{vmatrix} h & l \\ k & m \end{vmatrix} = -1$$

and

$$\begin{vmatrix} a & l \\ b & m \end{vmatrix} = \begin{vmatrix} h + l & l \\ k + m & m \end{vmatrix} = \begin{vmatrix} h & l \\ k & m \end{vmatrix} = -1 ,$$

where we have used (1.2).

Thus we have shown that Farey's theorem also holds for order $N + 1$, and therefore, through mathematical induction, for all N.

Let us call $\frac{h + l}{k + m}$ the "mediant" between $\frac{h}{k}$ and $\frac{l}{m}$. Then we have proved the following theorem at the same time.

THEOREM 2: *The fractions which belong to the Farey sequence of order $N + 1$ but not of order N are mediants of the Farey sequence of order N.*

Starting therefore from the Farey sequence $\frac{0}{1}, \frac{1}{1}$, the following sequences can be built up simply by inserting successively the mediants with the appropriate denominators. Since a mediant of the Farey sequence of order N must belong as a fraction to some Farey sequence of higher order, we have the following interesting theorem.

THEOREM 3: *The denominators of two adjacent fractions of a Farey sequence of order N add up to at least $N + 1$.*

We emphasize again that our discussion of the Farey sequences, in which, as we have seen, no equality sign can occur, has shown that two reduced fractions which are not identical cannot be equal (in the meaning of the definition of equality between fractions).†

The further theory of common fractions involving addition, subtraction, multiplication, and division is not of particular interest from our point of view, and we take it for granted as explained in elementary arithmetic.

† It is easily seen that this fact has "Euclid's lemma" (see Chapter 2) as a direct consequence.

2

Euclid's Lemma, Uniqueness of Prime Factorization

A linear Diophantine equation and Euclid's lemma. We now draw some important consequences from our theory of Farey sequences. First we discuss a linear Diophantine equation that is basic for many arguments in number theory.

THEOREM 4: *Let a and b be coprime: $(a, b) = 1$. Then the Diophantine equation*

$$ax + by = 1 \tag{2.1}$$

is solvable.

Remark: A Diophantine equation (called after Diophantus of Alexandria, third century) is an equation that has to be satisfied by integers.

Proof: Assume without loss of generality $0 < a < b$. Then, since $(a, b) = 1$, $\frac{a}{b}$ is a proper reduced fraction and consequently $\frac{a}{b}$ appears in some Farey sequence (e.g., that of order b). Let us now take an adjacent fraction

$$\frac{h}{k} < \frac{a}{b}.$$

Then by Farey's theorem, we have

$$\begin{vmatrix} h & a \\ k & b \end{vmatrix} = -1$$

and thus

$$ak - bh = 1 .$$

Therefore $x = k$, $y = -h$ is a solution of (2.1).

COROLLARY: *Suppose $(a, b) = 1$. Then we can also solve the Diophantine equation*

$$ax + by = c . \tag{2.2}$$

For let $ax_0 + by_0 = 1$. Then $a(x_0c) + b(y_0c) = c$, and $x = x_0c$, $y = y_0c$ is a solution of (2.2).

Euclid's lemma. The foregoing theorem now leads immediately to an important result which was proved by Euclid in a different manner.

THEOREM 5: (Euclid's lemma). *If a and b are coprime and $a \mid bc$ then $a \mid c$.*

Proof: Choose integers x and y so that $ax + by = 1$, which is possible according to Theorem 4. Then we have $acx + bcy = c$. Now since bc is a multiple of a, it can be expressed as $bc = ad$. Inserting this into the foregoing equation we have $a(cx + dy) = c$, which implies $a \mid c$. This theorem enables us to make Theorem 4 more specific.

THEOREM 4a: *Let $(a, b) = 1$. Then all solutions of (2.1) are contained in the formula*

$$x = x_0 - bt , \qquad y = y_0 + at , \tag{2.3}$$

where x_0, y_0 is any special solution of (2.1) and t is any integer.

Proof: Indeed (2.1) is fulfilled if we insert (2.3) in it, since we assume $ax_0 + by_0 = 1$. Now, conversely, if x_0, y_0 and x, y are two solutions of (2.1), then by subtraction it follows that

$$a(x - x_0) = -b(y - y_0) . \tag{2.4}$$

But $(a, b) = 1$ implies by Euclid's lemma that $a \mid (y - y_0)$. Accordingly if we put

$$y - y_0 = at ,$$

then (2.4) shows that

$$x - x_0 = -bt .$$

Thus (2.3) also appears as a necessary condition for solutions x, y.

The uniqueness of prime factorization. A special case of Theorem 5 is the following.

THEOREM 5a: *If p is a prime and p divides $b_1b_2 \cdots b_k$, then p divides at least one of the factors b_i.*

Proof: If $p \mid b_k$, the statement of the theorem is true. If $p \nmid b_k$, then $(p, b_k) = 1$, since for common divisors only 1 and p, the divisors of p, have to be tested, and $p \nmid b_k$.† Then by Euclid's lemma $p \mid b_1b_2 \cdots b_{k-1}$. This argument can be repeated with the conclusion that p divides either b_k or b_{k-1} or $\cdots$ or b_2 or b_1, as had to be shown.

Note that in our example for nonuniqueness of factorization, the multiplicative system of natural numbers $\equiv 1 \pmod 3$, this theorem is false. For we have

$$4 \cdot 25 = 100 = 10 \cdot 10 ,$$

† This means "p does not divide b_k."

and 4 divides 100 but does not divide 10, although 4 plays the role of a primitive number in this system.

With this preparation we are now ready to prove the following basic theorem, which is often referred to as the Fundamental Theorem of Number Theory.

THEOREM 6: *The factorization of a natural number into prime factors is, up to the order of the factors, unique.*

Proof: Suppose that $N = p_1 p_2 \cdots p_k$ is a factorization of the natural number N into primes. Clearly there exists such a factorization. If $N = q_1 q_2 \cdots q_l$ is a second factorization, then $p_1 \mid N$ and hence $p_1 \mid q_1 q_2 \cdots q_l$. Thus p_1 divides some q_i. By rearranging the q's we may assume $p_1 \mid q_1$. Since q_1 is prime we have $p_1 = q_1$. Then canceling p_1 and q_1 we have

$$p_2 p_3 \cdots p_k = q_2 q_3 \cdots q_l .$$

It is clear that we may continue in this way canceling p's and q's, and we see that the set of p's is, except for rearrangements in the order of the factors, exactly the same as the set of q's. This proves the uniqueness of factorization of N into primes.

Recently some mathematicians [Hasse, the physicist F. A. Lindemann (later Lord Cherwell), Zermelo] have given proofs through mathematical induction avoiding Euclid's lemma. The argument of these proofs runs as follows: Suppose there exists an integer N which has two distinct factorizations into primes. Let us choose the smallest such integer N and suppose that

$$N = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l$$

are two of its factorizations. Then all the p's are distinct from all the q's, for otherwise we could cancel a p and a q and have a smaller integer with two distinct factorizations. We may suppose

$$p_1 \leqq p_2 \leqq \cdots \leqq p_k$$
$$q_1 \leqq q_2 \leqq \cdots \leqq q_l .$$

Then let us consider p_1 and q_1. Since $p_1 \neq q_1$ we may assume $p_1 < q_1$ without any restrictions. Consider now

$$P = p_1 q_2 \cdots q_l .$$

Clearly $p_1 \mid P$ and $p_1 \mid N$ so that $p_1 \mid (N - P)$. Now $N - P$ is positive since $N - P = (q_1 - p_1) q_2 q_3 \cdots q_l$. Let us write $q_1 - p_1$ as a product of primes, say $q_1 - p_1 = r_1 \cdots r_s$. Then

$$N - P = r_1 \cdots r_s q_2 \cdots q_l .$$

We have seen from the beginning that p_1 is none of the q_i. Since $p_1 \nmid (q_1 - p_1)$, it follows that $p_1 \nmid r_i$ for all i. Thus all the q's and r's are distinct from p_1. On

the other hand, we have seen that $N - P$ is divisible by p_1 and, therefore, $N - P = p_1 t_1 \cdots t_j$ where the t's are primes. Thus we have two distinct factorizations of $N - P$ in one of which p_1 appears and in the other not. This contradicts the minimum property of N.

The uniqueness of prime factorization, or actually Euclid's lemma, which is equivalent to it, was important in Greek mathematics since it could be used for the discussion of irrational numbers, a topic of great interest in Greek mathematics and philosophy. Euclid proved the following theorem (probably using an idea of Theaetetus, a pupil of Plato).

THEOREM 7: *If p is a prime then $\sqrt{p}$ is irrational.*

Proof: The theorem says that $p = \left(\frac{m}{n}\right)^2 = \frac{m^2}{n^2}$ is impossible in natural numbers m, n. Suppose it were possible. Then we would have the equation $pn^2 = m^2$. But a square number has always an *even* number of prime factors. The preceding equation is therefore impossible, since on the left-hand side we have a number with an *odd* number of prime factors, and on the right hand the same number with an *even* number of prime factors, contrary to the uniqueness of prime factorization.

Greatest common divisor and least common multiple. This is the moment to amplify some concepts which we have casually defined in the Introduction and so far used merely as abbreviations.

If a and b are natural numbers, let d be the *greatest common divisor* (G.C.D.) of a and b, that is, the largest natural number d which is a divisor of both a and b. We write $d = (a, b)$. Then $\left(\frac{a}{d}, \frac{b}{d}\right) = 1$. Indeed, if $\left(\frac{a}{d}, \frac{b}{d}\right) = \delta$, then $d \cdot \delta$ divides both a and b, and δ must be 1 because of the maximality of d. In view of Theorem 4 there exist integers x and y with $\frac{a}{d}x + \frac{b}{d}y = 1$ and thus $ax + by = d$.

THEOREM 8: *The G.C.D. of a and b is a linear combination of a and b with integer coefficients.*

COROLLARY: *Any common divisor of a and b divides the G.C.D. (a, b) of a and b. Indeed if $\delta \mid a$, $\delta \mid b$ then $\delta \mid (ax + by)$, or $\delta \mid d$.*

Given two natural numbers a, b there exist natural numbers which are multiples of both a and b. For example, ab is such a number. By the *least common multiple* (L.C.M.) of a and b we mean the least natural number which is a multiple of both a and b. We denote the L.C.M. of a and b by $\{a, b\}$.

THEOREM 9: *If a and b are natural numbers then $ab = \{a, b\}(a, b)$.*

Proof: Let $\mu = \dfrac{ab}{(a, b)}$. Then μ is a multiple of both a and b, since $\mu = \dfrac{a}{(a, b)} b$ is a multiple of b and $\mu = \dfrac{b}{(a, b)} a$ is also a multiple of a. If ν is another common multiple of a and b, then there exist integers x, y such that

$$\frac{\nu}{\mu} = \frac{\nu(a, b)}{ab} = \frac{\nu(ax + by)}{ab} = \frac{\nu x}{b} + \frac{\nu y}{a}.$$

Hence $\dfrac{\nu}{\mu}$ is an integer. Thus μ is indeed the L.C.M. of a and b, and as a by-product we derive the fact that the L.C.M. of a and b divides any common multiple of a and b.

Many theorems of a similar nature about the G.C.D. and the L.C.M. can be proved. We mention some of them in the following exercises.

Exercises.

1. Let $a, b, c, \cdots, k$ be natural numbers. Let $\delta = (a, b, c, \cdots, k)$ be their greatest common divisor, that is, the largest natural number which divides all of $a, b, c, \cdots, k$. Prove that there exist integers $x, y, z, \cdots, w$ with $\delta = ax + by + cz + \cdots + kw$.
2. Prove that $c \cdot (a, b) = (ca, cb)$.
3. Prove that $(a, (b, c)) = (a, b, c)$.
4. Prove that
$$(a_1, a_2, \cdots, a_l)(b_1, b_2, \cdots, b_m) = (a_1 b_1, \cdots, a_l b_m)$$
where the parenthesis on the right side contains all lm products $a_h b_j$, $h = 1, \cdots, l$, $j = 1, \cdots, m$.
5. Let n natural numbers $a_1, a_2, \cdots, a_n$ be given. Put
$$(a_1, a_2, \cdots, a_n) = d_1^{(n)},$$
$$(a_1 a_2, a_1 a_3, \cdots, a_{n-1} a_n) = d_2^{(n)},$$
and in general
$$(a_1 a_2 \cdots a_k, a_1 a_2 \cdots a_{k-1} a_{k+1}, \cdots, a_{n-k+1} a_{n-k+2} \cdots a_n) = d_k^{(n)},$$
where this parenthesis contains all products of k distinct a's as factors. We have by this definition
$$a_1 a_2 \cdots a_n = d_n^{(n)}.$$
Then show that
$$e_1 = d_1, e_2 = \frac{d_2}{d_1}, e_3 = \frac{d_3}{d_2}, \cdots, e_n = \frac{d_n}{d_{n-1}}$$
are integers. (We have written for the sake of brevity simply d_k for $d_k^{(n)}$.) Show moreover that the quotients
$$\frac{e_2}{e_1}, \frac{e_3}{e_2}, \cdots, \frac{e_n}{e_{n-1}}$$
are also integers.

3

Congruences

Congruence as equivalence relation. Let us examine more carefully the concept of *congruence*, which we have used informally in some of our preceding remarks. Let a, b, m be integers. Usually m is taken to be positive. We say *a is congruent to b modulo m* and write $a \equiv b \pmod m$ if $m \mid (a - b)$. The congruence relation $\equiv$ is an *equivalence relation* on the set of integers. That is, it is *reflexive:*

$$a \equiv a \pmod m,$$

symmetric:

$$\text{if } a \equiv b \pmod m \text{ then } b \equiv a \pmod m,$$

and *transitive:*

$$\text{if } a \equiv b \pmod m \text{ and } b \equiv c \pmod m, \text{ then } a \equiv c \pmod m.$$

This is clear. For $a \equiv a \pmod m$ merely says $m \mid (a - a)$. If $a \equiv b \pmod m$, then $m \mid (a - b)$, so that $m \mid (b - a)$ and $b \equiv a \pmod m$. Finally, if $a \equiv b \pmod m$ and $b \equiv c \pmod m$, then $m \mid (a - b)$ and $m \mid (b - c)$ so $m \mid (a - b) + (b - c)$, $m \mid (a - c)$, and $a \equiv c \pmod m$. Thus congruence is an equivalence relation, and the set of integers is partitioned into disjoint classes. Any two integers in the same class are congruent to one another, and no two integers in distinct classes are congruent to one another. How many congruence classes are there? We can answer this easily by exhibiting a specimen from each class. Clearly the integers $0, 1, \cdots, m - 1$ lie in different classes. And given any integer, we may add or subtract a suitable multiple of m and arrive at one of $0, 1, \cdots, m - 1$. Thus there are just m congruence classes modulo m, and the integers $0, 1, \cdots, m - 1$ form a set of representatives, one from each of the classes.

Congruences, like equalities, are equivalences. They behave like equalities with respect to addition and multiplication. If $a \equiv b \pmod m$ and $c \equiv d \pmod m$, then we have

$$a + c \equiv b + d \pmod m, \qquad a - c \equiv b - d \pmod m \tag{3.1}$$

and

$$ac \equiv bd \pmod m. \tag{3.2}$$

To prove the additive rule we note merely that if $m \mid (a - b)$ and $m \mid (c - d)$ then $m \mid ((a - b) + (c - d))$ or $m \mid ((a + c) - (b + d))$. For the multiplicative rule we observe $ac \equiv bc \pmod m$ because $m \mid (a - b)c$ and that

$bc \equiv bd$ because $m \mid b(c-d)$. Then $ac \equiv bd$ follows through the transitivity of the congruence relation.

Properties (3.1) and (3.2) can be expressed more concisely in terms of the congruence classes. They show that whichever element a we may choose from the class A and whichever b from class B, the sum will always lie in one class C which we may symbolically designate as $C = A + B$. Analogous remarks apply to $A - B$ and $A \cdot B$.

Among the classes is the class Z which contains the elements $z \equiv 0 \pmod m$. In the language of group theory we can say that the congruence classes form an additive Abelian group with the class Z as "zero element." The inverse of class A is the class A' which contains the negatives of all members of A.

What is true about addition, subtraction, and multiplication is, in general, not true about division. We cannot unrestrictedly divide congruences. For we have $2 \equiv 12 \pmod{10}$ but $1 \not\equiv 6 \pmod{10}$. On the other hand, $2 \equiv 24 \pmod{11}$ and $1 \equiv 12 \pmod{11}$, so that in this case we can safely divide by 2. In fact, if $ab \equiv ac \pmod m$ and $(a, m) = 1$, then $b \equiv c \pmod m$. For if $m \mid (ab - ac)$ then $m \mid a(b-c)$ and it follows from Euclid's lemma that $m \mid (b-c)$.

The congruence classes of elements prime to m form an Abelian group with respect to multiplication. The unit class U is evidently that one which contains the number 1. The existence of an inverse is also assured, since to any given number a prime to m there exists a number a^* such that

$$aa^* \equiv 1 \pmod m .$$

Indeed we need only refer to the Diophantine equation

$$aa^* + my = 1$$

which is solvable for a^* and y since $(a, m) = 1$ is assumed (Theorem 4).

Moreover, if we take a prime number p as modulus, then the congruence classes form a *finite field* (of order p). Indeed the classes form an additive group, and all classes with the exception of the zero class Z contain only elements prime to the modulus, and therefore form a multiplicative group.

Euler's function. The number of congruence classes modulo m is, as we have seen, m itself. The number of classes with elements prime to m is designated as $\varphi(m)$. The function φ is also called Euler's function. It is clear that for a prime number p we have $\varphi(p) = p - 1$. In addition, for $m = p^\alpha$ we see immediately that $\varphi(p^\alpha) = (p-1)p^{\alpha-1}$, since of the numbers $1, 2, \cdots, p^\alpha$ representing the p^α different congruence classes modulo p^α, only those are not prime to p^α which are divisible by p, in number $p^{\alpha-1}$. Thus $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$.

The explicit number $\varphi(m)$ for general m will be found if we prove

THEOREM 10: *For $(m, n) = 1$ we have $\varphi(mn) = \varphi(m) \cdot \varphi(n)$.*

This can be proved in several ways. Since we have the Farey theory of common fractions at our disposal, we may use it for this purpose. We observe first that the number of reduced proper fractions of denominator m is just $\varphi(m)$, since there are $\varphi(m)$ numerators $1 \leqq h \leqq m$ prime to m. Instead of insisting on *proper* fractions, we may just as well count the number of reduced fractions h/m with

$$a \leqq \frac{h}{m} < a + 1$$

for any given a. Now let $m = m_1 m_2$ with $(m_1, m_2) = 1$. We can decompose the reduced proper fraction h/m in a unique way into a sum of the partial fractions with denominators m_1 and m_2 as follows. Since $(m_1, m_2) = 1$, we may find integers u, v with

$$h = m_2 u + m_1 v. \tag{3.3}$$

The solution u, v is of course not unique, but we have observed (Theorem 4a) that the general solution is obtained from a special one (u_0, v_0) by

$$u = u_0 - t m_1, \qquad v = v_0 + t m_2$$

with an arbitary integer t. It is clear that there is one and only one solution u, v with such a t that

$$0 < u < m_1 .$$

($u = 0$ is excluded because it would lead in (3.3) to $m_1 \mid h$, which is against the assumption $(h, m_1 m_2) = 1$, h/m being a reduced fraction.)

We then have from (3.3)

$$\frac{h}{m} = \frac{h}{m_1 m_2} = \frac{u}{m_1} + \frac{v}{m_2}, \quad 0 < \frac{u}{m_1} < 1 . \tag{3.4}$$

Both fractions are reduced, since a common divisor of, let us say, u and m_1 would divide $m_1 m_2$ and according to (3.3) also h, contrary to $(h, m_1 m_2) = 1$. Although v/m_2 is certainly reduced, it is not necessarily a proper fraction; from

$$\frac{v}{m_2} = \frac{h}{m} - \frac{u}{m_1}$$

we can only infer

$$-\frac{u}{m_1} < \frac{v}{m_2} < 1 - \frac{u}{m_1} . \tag{3.5}$$

Now, conversely, let two proper reduced fractions u/m_1, v/m_2 be given. If this v/m_2 does not already satisfy (3.5), we replace it by

$$\frac{v'}{m_2} = \frac{v - m_2}{m_2} = \frac{v}{m_2} - 1 ,$$

which must then fit in (3.5). Then computing h/m from (3.4), there possibly replacing v/m_2 by v'/m_2, we obtain $(h/m) = (h/m_1m_2)$ as a proper reduced fraction.

In this way we obtain a one-to-one correspondence of proper reduced fractions h/m_1m_2 with pairs of proper reduced fractions u/m_1, v/m_2 (called *partial fractions*). That is, the number of admissible numerators h is the same as the number of pairs u, v' of numerators of the partial fractions. This statement can be expressed as $\varphi(m_1m_2) = \varphi(m_1) \cdot \varphi(m_2)$, proving our theorem.

It now follows by iteration that, for any finite set of pairwise coprime numbers $m_1, m_2, \cdots, m_l$ we have

$$\varphi(m_1m_2 \cdots m_l) = \varphi(m_1) \cdot \varphi(m_2) \cdots \varphi(m_l).$$

If we break down a given number n into powers of primes as factors, we can employ our knowledge of $\varphi(p^\alpha)$ and obtain for $n = p_1^{\alpha_1}p_2^{\alpha_2} \cdots p_k^{\alpha_k}$

$$\begin{aligned}\varphi(n) &= p_1^{\alpha_1}\left(1 - \frac{1}{p_1}\right) p_2^{\alpha_2}\left(1 - \frac{1}{p_2}\right) \cdots p_k^{\alpha_k}\left(1 - \frac{1}{p_k}\right) \\ &= n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right) \qquad (3.6) \\ &= n \prod_{p|n} \left(1 - \frac{1}{p}\right),\end{aligned}$$

which is Euler's formula.

Let us write down all fractions l/n of fixed denominator n, $0 < l/n \leqq 1$, whether reduced or not:

$$\frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \cdots, \frac{n-1}{n}, \frac{n}{n}. \qquad (3.7)$$

They are n in number. Some, as $1/n$ and $(n-1)/n$, are in reduced form. In the others we cancel the common divisors of numerator and denominator. The resulting reduced denominators are then divisors d of n; for each divisor d of n all reduced proper fractions of denominator d will appear in the list. They are $\varphi(d)$ in number, as we know. Thus counting the fractions in (3.7) according to their reduced denominators, we have the following important theorem.

THEOREM 11: *The Euler function $\varphi(n)$ has the property*

$$n = \sum_{d|n} \varphi(d) \qquad (3.8)$$

where the summation is extended over all divisors d of n.

If the numbers

$$(R) \qquad r_1, r_2, \cdots, r_\nu,$$

with $\nu = \varphi(m)$, are representatives of the classes prime to m, and if $(a, m) = 1$, then the numbers

$$(R') \qquad ar_1, ar_2, \cdots, ar_\nu$$

form another such system, only in a different arrangement. Any symmetric function of (R) will remain in its congruence class if (R) is replaced by (R'). In particular, therefore,

$$ar_1 \cdot ar_2 \cdot \cdots \cdot ar_\nu \equiv r_1 \cdot r_2 \cdot \cdots \cdot r_\nu \pmod m$$

or

$$(a^\nu - 1)\, r_1r_2 \cdots r_\nu \equiv 0 \pmod m.$$

Since $(r_1r_2 \cdots r_\nu, m) = 1$, we have $a^\nu - 1 \equiv 0 \pmod m$ by Euclid's lemma and have thus proved the following theorem.

THEOREM 12: *If $(a, m) = 1$ then*

$$a^{\varphi(m)} \equiv 1 \pmod m.$$

This theorem was derived by Euler as a generalization of the special case $m = p$, a prime number, $p \nmid a$

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

which was given by Fermat.

Higher congruences. The solvability of the linear congruence

$$ax \equiv c \pmod m$$

with $(a, m) = 1$ is implied by (2.2).

Let us consider congruences in much the same way that one considers equations in algebra. We look for solutions for congruences

$$a_0x^n + a_1x^{n-1} + \cdots + a_n \equiv 0 \pmod p$$

where p is prime. Since $x^{p-1} \equiv 1 \pmod p$ if $(x, p) = 1$, it follows that $x^p \equiv x \pmod p$ for all x. Using this fact we may get rid of all powers x^p, x^{p+1}, x^{p+2}, $\cdots$, by replacing them by lower powers of x. Hence we may suppose from the beginning that in all our congruences we have $n < p$.

Such a congruence need not have a solution. For example, $x^2 \equiv 3 \pmod 7$ has no solution since $0^2 \equiv 0$, $1^2 \equiv 1$, $2^2 \equiv 4$, $3^2 \equiv 2$, $4^2 \equiv 2$, $5^2 \equiv 4$, and $6^2 \equiv 1 \pmod 7$. However, just as in the theory of algebraic equations, a congruence of degree n and prime number modulus p can have at most n solutions.

THEOREM 13: *The number of solutions of the congruence*

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \equiv 0 \pmod p, \quad (a_0, p) = 1, \qquad (3.9)$$

is at most n.

Proof: The statement is correct if (3.9) has no solution. Suppose it has a solution x_1. Then

$$a_0x_1^n + a_1x_1^{n-1} + \cdots + a_n \equiv 0 \pmod{p}.$$

We subtract this from (3.9) and obtain

$$a_0(x^n - x_1^n) + a_1(x^{n-1} - x_1^{n-1}) + \cdots + a_{n-1}(x - x_1) \equiv 0 \pmod{p}, \tag{3.10}$$

which any x satisfying (3.9) must also satisfy. The congruence (3.10), however, can be rewritten as

$$(x - x_1) \cdot (a_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-1}) \equiv 0 \pmod{p}, \tag{3.11}$$

where the b's are certain expressions obtained from x_1 and the a's. Since p divides the product in (3.11), it must divide one of the factors. Any x satisfying (3.9) must therefore satisfy either

$$x - x_1 \equiv 0 \pmod{p}$$

or

$$a_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-1} \equiv 0 \pmod{p}.$$

The first alternative yields again x_1. The second may or may not yield a solution. That second congruence is of degree $(n - 1)$, the highest coefficient is again a_0, and, since the theory is true for the first degree congruence

$$a_0x + f_1 \equiv 0 \pmod{p},$$

it follows by induction that the theorem is true for any degree n.

Exercise: Prove Wilson's theorem: For any prime number p we have

$$(p - 1)! \equiv -1 \pmod{p}.$$

Hint: The congruences

$$(x - 1)(x - 2)\cdots(x - p + 1) \equiv 0 \pmod{p}$$

and

$$x^{p-1} \equiv 0 \pmod{p}$$

have the same solutions. Apply Theorem 13 on their difference.

Chinese Remainder Theorem. We shall sometimes need the so-called "Chinese Remainder Theorem," which for the sake of brevity we enunciate only for 3 moduli. Its generalization to any number of moduli is obvious.

THEOREM 14: *If the moduli m_1, m_2, m_3 are pairwise coprime*

$$(m_1, m_2) = (m_1, m_3) = (m_2, m_3) = 1,$$

and if

$$(a, m_1) = (b, m_2) = (c, m_3) = 1,$$

then the system of simultaneous congruences

$$\begin{aligned} ax &\equiv A \pmod{m_1} \\ bx &\equiv B \pmod{m_2} \\ cx &\equiv C \pmod{m_3} \end{aligned} \tag{3.12}$$

has a solution modulo $m_1m_2m_3$, for any A, B, C.

We leave the details of the proof to the reader. We observe only that, if the three systems

$$\begin{aligned} au &\equiv 1 \pmod{m_1} & av &\equiv 0 \pmod{m_1} & aw &\equiv 0 \pmod{m_1} \\ bu &\equiv 0 \pmod{m_2} & bv &\equiv 1 \pmod{m_2} & bw &\equiv 0 \pmod{m_2} \\ cu &\equiv 0 \pmod{m_3} & cv &\equiv 0 \pmod{m_3}, & cw &\equiv 1 \pmod{m_3} \end{aligned}$$

are solved, then $x = Au + Bv + Cw$ is obviously a solution of (3.12).

4

Decimal Fractions

We return again to the study of fractions. Common fractions, in the assemblage of Farey sequences, led us to linear congruences, Euclid's lemma, and thus the uniqueness of prime factorization.

This time we investigate the representation of common fractions as decimals. These are for themselves worthy of arithmetical studies, and moreover they will lead to a new approach to the Fermat-Euler theorem.

It is a familiar fact that any decimal that is either terminating or periodic is the decimal expansion of a rational number. We shall be concerned with the converse problem. Given a rational number, we shall see that its decimal expansion is either terminating or periodic and that the decimal digits have interesting arithmetic properties. Let us look at some examples:

$$0.2 = \frac{2}{10} = \frac{1}{5}$$

$$0.025 = \frac{25}{1000} = \frac{1}{40}.$$

These are finite decimals. But examine $\frac{1}{3} = 0.333\cdots$

$$\begin{array}{r} 0.333\cdots \\ 3\overline{)\ 1.000\cdots} \\ \underline{9} \\ 10 \\ \underline{9} \\ 10 \\ \underline{9} \\ 1 \end{array}$$

or

$$\frac{1}{7} = 0.142857\cdots$$

$$\begin{array}{r} 0.142857 \\ 7\overline{)\ 1.000000} \\ \underline{7} \\ 30 \\ \underline{28} \\ 20 \\ \underline{14} \\ 60 \\ \underline{56} \\ 40 \\ \underline{35} \\ 50 \\ \underline{49} \\ 1 \end{array}$$

both obtained by long division.

We say that these decimals are periodic: $\frac{1}{3}$ has a period of length 1, and $\frac{1}{7}$ has a period of length 6. In theory, in order to write down $\frac{1}{3}$, we must write an infinity of 3's, but we get around this by drawing a line over the period, thus:

$$\frac{1}{3} = 0.3333\cdots = 0.\overline{3}$$

$$\frac{1}{7} = 0.142857\ 142857\ 142857\cdots = 0.\overline{142857}.$$

This shows that 3 and 142857, respectively, are to be written out indefinitely after the decimal point. But now let us look at $\frac{1}{6}$

$$\begin{array}{r} 0.1666\cdots \\ 6\overline{)\ 1} \\ 10 \\ 40 \\ 40 \\ 40 \end{array}$$

This is written as $0.1\overline{6}$. The periodic decimals that start their period at the decimal point, such as $\frac{1}{3}$ and $\frac{1}{7}$, are called *pure* periodic decimals. The decimal

$\frac{1}{6}$ does not start its period at the decimal point, and thus it is not pure. This case is of little interest. We shall consider the two other cases.

I. FINITE DECIMALS: In general, if $A/2^\alpha 5^\beta$ is a proper reduced fraction (A integral), it has a finite decimal expansion because it can always be changed to a fraction with a denominator that is a power of 10. For if, say, $\alpha \geqq \beta$, then

$$\frac{A}{2^\alpha 5^\beta} = \frac{A \cdot 5^{\alpha-\beta}}{2^\alpha \cdot 5^\alpha} = \frac{A \cdot 5^{\alpha-\beta}}{10^\alpha} = 0.\underbrace{\cdots}_{\alpha \text{ places}}.$$

If $\beta > \alpha$, we multiply by $2^{(\beta-\alpha)}$, and there are β places. For example,

$$\frac{147}{200} = \frac{735}{1000} = 0.735 .$$

Conversely, if a fraction has a finite decimal expansion, multiplication by a suitable power of 10 gives an integer. In reduced form, therefore, its denominator contains only 2 and 5 as prime factors.

II. PURE PERIODIC DECIMALS: We shall show that, if A/m is a proper reduced fraction with $(10, m) = 1$, then A/m has a pure periodic decimal expansion. To investigate this, we first ask the question, "How do you find the successive digits in the decimals?" When we divide A by m we have

$$\begin{array}{rll} & 0.q_1q_2q_3\cdots & \\ m) & A & 0 < A = r_1 < m \\ & 10A & \\ & -mq_1 & 0 \leqq q_1 < 10 \\ \hline & r_2 & 0 < r_2 < m \\ & 10r_2 & \\ & -mq_2 & 0 \leqq q_2 < 10 \\ \hline & r_3 & 0 < r_3 < m \\ & \cdot & \\ & \cdot & \\ & \cdot & \end{array}$$

In general,

$$r_{j+1} = 10r_j - q_jm$$

where $0 \leqq q_j < 10$ and $0 < r_j < m$.

All the remainders r_j are prime to m. We can seeth is by induction. For $r_1 = A$ is prime to m. Since $r_{j+1} = 10r_j - q_jm$, we have $r_{j+1} \equiv 10r_j \pmod m$. Since $(r_j, m) = 1$ and $(10, m) = 1$, it follows that $(10r_j, m) = 1$, and hence that $(r_{j+1}, m) = 1$.

Since there are infinitely many r's and all the r's lie between 1 and $m - 1$, there must be two equal r's. Let the first r's that are equal be r_j and r_{j+l}. We must show $j = 1$. If $j > 1$, then r_{j-1} is defined. Then

$$r_j \equiv 10r_{j-1} \pmod m$$

and

$$r_{j+l} \equiv 10r_{j+l-1} \pmod m$$

so that

$$10r_{j-1} \equiv 10r_{j+l-1} \pmod m .$$

But $(m, 10) = 1$, and thus

$$r_{j-1} \equiv r_{j+l-1} \pmod m .$$

Now r_{j-1} and r_{j+l-1} lie between 1 and $m - 1$ inclusive and therefore are too small to differ by a multiple of m, unless they are equal. Thus $r_{j-1} = r_{j+l-1}$, and this contradicts the fact that r_j, r_{j+l} was the first equal pair. The first case of $r_j = r_{j+l}$ occurs when $j = 1$, and the sequence of r's is purely periodic with period l. It follows that the sequence of q's (that is, the sequence of decimal digits) is also purely periodic with period l.

We let $\lambda(m)$ denote the length of the period of the decimal expansion of $1/m$. Later we shall see that all proper reduced fractions A/m have the same period length $\lambda(m)$. We first show that $\lambda(m) \leqq \varphi(m)$. This is true because there are only $\varphi(m)$ residue classes prime to m, and at least two of $r_1, r_2, \cdots, r_{\varphi(m)+1}$ must be equal. Let us look at some numerical evidence. It can happen that $\lambda(m) = \varphi(m)$, for

$$\lambda(7) = \varphi(7) = 6$$

and

$$\lambda(17) = \varphi(17) = 16$$

as direct computation will show. We have mentioned $\frac{1}{7}$ above, and we find

$$\frac{1}{17} = 0.\overline{0588235294117647} .$$

But $\frac{1}{3} = 0.\overline{3}$, so that $\lambda(3) = 1$, while $\varphi(3) = 2$, so that $\lambda(3) < \varphi(3)$. In addition, we have $\frac{1}{21} = 0.\overline{047619}$ so that $\lambda(21) = 6$ while

$$\varphi(21) = 21 \prod_{p|21} \left(1 - \frac{1}{p}\right)$$

$$= 21\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{7}\right) = 12 .$$

We shall see later that if m has two or more different prime divisors then $\lambda(m) \leqq \frac{1}{2}\varphi(m)$. Thus for the maximum period $\varphi(m)$ to be attained, it is necessary that m be a power of a prime p. But this is not sufficient as we have seen for $p = 3$. In fact it is not even known whether there are infinitely many primes p for which the maximum period $\varphi(p) = p - 1$ is attained. We shall have more to say about this later.

If

$$\frac{A}{m} = \frac{r_1}{m} = 0.\overline{q_1q_2\cdots q_l}$$

then

$$\frac{r_2}{m} = 0.\overline{q_2 q_3 \cdots q_\lambda q_1}$$

$$\cdots$$

$$\frac{r_\lambda}{m} = 0.\overline{q_\lambda q_1 q_2 \cdots q_{\lambda-1}}$$

This is so because, if we had started with $A = r_j$ instead of $A = r_1$, we would have followed the same procedure and would have obtained the same q's and r's in periodic order, the period starting, however, at a different place. For example, we have already had (written now in abbreviated form)

$$\begin{array}{r|l} & 0.\overline{142857} \\ \hline 7 & 1.0 \\ & 30 \\ & 20 \\ & 60 \\ & 40 \\ & 50 \\ & 1 \\ & \cdot \\ & \cdot \\ & \cdot \end{array}$$

Corresponding to this we have

$$1/7 = 0.\overline{142857}$$
$$3/7 = 0.\overline{428571}$$
$$2/7 = 0.\overline{285714}$$
$$6/7 = 0.\overline{857142}$$
$$4/7 = 0.\overline{571428}$$
$$5/7 = 0.\overline{714285}.$$

Now let us look at $\frac{1}{41}$

$$\begin{array}{r|l} & 0.\overline{02439} \\ \hline 41 & 1.0 \\ & 100 \\ & 180 \\ & 160 \\ & 370 \\ & 1 \\ & \cdot \\ & \cdot \\ & \cdot \end{array}$$

The five residues r are 1, 10, 18, 16, 37.

THEOREM 15: *The residue classes (mod 41) containing* 1, 10, 18, 16, and 37 *form a multiplicative group.*

Proof:

$$1 \equiv 1 \cdot 10^0 \pmod{41}$$
$$10 \equiv 1 \cdot 10^1 \pmod{41}$$
$$18 \equiv 1 \cdot 10^2 \pmod{41}$$
$$16 \equiv 1 \cdot 10^3 \pmod{41}$$
$$37 \equiv 1 \cdot 10^4 \pmod{41}$$
$$1 \equiv 1 \cdot 10^5 \pmod{41}.$$

Thus all these r's are congruent to powers of 10 (mod 41). If

$$r \equiv 10^k \pmod{41}$$

and

$$s \equiv 10^l \pmod{41}$$

then

$$rs \equiv 10^{k+l} \pmod{41}, \text{ which is also a power of } 10.$$

If $r \equiv 10^k \pmod{41}$, $k = 1, 2, 3, 4, 5$, then $s \equiv 10^{5-k} \pmod{41}$ is the inverse, since

$$rs \equiv 10^k \cdot 10^{5-k} \equiv 10^{k+5-k} \equiv 10^5 \equiv 1 \pmod{41}.$$

This proves that the residue classes of 1, 10, 18, 16, 37 form a subgroup of the group of residue classes prime to 41.

There is no magic about the number 41. In general, if the r's which occur in the decimal expansion of $1/m$ are $r_1, r_2, \cdots, r_\lambda$, then, since $r_1 = 1$ and $r_{j+1} \equiv 10r_j \pmod{m}$, it follows under the assumption $(10, m) = 1$ that

$$r_j \equiv 10^{j-1} \pmod{m}, j = 1, 2, 3, \cdots.$$

Since all the r_j, $j = 1, 2, \cdots, \lambda$, are incongruent modulo m, it follows that all the powers 10^j, $j = 1, 2, \cdots, \lambda$, are incongruent modulo m. Further, $r_{\lambda+1} = r_1 = 1$, so that $10^\lambda \equiv 1 \pmod{m}$ and 10^λ is the least power of 10 which is congruent to 1 (mod m). In particular *there exists a power of* 10 *which is congruent to* 1 *(mod m)*. The residue classes containing $r_1, r_2, \cdots, r_\lambda$ are just the residue classes containing $10^0, 10^1, \cdots, 10^{\lambda-1}$. In view of the congruence $10^\lambda \equiv 1 \pmod{m}$, the residue classes containing $10^0, 10^1, \cdots, 10^{\lambda-1}$ form, just as in the case $m = 41$, a subgroup of the group of residue classes prime to m. Since the group of residue classes modulo m which are prime to m has order $\varphi(m)$, it follows from the elements of group theory that the order of the subgroup consisting of the classes containing $r_1, r_2, \cdots r_\lambda$ divides $\varphi(m)$. Thus we have the following important theorem.

THEOREM 16: *$\lambda(m)$ divides $\varphi(m)$.*

We shall, however, prove this theorem without assuming anything about groups and get some additional information about decimal expansions. For concreteness let us consider the case $m = 41$. One can give a proof for the general case along the same lines. Since 1, 10, 18, 16, 37 are the r's associated with the fraction 1/41, we see that 2, 20, 36, 32, 74 would be the r's associated with 2/41 except for the fact that $74 > 41$ is too large. But our scheme for generating the q's and r's shows that we are to reduce 74 modulo 41, and thus we obtain 2, 20, 36, 32, 33 as the sequence of r's for 2/41. Note that 2, 20, 36, 32, 33 represent just those residue classes modulo 41 which contain elements $\equiv 2 \cdot 10^k \pmod{41}$ for some k. Similarly, 3, 30, 13, 7, 29 are the r's for 3/41. Continuing in this way we see that any proper reduced fraction $A/41$ gives rise to just five r's and hence that the period length of all fractions $A/41$ is the same as that of 1/41. We also see that the $\varphi(41)$ numbers less than 41 and prime to 41 are divided in this way into a certain number of disjoint sets of $\lambda(41)$ integers each. If there are k sets, then $k \cdot \lambda(41) = \varphi(41)$, and hence $\lambda(41)$ divides $\varphi(41)$. This proves our assertion.

We have seen above that $10^{\lambda(m)} \equiv 1 \pmod{m}$, so that $m \mid (10^{\lambda(m)} - 1)$. If $\varphi(m) = k\lambda(m)$, then $10^{\varphi(m)} - 1 = 10^{k\lambda(m)} - 1$. If we use the identity $x^k - 1 = (x-1)(x^{k-1} + x^{k-2} + \cdots + 1)$ with $x = 10^{\lambda(m)}$, we see that $m \mid (10^{\varphi(m)} - 1)$ or that $10^{\varphi(m)} \equiv 1 \pmod{m}$. Thus we have proved a special case of the theorem of Fermat-Euler (Theorem 11):

If $(m, 10) = 1$ then

$$10^{\varphi(m)} \equiv 1 \pmod{m}.$$

In particular, if m is a prime p and $p \neq 2, 5$ we have

$$10^{p-1} \equiv 1 \pmod{p}.$$

Actually the significance of 10 in the statements of these theorems is only that we have used decimal expansions, which are expansions to the base 10. If we agree to work to the base q, then we again obtain the full Theorem 12 with a different proof.

FERMAT-EULER THEOREM: *If $(q, m) = 1$ then $q^{\varphi(m)} = 1 \pmod{m}$.*

5

Approximation of Real Numbers by Rational Numbers, Application to Sums of Two Squares; Prime Numbers in Certain Arithmetic Progressions

Dirichlet's pigeon-hole principle. So far we have dealt only with rational numbers. We are now going to establish some inequalities relating irrational numbers to rational numbers. We shall obtain some theorems which will be useful in giving other proofs of previously established results and in giving more insight into the nature of rational numbers.

Let γ be an irrational number. We wish to find out how closely we can approximate γ by a rational number h/k with denominator k. For any k the irrational γ will lie between h/k and $(h+1)/k$ for some h. Then either

$$0 < \gamma - \frac{h}{k} < \frac{1}{2k},$$

or

$$0 < \frac{h+1}{k} - \gamma < \frac{1}{2k},$$

where, because of the irrationality of γ, equality signs are certainly excluded.

For certain k, however, much better approximations are possible.

We use the "pigeon-hole principle" of Dirichlet which states that, if N objects are placed in $N-1$ pigeon-holes, then at least one hole will contain two or more objects. If, as usual, $[x]$ denotes the greatest integer not exceeding x, we take as objects the N real numbers

$$0 < n\gamma - [n\gamma] < 1 \qquad n = 1, 2, 3, \cdots, N.$$

As pigeon-holes we take the intervals $\left(0, \frac{1}{N}\right), \left(\frac{1}{N}, \frac{2}{N}\right), \cdots, \left(\frac{N-1}{N}, 1\right)$.

There are also N of these. We now have the following possibilities before us:

1. There is one "object" in every interval. Then there is in particular an integer $m \leqq N$ such that

$$0 < m\gamma - [m\gamma] < \frac{1}{N}.$$

Dividing by m we obtain

$$0 < \gamma - \frac{[m\gamma]}{m} < \frac{1}{Nm} \leqq \frac{1}{m^2}.$$

2. The first interval $(0, 1/N)$ does not contain an "object." All N objects will then find their place in the remaining $N-1$ intervals, and thus one interval will contain two different objects. Therefore, with

$$0 < m < n \leqq N,$$

we have

$$|(n\gamma - [n\gamma]) - (m\gamma - [m\gamma])| < \frac{1}{N}$$

or

$$|(n-m)\gamma - ([n\gamma] - [m\gamma])| < \frac{1}{N}.$$

Here we have found two positive integers $k = n - m < N$ and

$$h = [n\gamma] - [m\gamma] < N$$

which fulfill

$$|k\gamma - h| < \frac{1}{N}.$$

It again follows that

$$\left|\gamma - \frac{h}{k}\right| < \frac{1}{Nk} \leqq \frac{1}{k^2}.$$

Thus, in either case, γ is approximated to within $1/k^2$ by a fraction h/k. More precisely, we found a fraction h/k with denominator $k \leqq N$ so that

$$\left|\gamma - \frac{h}{k}\right| < \frac{1}{kN}. \tag{5.1}$$

We can also obtain this result by the use of Farey fractions, with even a slight improvement. In the Farey sequence of order N, we know that we can find two consecutive terms, a/b and c/d, such that

$$\frac{a}{b} < \gamma < \frac{c}{d}.$$

We consider the mediant $\frac{a+c}{b+d}$. Then γ lies on one or the other side of $\frac{a+c}{b+d}$:

$$\frac{a}{b} < \gamma < \frac{a+c}{b+d} \tag{5.2}$$

or

$$\frac{a+c}{b+d} < \gamma < \frac{c}{d}. \tag{5.3}$$

Since the mediant is not present in the Farey sequence of order N, we have $b + d \geqq N + 1$. Thus either

$$0 < \gamma - \frac{a}{b} < \frac{a+c}{b+d} - \frac{a}{b} = \frac{1}{b(b+d)} \leqq \frac{1}{(N+1)b} \tag{5.4}$$

or

$$0 < \frac{c}{d} - \gamma < \frac{c}{d} - \frac{a+c}{b+d} = \frac{1}{d(b+d)} \leqq \frac{1}{(N+1)d}, \tag{5.5}$$

which again implies (5.1), with even the improvement $(N+1)$ instead of N. Altogether, we have proved the following theorem.

THEOREM 17: *For an irrational γ and a positive integer N there always exists a fraction h/k, with denominator $k \leqq N$ such that*

$$\left|\gamma - \frac{h}{k}\right| < \frac{1}{(N+1)k}. \tag{5.6}$$

Remark: Our reasoning remains valid if γ is not irrational but is replaced by a reduced fraction l/m with denominator $m > N$ so that l/m is not found in the Farey sequence of order N. Then, however, it may happen that $\frac{l}{m} = \frac{a+c}{b+d}$, and this possibility will not allow us to state a strict inequality in (5.6). We therefore obtain another theorem.

THEOREM 17a: *If l/m is a reduced fraction of denominator $m > N$, then there always exists a fraction h/k with denominator $k \leqq N$ such that*

$$\left|\frac{l}{m} - \frac{h}{k}\right| \leqq \frac{1}{(N+1)k}. \tag{5.7}$$

Equality takes place here for $m = N + 1$.

Sums of two squares. Before we pursue the question of approximation further, let us look at some applications of this theorem.

THEOREM 18: *If n and A are positive integers such that $n \mid (A^2 + 1)$, $n \geqq 2$, then there exist integers s and t such that $n = s^2 + t^2$.*

Proof: Let us take $N = [\sqrt{n}] < n$ in Theorem 17a. It is evident that our hypothesis implies $(n, A) = 1$. Therefore, according to Theorem 17a, there exists a fraction r/s in lowest terms such that

$$\left|\frac{A}{n} - \frac{r}{s}\right| \leqq \frac{1}{(N+1)s} \qquad 0 < s \leqq N.$$

This yields

$$|As - rn| \leq \frac{n}{N+1} = \frac{n}{[\sqrt{n}]+1} < \sqrt{n}\,.$$

If we put $t = As - rn$, we have

$$t^2 + s^2 = (As - rn)^2 + s^2 = s^2(A^2+1) - 2Asrn + r^2n^2$$

and thus

$$t^2 + s^2 \equiv 0 \pmod{n}\,.$$

On the other hand,

$$t^2 + s^2 < (\sqrt{n})^2 + ([\sqrt{n}])^2 \leq 2n.$$

Thus $t^2 + s^2$ is a positive multiple of n but less than $2n$. This leaves only the possibility

$$n = t^2 + s^2\,.$$

Remark: Note that $(s, t) = 1$:

$$(s, t) = (s, As - rn) = (s, rn) = (s, n)$$

and

$$n = s^2 + t^2 = s^2(A^2+1) - 2Asrn + r^2n^2$$

or

$$1 = s^2\frac{A^2+1}{n} - 2Asr + r^2n\,.$$

We recall that $(A^2+1)/n$ is an integer, and see that any common divisor of s and n must divide 1. Thus

$$(s, t) = (s, n) = 1.$$

COROLLARY: *If $n \mid (A^2 + B^2)$, $n \geq 2$, and $(A, B) = 1$, then there exist integers s, t with $n = s^2 + t^2$.*

We have the algebraic identity

$$(A^2 + B^2)(C^2 + D^2) = (AC + BD)^2 + (AD - BC)^2\,.$$

But, since $(A, B) = 1$, we know that we can pick C and D such that $AD - BC = 1$. Thus we have

$$(A^2 + B^2)(C^2 + D^2) = (AC + BD)^2 + 1.$$

If $n \mid (A^2 + B^2)$, therefore, then $n \mid (T^2 + 1)$, where $AC + BD = T$. Then, however n is of the form $t^2 + s^2$ by the theorem we have just proved.

Prime numbers of the form $4n + 1$. Recall that we have proved the existence of an infinite number of primes $\equiv 3 \pmod 4$. Now we are able to prove the existence of an infinite number of primes $p \equiv 1 \pmod 4$. We use an extension of Euclid's argument and Theorem 18.

THEOREM 19: *There exists an infinite number of primes $p \equiv 1 \pmod 4$.*

Proof: Suppose there are only a finite number, and let them be 5, 13, 17, $\cdots$, p. Form the number $N = (2 \cdot 5 \cdot 13 \cdot \cdots \cdot p)^2 + 1$. This N is congruent to 1 (mod 4), but it cannot be prime because all the primes $\equiv 1 \pmod 4$ are less than it. In addition, it has no even factors, because it is odd. Thus any factor of N must be of the form $a^2 + b^2$, because it divides a number of the form $A^2 + 1$. It is $\equiv 1 \pmod 4$, therefore, because every square is congruent to 1 or 0 (mod 4), and, since $a^2 + b^2$ is odd, $a^2 + b^2 \equiv 1 \pmod 4$. Consequently, any prime divisor of N is also a prime $\equiv 1 \pmod 4$. Such a prime divisor cannot equal any one of the set 5, 13, 17, $\cdots$, p, since these obviously do not divide N, and so this contradicts the assumption that there are only a finite number of primes $\equiv 1 \pmod 4$.

Some similar theorems. We may use the method of Theorem 18 to prove a similar theorem about numbers of the form $s^2 + 3t^2$.

THEOREM 20: *If $n \mid (A^2 + 3)$ where n is odd and $(3, n) = 1$, then there exist integers s and t with $n = s^2 + 3t^2$.*

Proof: Our hypothesis implies $(A, n) = 1$. Let us pick an $N < n$, but leave it temporarily unspecified. From (5.7) we have the existence of coprime integers r, s such that

$$\left|\frac{A}{n} - \frac{r}{s}\right| \leq \frac{1}{(N+1)s} \qquad 0 < s \leq N.$$

Continuing as before,

$$|As - rn| \leq \frac{n}{N+1}\,,$$

$$t = As - rn\,,$$

$$t^2 + 3s^2 = (As - rn)^2 + 3s^2 = s^2(A^2+3) - 2Asrn + r^2n^2\,,$$

$$t^2 + 3s^2 \equiv 0 \pmod{n}\,.$$

Now following our old proof, we know that

$$t^2 + 3s^2 \leq \frac{n^2}{(N+1)^2} + 3s^2 < \frac{n^2}{N^2} + 3N^2\,.$$

If we are to prove this theorem by methods similar to those used to prove Theorem 18, we must make $t^2 + 3s^2$ small. This amounts to choosing an

integer N for which $n^2/N^2 + 3N^2$ is small. Let us set $F(x) = (n^2/x^2) + 3x^2$ and use differential calculus to find the real number $x > 0$ which makes $F(x)$ a minimum. We have

$$F'(x) = \frac{-2n^2}{x^3} + 6x\,.$$

Note that $F'(x) = 0$ when $2n^2 = 6x^4$, or when $x = \sqrt{n}\,3^{-\frac{1}{4}}$. Thus, choosing $N = [\sqrt{n}\,3^{-\frac{1}{4}}]$ is probably good enough for our purposes.

We have then

$$t^2 + 3s^2 \leqq \frac{n^2}{([\sqrt{n}\,3^{-\frac{1}{4}}] + 1)^2} + 3[\sqrt{n}\,3^{-\frac{1}{4}}]^2$$

$$< \frac{n^2}{n\,3^{-\frac{1}{4}}} + 3^{\frac{1}{4}}\,n = 2\sqrt{3}n\,.$$

Since $3 < 2\sqrt{3} < 4$, and since $n \mid (t^2 + 3s^2)$, we have

$$t^2 + 3s^2 = n,\ 2n,\ \text{or}\ 3n\,.$$

If $t^2 + 3s^2 = n$, we are finished. If $t^2 + 3s^2 = 3n$, then $3 \mid t$: say $3T = t$. Then $n = s^2 + 3T^2$, which proves our theorem in this case. Finally, $t^2 + 3s^2 = 2n$ is impossible. For, since n is odd, $2n \equiv 2 \pmod 4$. But $s^2, t^2 \equiv 0$ or $1 \pmod 4$, and no combination of $\binom{0}{1} + \binom{0}{3} \pmod 4$ is congruent to 2 (mod 4). Our theorem is proved.

We can now state a corollary as we did before.

COROLLARY: *If* $2 \nmid n$, $3 \nmid n$, $(A, B) = 1$, *and* $n \mid (A^2 + 3B^2)$, *then* $n = t^2 + 3s^2$.

For, as before,

$$(A^2 + 3B^2)(C^2 + 3D^2) = (AC + 3BD)^2 + 3(AD - BC)^2\,,$$

and, as before, we can set $AD - BC = 1$. Then $n \mid (T^2 + 3)$, and hence $n = t^2 + 3s^2$, by our theorem.

Now we can prove the existence of an infinite number of primes $p \equiv 1 \pmod 3$.

THEOREM 21: *There exists an infinite number of primes* $p \equiv 1 (mod\ 3)$.

Proof: The proof is like the proof of Theorem 19. Here we consider

$$(2 \cdot 7 \cdot 13 \cdot 19 \cdots p)^2 + 3\,.$$

This is divisible by neither 2 nor 3; hence any prime divisor q has the form $q = t^2 + 3s^2 \equiv 1 \pmod 3$. But $q \neq 7, 13, 19, \cdots, p$.

THEOREM 22: *If* $n \mid (A^2 + 2)$, A *odd, then there exist natural numbers* x, y *such that* $n = x^2 + 2y^2$.

We leave the proof as an exercise for the reader.

The situation is, however, quite different in the following theorem.

THEOREM 23: *If* $n \mid (A^2 + 5)$ *with* $(A, 5) = 1$, *then either the equation* $n = x^2 + 5y^2$ *or the equation* $2n = x^2 + 5y^2$ *is solvable, but not both.*

Proof: We shall consider the clause "but not both" later (p. 72). Otherwise we proceed as before. We choose $N = [n^{\frac{1}{4}} \cdot 5^{-\frac{1}{4}}]$ and find a fraction r/s such that

$$\left|\frac{A}{n} - \frac{r}{s}\right| \leqq \frac{1}{(N+1)s}\,, 0 < s \leqq N.$$

We then have

$$|As - rn| \leqq \frac{n}{N+1} < n^{\frac{1}{4}}5^{\frac{1}{4}}.$$

We put $t = As - rn$ and have

$$t^2 + 5s^2 = (As - rn)^2 + 5s^2 = (A^2 + 5)s^2 - 2Asrn + r^2n^2\,.$$

Since n divides $A^2 + 5$,

$$t^2 + 5s^2 \equiv 0 \pmod n\,.$$

On the other hand,

$$t^2 + 5s^2 < n\sqrt{5} + 5n\,5^{-\frac{1}{4}} = 2\sqrt{5}\,n\,.$$

Now since $4 < 2\sqrt{5} < 5$, and $t^2 + 5s^2$ is a multiple of n, we can have only the cases

$$t^2 + 5s^2 = n \text{ or } 2n \text{ or } 3n \text{ or } 4n\,.$$

The first two cases are noted in the theorem. The fourth case implies

$$t^2 + 5s^2 \equiv t^2 + s^2 \equiv 0 \pmod 4\,.$$

Therefore t and s are both even, say $t = 2t_1$, $s = 2s_1$, and thus

$$4t_1^2 + 20s_1^2 = 4n\,,$$

$$t_1^2 + 5s_1^2 = n\,.$$

The theorem is also proved in this case. There remains the case

$$t^2 + 5s^2 = 3n\,. \tag{5.8}$$

Now the following is an identity in x and y:

$$2\{(x - 2y)^2 + 5(x + y)^2\} = 3\{(2x + y)^2 + 5y^2\} \tag{5.9}$$

as direct computation shows. Equation (5.8) entails

$$0 \equiv t^2 + 5s^2 \equiv t^2 - s^2 \equiv (t + s)(t - s) \pmod 3\,,$$

and, since the signs of s and t are arbitrary, we can assume

$$t \equiv s \pmod 3 \tag{5.10}$$

We now put

$$\begin{aligned} t &= x - 2y \\ s &= x + y \end{aligned} \tag{5.11}$$

or

$$x = \frac{t+2s}{3}, \qquad y = \frac{s-t}{3} \tag{5.12}$$

where x and y are both *integers* in view of (5.10). If we now insert (5.11) and (5.12) in (5.9), we obtain

$$2(t^2 + 5s^2) = 3\{(2x+y)^2 + 5y^2\},$$

and because of (5.8)

$$2n = (2x+y)^2 + 5y^2 = x_1^2 + 5y^2,$$

which reduces case 3 to case 2 and finishes the proof.

That the alternative in Theorem 23 is not due to a fault of our method and cannot be avoided is shown by the pair of examples:

(*a*) $A = 11 \qquad A^2 + 5 = 126$

$21 \mid 126 \qquad n_1 = 21 = 1^2 + 5 \cdot 2^2,$

but $2n_1 = 42$ is not expressible as $x^2 + 5y^2$;
(*b*) $18 \mid 126$, $n_2 = 18$ is not expressible as $x^2 + 5y^2$ but $2n_2 = 36 = 4^2 + 5 \cdot 2^2$.†

† In the background of Theorem 23 lies, of course, the fact that the discriminant -20 has 2 classes (as a matter of fact 2 genera) of binary quadratic forms. The identity

$$2(2x^2 + 2xy + 3y^2) = (2x+y)^2 + 5y^2$$

together with the alternative of the theorem show that the forms $x^2 + 5y^2$ and $2x^2 + 2xy + 3y^2$ belong to different classes.

6

Better Rational Approximation of Irrational Numbers; Ford Circles and Hurwitz's Theorem

Goodness of approximation. Let us return to our approximation theory. If γ is irrational and N is a given positive integer, we have shown the existence of a rational approximation h/k, $k \leq N$, such that

$$\left|\gamma - \frac{h}{k}\right| < \frac{1}{k(N+1)}.$$

In particular, we then have

$$\left|\gamma - \frac{h}{k}\right| < \frac{1}{k^2}.$$

These elementary arguments may be used to show the existence of *infinitely many* fractions h/k such that

$$\left|\gamma - \frac{h}{k}\right| < \frac{1}{k^2}.$$

However, using Farey sequences we can prove a much stronger theorem.

THEOREM 24: *If γ is irrational, there exist infinitely many fractions h/k such that*

$$\left|\gamma - \frac{h}{k}\right| < \frac{1}{2k^2}.$$

Proof: Suppose that in the Farey sequence of order N we have

$$\frac{a}{b} < \gamma < \frac{c}{d}.$$

We wish to show that either

$$\gamma - \frac{a}{b} < \frac{1}{2b^2} \quad \text{or} \quad \frac{c}{d} - \gamma < \frac{1}{2d^2}. \tag{6.1}$$

Assume the contrary,

$$\gamma - \frac{a}{b} \geq \frac{1}{2b^2} \quad \text{and} \quad \frac{c}{d} - \gamma \geq \frac{1}{2d^2}. \tag{6.2}$$

Then

$$\frac{c}{d} - \frac{a}{b} \geq \frac{1}{2b^2} + \frac{1}{2d^2} = \frac{b^2 + d^2}{2b^2d^2} .$$

Now we know from the theory of Farey sequences that

$$\frac{c}{d} - \frac{a}{b} = \frac{1}{bd} ,$$

so that

$$0 \geq \frac{b^2 + d^2}{2b^2d^2} - \frac{1}{bd} = \frac{(b-d)^2}{2b^2d^2} .$$

But this is only possible for $b = d$. In addition, this implies $b = d = 1$, since $ad - bc = -1$. Therefore, for a Farey sequence of order $N > 1$, (6.2) must be false and (6.1) true, which proves the theorem.

Thus we are led to the question whether infinitely many fractions h/k exist with still better approximation

$$\left| \gamma - \frac{h}{k} \right| < \frac{1}{ck^2} \tag{6.3}$$

with $c > 2$, and what the greatest value of c may be.

This question was answered completely by A. Hurwitz.

THEOREM 25: *For any positive $c \leq \sqrt{5}$, the inequality* (6.3) *has infinitely many solutions. There exists, however, an irrational γ for which* (6.3) *has only finitely many solutions in case $c > \sqrt{5}$.*

The second half of this theorem is easily settled. We take $\gamma = \dfrac{1+\sqrt{5}}{2}$ and $0 < \alpha < 1$ and ask for fractions h/k such that

$$\left| \frac{h}{k} - \frac{1+\sqrt{5}}{2} \right| < \frac{\alpha}{\sqrt{5}k^2} . \tag{6.4}$$

If we write

$$\frac{h}{k} - \frac{1+\sqrt{5}}{2} = \frac{\theta}{\sqrt{5}k^2}$$

the preceding inequality means $|\theta| < \alpha < 1$. We have

$$h - \frac{k}{2} = \frac{\sqrt{5}k}{2} + \frac{\theta}{\sqrt{5}k}$$

and after squaring and rearranging

$$h^2 - hk - k^2 = \theta + \frac{\theta^2}{5k^2} .$$

The integer on the left-hand side cannot be 0 for integers h, k not both zero, so we have

$$1 \leq \left| \theta + \frac{\theta^2}{5k^2} \right| < \alpha + \frac{\alpha^2}{5k^2}$$

or

$$k^2 < \frac{\alpha^2}{5(1-\alpha)} .$$

This restricts the denominators k, and (6.4) then also permits only finitely many h to each of the finitely many k. Therefore, (6.3) has indeed only finitely many solutions h/k for

$$c = \frac{\sqrt{5}}{\alpha} > \sqrt{5} .$$

The Ford circles. We postpone for a while the proof of the first half of Theorem 25 in order to prepare a new tool. This is a geometric figure introduced by L. R. Ford, consisting of certain circles which have something to do with the Farey sequences. It is useful to think of these "Ford circles" as lying in the complex z-plane of $z = x + iy$.

Let $C(h/k)$ be the circle with center at $h/k + i/2k^2$ and radius $1/2k^2$. Thus $C(h/k)$ is the circle

$$\left| z - \left(\frac{h}{k} + \frac{i}{2k^2} \right) \right| = \frac{1}{2k^2}$$

which lies in the upper half-plane and is tangent to the x-axis at $x = (h/k)$. These circles have an important property.

THEOREM 26: *Two distinct Ford circles never intersect. They are tangent if and only if their fractions are adjacent ones in some Farey sequence.*

Proof: The centers of two distinct circles $C(h/k)$, $C(l/m)$ are $h/k + i/2k^2$, $l/m + i/2m^2$ with $hm - kl \neq 0$. See Fig. 1. The square of the distance between their centers is therefore

$$d^2 = \left(\frac{h}{k} - \frac{l}{m} \right)^2 + \left(\frac{1}{2k^2} - \frac{1}{2m^2} \right)^2 .$$

The square of the sum of their radii is

$$t^2 = \left(\frac{1}{2k^2} + \frac{1}{2m^2} \right)^2 .$$

Since

$$d^2 - t^2 = \left(\frac{h}{k} - \frac{l}{m} \right)^2 - \frac{1}{k^2m^2} = \frac{(hm-kl)^2 - 1}{k^2m^2} \geq 0 ,$$

the circles do not intersect. The circles are tangent if and only if equality holds, that is, if and only if $hm - kl = \pm 1$. In this case, according to Theorem 1, h/k and l/m are adjacent fractions of some Farey sequence (e.g., that of order $N = k + m - 1$).

Note that, if $C(h/k)$ and $C(l/m)$ are tangent and if their point of tangency is $w = u + iv$, then u divides the segment $(h/k, l/m)$ in the ratio $1/k^2 : 1/m^2 = m^2 : k^2$. Therefore,

$$u = \frac{(h/k)k^2 + (l/m)m^2}{m^2 + k^2} = \frac{hk + lm}{m^2 + k^2} \tag{6.5}$$

is a rational number, and similarly v is also rational. The points of tangency of Ford circles have rational coordinates.

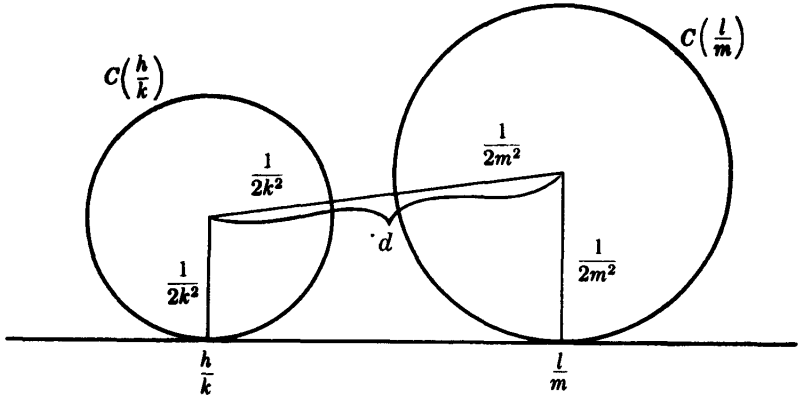

Figure 1

Circular triangles. Now the configuration of all Ford circles shows circular triangles which are formed by arcs of mutually tangent circles (see Figs. 2 and 3). Let the circles be $C(H/K)$, $C(h/k)$, $C(h_1/k_1)$ with $0 < K < k < k_1$. The fractions H/K and h/k appear as adjacent in the Farey sequence of order k. The fraction h_1/k_1 is not in this Farey sequence. However, since h_1/k_1 is adjacent to H/K as well as to h/k (because of the tangency of their Ford circles), it must be the mediant $\frac{h_1}{k_1} = \frac{H + h}{K + k}$, in a Farey sequence of higher order,

$$h_1 = H + h, \qquad k_1 = K + k. \tag{6.6}$$

We now return to the proof of Theorem 25.

Let γ be irrational. The vertical line $x = \gamma$ cannot reach any point of tangency of the Ford circles. It must pass, therefore, through the interior of infinitely many of the circular triangles.

Let one such triangle be formed by arcs of the circles $C(h/k)$, $C(H/K)$, and $C(h_1/k_1)$. Since the configuration of Farey circles is symmetric about the line

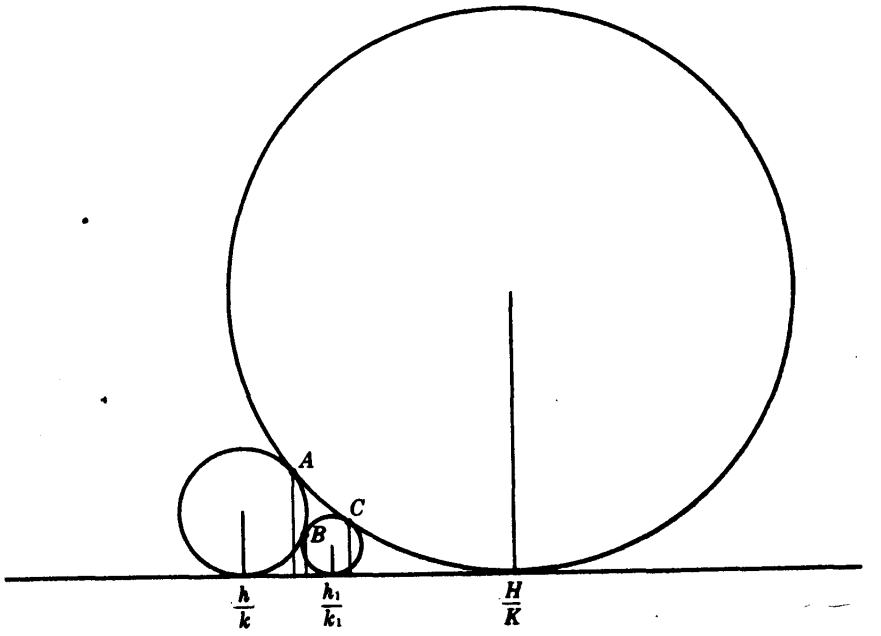

Figure 2

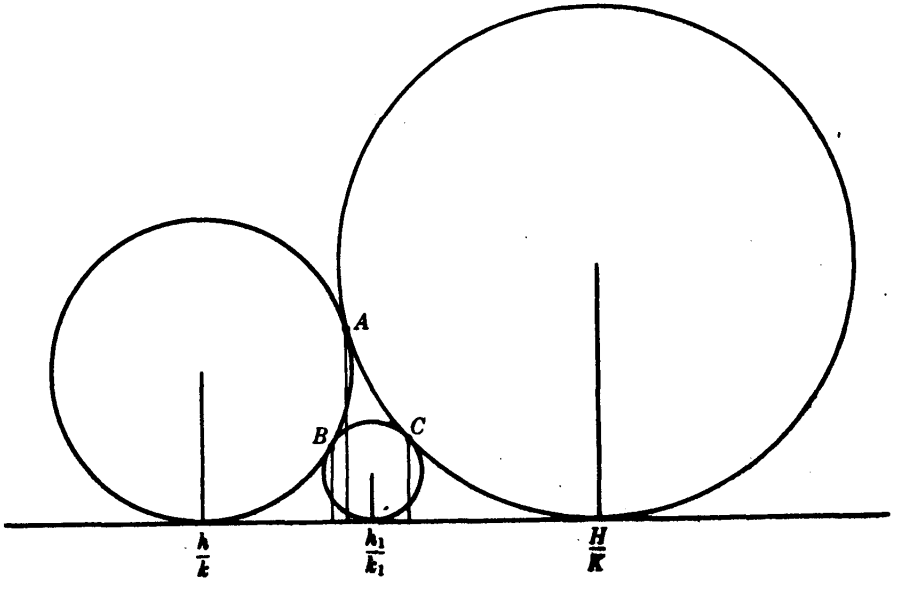

Figure 3

$x = \frac{1}{2}$, we may replace γ by $1 - \gamma$ if necessary and then assume without loss of generality

$$\frac{h}{k} < \gamma < \frac{H}{K} \tag{6.7}$$

and also, of course,

$$\frac{h}{k} < \frac{h_1}{k_1} < \frac{H}{K} \tag{6.8}$$

Let A be the point of tangency of $C(h/k)$ and $C(H/K)$; B of $C(h/k)$ and $C(h_1/k_1)$; and C of $C(h_1/k_1)$ and $C(H/K)$. Let a, b, c, be the x-coordinates of A, B, C. Then, in view of (6.5), we have

$$a = \frac{hk + HK}{k^2 + K^2}, \; b = \frac{hk + h_1k_1}{k^2 + k_1^2}, \; c = \frac{h_1k_1 + HK}{k_1^2 + K^2}.$$

Thus

$$\begin{aligned} c - a &= \frac{HK(k^2 - k_1^2) + h_1k_1(k^2 + K^2) - hk(k_1^2 + K^2)}{(k^2 + K^2)(k_1^2 + K^2)} \\ &= \frac{kK(Hk - hK) + k_1k(h_1k - hk_1) - Kk_1(Hk_1 - h_1K)}{(k^2 + K^2)(k_1^2 + K^2)} \\ &= \frac{kK + k_1k - Kk_1}{(k^2 + K^2)(k_1^2 + K^2)} \end{aligned}$$

because of Theorem 1 and the ordering (6.8). Finally with (6.6) we obtain

$$c - a = \frac{kK + k^2 - K^2}{(k^2 + K^2)(k_1^2 + K^2)}.$$

Let us put

$$s = \frac{k}{K} > 1.$$

Then we have

$$c - a = \frac{s^2 + s - 1}{K^2(s^2 + 1)((s + 1)^2 + 1)},$$

and, since

$$s^2 + s - 1 > 1 + 1 - 1 = 1,$$

we conclude

$$c - a > 0.$$

Similarly we obtain

$$\begin{aligned} c - b &= \frac{kk_1 + Kk + Kk_1}{(k^2 + k_1^2)(k_1^2 + K^2)} = \frac{K^2 + 3Kk + k^2}{(k^2 + k_1^2)(K^2 + k_1^2)} \\ &= \frac{s^2 + 3s + 1}{K^2((s + 1)^2 + 1)(s^2 + (s + 1)^2)} > 0. \end{aligned}$$

However, $b - a$ can be positive or negative. Indeed an analogous computation yields

$$b - a = \frac{k^2 - kK - K^2}{(k^2 + K^2)(k^2 + k_1^2)} = \frac{s^2 - s - 1}{K^2(s^2 + 1)(s^2 + (s + 1)^2)}.$$

Here

$$s^2 - s - 1 = \left(s - \frac{1}{2} - \frac{\sqrt{5}}{2}\right)\left(s - \frac{1}{2} + \frac{\sqrt{5}}{2}\right), \tag{6.9}$$

and, since

$$s - \frac{1}{2} + \frac{\sqrt{5}}{2} > \frac{1}{2} + \frac{\sqrt{5}}{2} > 0,$$

we see that the sign of $b - a$ is the same as the sign of

$$s - \frac{1 + \sqrt{5}}{2}.$$

We now consider separately the cases $b > a$ and $b < a$ ($b = a$ is impossible since the polynomial in (6.9) cannot vanish for rational values of s.)

Case I: $b > a$ or $s > \frac{1 + \sqrt{5}}{2}$.

We intend to show that in this case

$$\left|\gamma - \frac{H}{K}\right| < \frac{1}{\sqrt{5}K^2}.$$

Indeed, because in this case a and c are the extreme abscissas of the circular triangle which is hit by $x = \gamma$, we have

$$a < \gamma < c < \frac{H}{K}$$

and

$$\begin{aligned} 0 < \frac{H}{K} - \gamma < \frac{H}{K} - a &= \frac{H}{K} - \frac{hk + HK}{k^2 + K^2} \\ &= \frac{k}{K(k^2 + K^2)} = \frac{s}{K^2(s^2 + 1)}. \end{aligned}$$

Now

$$\left(s - \frac{1 + \sqrt{5}}{2}\right)\left(s + \frac{1 - \sqrt{5}}{2}\right) > 0,$$

since the first factor and *a fortiori* the second factor are positive under our assumption. Thus

$$s^2 - \sqrt{5}s + 1 > 0$$

or

$$s^2 + 1 > \sqrt{5}s\,,$$

$$\frac{s}{s^2+1} < \frac{1}{\sqrt{5}}$$

so that we have in this case:

$$0 < \frac{H}{K} - \gamma < \frac{1}{\sqrt{5}K^2}\,,$$

as announced.

Case II: $b < a$ or $1 < s < \dfrac{1+\sqrt{5}}{2}$.

This time we intend to show

$$\left|\frac{h_1}{k_1} - \gamma\right| < \frac{1}{\sqrt{5}k_1^2}\,.$$

Under our conditions

$$b < a < c$$

and therefore

$$b < \gamma < c\,. \tag{6.10}$$

We derive first that

$$\frac{h_1}{k_1} - b > c - \frac{h_1}{k_1}\,. \tag{6.11}$$

This is fairly clear geometrically since C is higher on the circle $C(h_1/k_1)$ than B, $C(H/K)$ having a radius larger than that of $C(h/k)$. Explicitly we have

$$\frac{h_1}{k_1} - b = \frac{h_1}{k_1} - \frac{hk + h_1k_1}{k^2 + k_1^2} = \frac{k}{k_1(k^2+k_1^2)}$$

$$c - \frac{h_1}{k_1} = \frac{h_1k_1 + HK}{k_1^2 + K^2} - \frac{h_1}{k_1} = \frac{K}{k_1(K^2+k_1^2)}$$

and so indeed

$$\frac{k}{k_1(k^2+k_1^2)} - \frac{K}{k_1(K^2+k_1^2)} = \frac{(k-K)(k_1^2 - kK)}{k_1(k^2+k_1^2)(K^2+k_1^2)}$$

$$= \frac{k^3 - K^3}{k_1(k^2+k_1^2)(K^2+k_1^2)} > 0\,.$$

Thus because of (6.10) we have

$$\frac{h_1}{k_1} - c < \frac{h_1}{k_1} - \gamma < \frac{h_1}{k_1} - b$$

and because of (6.11)

$$\left|\frac{h_1}{k_1} - \gamma\right| < \frac{h_1}{k_1} - b = \frac{k}{k_1(k^2+k_1^2)} = \frac{1}{k_1^2}\frac{s(s+1)}{s^2 + (s+1)^2}\,.$$

Here, again, however,

$$\frac{s(s+1)}{s^2 + (s+1)^2} < \frac{1}{\sqrt{5}}\,.$$

Indeed, the opposite

$$\frac{s(s+1)}{s^2 + (s+1)^2} > \frac{1}{\sqrt{5}}$$

would entail $s(s+1)\sqrt{5} > 2s^2 + 2s + 1$

or

$$s^2(\sqrt{5}-2) + s(\sqrt{5}-2) - 1 > 0$$

or

$$0 < s^2 + s - \frac{1}{\sqrt{5}-2} = s^2 + s - (\sqrt{5}+2)$$

$$= \left(s + \frac{1}{2} + \frac{2+\sqrt{5}}{2}\right)\left(s + \frac{1}{2} - \frac{2+\sqrt{5}}{2}\right)$$

$$= \left(s + \frac{3}{2} + \frac{\sqrt{5}}{2}\right)\left(s - \frac{1}{2} - \frac{\sqrt{5}}{2}\right),$$

which is wrong since Case II presupposes $s - \frac{1}{2} - (\sqrt{5}/2) < 0$.

Hence in both cases we have found a fraction l/m so that

$$\left|\gamma - \frac{l}{m}\right| < \frac{1}{\sqrt{5}m^2}\,. \tag{6.12}$$

Here l/m was determined by the triangle ABC traversed by the line $x = \gamma$. Since this line cuts across infinitely many such triangles, there are infinitely many fractions l/m satisfying (6.12), which proves Hurwitz's theorem.

7

Primitive Congruence Roots; The Regular Heptadecagon

Primitive congruence roots. We have seen in Theorem 13 that a congruence of degree n modulo a prime number p cannot have more than n solutions modulo p.

This maximal number can be attained as the example $x^{p-1} - 1 \equiv 0 \pmod{p}$ shows, which has the solutions $x \equiv 1, 2, \cdots, (p-1)$. It follows that if $d \mid p-1$ then $x^d \equiv 1 \pmod{p}$ has d solutions. For let $p - 1 = md$. We then have the identity

$$x^{p-1} - 1 = (x^d - 1)(x^{(m-1)d} + x^{(m-2)d} + \cdots + 1).$$

But the congruence $x^d - 1 \equiv 0 \pmod{p}$ has at most d solutions, and the congruence $x^{(m-1)d} + x^{(m-2)d} + \cdots + 1 \equiv 0 \pmod{p}$ has at most $(m-1)d$ solutions. If $x^d - 1 \equiv 0 \pmod{p}$ had less than d solutions, then $x^{p-1} - 1 \equiv 0 \pmod{p}$ would have less than $d + (m-1)d = p - 1$ solutions, which is not the case. Thus $x^d - 1 \equiv 0 \pmod{p}$ does indeed have d solutions.

If $\delta \mid d$, then $x^\delta - 1$ divides $x^d - 1$ algebraically. Thus any solution of $x^\delta - 1 \equiv 0 \pmod{p}$ is also a solution of $x^d - 1 \equiv 0 \pmod{p}$. We say that a solution x_0 of $x^{p-1} - 1 \equiv 0 \pmod{p}$ *belongs to the exponent* d if it is a solution of $x^d - 1 \equiv 0 \pmod{p}$, but is not a solution of $x^\delta - 1 \equiv 0 \pmod{p}$ for any $\delta < d$. We say then also that x_0 is a *primitive solution* of $x^d - 1 \equiv 0 \pmod{p}$. If a solution belongs to the exponent d; then necessarily $d \mid (p-1)$. For if $e = (d, p-1)$, then $e = md + r(p-1)$ for suitable integers m and r (see Theorem 8), and then $x^e \equiv (x^d)^m (x^{p-1})^r \equiv 1 \cdot 1 \equiv 1 \pmod{p}$. Since $e \leqq d$, it follows from the minimality of d that $e = d$ and hence $d \mid (p-1)$. Thus we may separate all the solutions of $x^{p-1} - 1 \equiv 0$ into classes of solutions, each class containing those solutions which belong to the exponent d. We need only consider divisors d of $p-1$. We let $\psi(d)$ denote the number of solutions which belong to the exponent d, that is. the number of primitive solutions of $x^d - 1 \equiv 0 \pmod{p}$.

THEOREM 27: $\psi(d) = \varphi(d)$.

Proof: The statement is true for $d = 1, 2$. For $\varphi(1) = 1$, and the congruence $x - 1 \equiv 0 \pmod{p}$ has the unique solution $x \equiv 1$. Also $\varphi(2) = 1$,

and the congruence $x^2 - 1 = (x-1)(x+1) \equiv 0 \pmod{p}$ has the unique primitive solution $x \equiv p - 1$. Let us prove our theorem by induction. Suppose $\psi(\delta) = \varphi(\delta)$ for all $\delta < d$. Since every solution of $x^d - 1 \equiv 0 \pmod{p}$ is a primitive solution of $x^\delta - 1 \equiv 0 \pmod{p}$ for some divisor δ of d, we have

$$d = \sum_{\delta \mid d} \psi(\delta).$$

Since, by induction, we have $\psi(\delta) = \varphi(\delta)$ for all divisors δ of d except perhaps d itself, we may write

$$d = \sum_{\delta \mid d} \varphi(\delta) + \psi(d) - \varphi(d).$$

On the other hand we have from Theorem 11

$$d = \sum_{\delta \mid d} \varphi(\delta).$$

Thus

$$\psi(d) - \varphi(d) = 0,$$

and this is the assertion of our theorem.

As a corollary to our theorem we see that primitive solutions of $x^d - 1 \equiv 0 \pmod{p}$ do indeed exist for all divisors d of $p-1$. In particular, there does exist a primitive solution to the congruence $x^{p-1} - 1 \equiv 0 \pmod{p}$ and in fact our theorem insures that there exist $\varphi(p-1)$ primitive solutions. A primitive solution of $x^{p-1} - 1 \equiv 0 \pmod{p}$ is called a *primitive root modulo* p.

These notions have some bearing on the length of the periods of decimals. We found $\lambda(m)$ to be the smallest exponent $d > 0$ so that $10^d \equiv 1 \pmod{m}$. (See the paragraph preceding Theorem 16.) In particular, for the modulus $m = p$ we have $\lambda(p) = \varphi(p) = p - 1$ if 10 is a primitive root modulo p. For example, 10 is a primitive root modulo 17 since $10^{16} \equiv 1 \pmod{17}$, but no lower power of 10 is congruent to 1 modulo 17. Thus $\lambda(17) = 16$ as we have found on page 27. Similarly 10 is a primitive root modulo 7, i.e., $\lambda(7) = 6$. But 10 is not a primitive root modulo 41, since we had $\lambda(41) = 5 < 40$.

In view of the Euler theorem

$$a^{\varphi(m)} \equiv 1 \pmod{m} \text{ for } (a, m) = 1,$$

it is reasonable to ask whether there exist *primitive roots modulo* m for composite m, that is, solutions of the congruence

$$x^{\varphi(m)} - 1 \equiv 0 \pmod{m}$$

which are not also solutions of $x^d - 1 \equiv 0 \pmod{m}$ with $d < \varphi(m)$. Here several cases have to be treated separately.

THEOREM 28: *If p is an odd prime, there exist primitive roots mod p^α for any α.*

Proof: Let r be a primitive root modulo p. Then $r^{p-1} \equiv 1 \pmod{p}$ and so

$$r^{p-1} = 1 + pt$$

where t is an integer. For any integer x we have

$$(r + px)^{p-1} = \binom{p-1}{0} r^{p-1} + \binom{p-1}{1} r^{p-2} px + \binom{p-1}{2} r^{p-3} (px)^2 + \cdots + \binom{p-1}{p-1} (px)^{p-1}$$

$$= 1 + p(t - r^{p-2}x + pa)$$

where a is an integer. Now let x be a solution x_0 of the congruence

$$t - r^{p-2}x \equiv 1 \pmod{p}.$$

Since $(r, p) = 1$, we know that such a solution exists. Then

$$(r + px_0)^{p-1} = 1 + pt_0, \qquad (t_0, p) = 1 .$$

We assert that $r + px_0$ is a primitive root modulo p^α. To see this we must compute $(r + px_0)^{p^\beta(p-1)}$ for any positive integer β. We have first

$$(r + px_0)^{p(p-1)} = (1 + pt_0)^p$$

$$= 1 + \binom{p}{1} pt_0 + \binom{p}{2} (pt_0)^2 + \cdots + \binom{p}{p} (pt_0)^p$$

$$= 1 + p^2 t_1$$

where $(t_1, p) = 1$ since $t_1 \equiv t_0 \pmod{p}$. Similarly we have

$$(r + px_0)^{p^2(p-1)} = (1 + p^2 t_1)^p = 1 + p^3 t_2$$

where $(t_2, p) = 1$. Continuing in this way, we arrive at the general formula

$$(r + px_0)^{p^\beta(p-1)} = 1 + p^{\beta+1} t_\beta, \qquad (t_\beta, p) = 1 .$$

Let d be the least positive integer with

$$(r + px_0)^d \equiv 1 \pmod{p^\alpha}. \tag{7.1}$$

We must show $d = p^{\alpha-1}(p-1) = \varphi(p^\alpha)$. From the preceding congruence we have *a fortiori* $(r + px_0)^d \equiv 1 \pmod{p}$ and hence $r^d \equiv 1 \pmod{p}$. Thus, since r is a primitive root mod p, we have $(p-1) \mid d$. But on the other hand we have $d \mid p^{\alpha-1}(p-1)$ because of (7.1), and it follows that $d = (p-1)p^\beta$ where $\beta \leqq \alpha - 1$. Thus we have

$$(r + px_0)^{p^\beta(p-1)} \equiv 1 \pmod{p^\alpha}.$$

On the other hand, our previous computations show

$$(r + px_0)^{p^\beta(p-1)} = 1 + p^{\beta+1} t_\beta, \qquad (t_\beta, p) = 1,$$

from which it follows that $\alpha \leqq \beta + 1$. Thus $\beta + 1 \leqq \alpha \leqq \beta + 1$, so that $\alpha = \beta + 1$ and then $d = p^\beta(p-1) = p^{\alpha-1}(p-1)$. This completes the proof.

Note that if r is a primitive root modulo p^2 then we do indeed have $r^{p-1} = 1 + pt$ where t is prime to p, and our construction shows that r is a primitive root mod p^α for all $\alpha = 1, 2, 3, \cdots$.

Let us give an example. 7 is a primitive root modulo 5, but since $7^4 = 1 + 2400 \equiv 1 \pmod{25}$, 7 is not a primitive root modulo 25. Our construction tells us we must solve the congruence

$$\frac{2400}{5} - 7^2 x \equiv 1 \pmod{5}$$

or

$$480 - 343x \equiv 1 \pmod{5}$$

$$-3x \equiv 1 \pmod{5}$$

so that $x_0 = 3$ is a solution. Then $r + px_0 = 7 + 15 = 22$ is a primitive root modulo 25 and in fact modulo all powers 5^α.

Our theorem may be given a group theoretic interpretation: It says that the group of $\varphi(p^\alpha)$ residue classes modulo p^α prime to p^α is cyclic. Any residue class containing a primitive root modulo p^α is a generator of the group.

The theorem is false for $p = 2$:

2 has a primitive root: $\varphi(2) = 1$, $1^1 = 1$

4 has a primitive root: $\varphi(4) = 2$, $3^1 \equiv 3$, $3^2 \equiv 1 \pmod{4}$. But 8 has no primitive root since $\varphi(8) = 4$, while

$$1^2 \equiv 1, \quad 3^2 \equiv 1, \quad 5^2 \equiv 1, \quad 7^2 \equiv 1 \pmod{8} .$$

If d is the smallest exponent such that $a^d \equiv 1 \pmod{m}$, $(a, m) = 1$, then we say that *a belongs to the exponent d modulo m.*

Problem: Show that the highest exponent to which an odd number can belong modulo 2^γ, $\gamma \geqq 3$, is not $\varphi(2^\gamma) = 2^{\gamma-1}$, but is $\frac{1}{2}\varphi(2^\gamma) = 2^{\gamma-2}$. Show that the number 5 belongs to the exponent $2^{\gamma-2}$ modulo 2^γ.

If m is divisible by two distinct primes, and $m \neq 2p^\alpha$ where p is an odd prime, then there is no primitive root modulo m. For suppose $m = m_1 m_2$ where $(m_1, m_2) = 1$ and neither of m_1, m_2 is 1 or 2. Suppose $(r, m) = 1$. We have $r^{\varphi(m_1)} \equiv 1 \pmod{m_1}$ and $r^{\varphi(m_2)} \equiv 1 \pmod{m_2}$. Now $\varphi(x)$ is even unless $x = 1$ or $x = 2$. Thus

$$r^{\varphi(m_1)\varphi(m_2)/2} \equiv 1 \pmod{m_1} \quad \text{and} \quad r^{\varphi(m_2)\varphi(m_1)/2} \equiv 1 \pmod{m_2} .$$

It follows that

$$r^{\varphi(m_1)\varphi(m_2)/2} \equiv 1 \pmod{m_1 m_2}$$

and, since $\varphi(m_1 m_2) = \varphi(m_1)\varphi(m_2)$, this means

$$r^{\varphi(m)/2} \equiv 1 \pmod{m} .$$

Thus any number r prime to m belongs to an exponent at most $\varphi(m)/2$ modulo m. Therefore there can be no primitive roots modulo m. A further consequence of what we have proved is the fact that $\lambda(m)$, the length of the decimal period of $1/m$, is at most $\varphi(m)/2$ in this case.

Problem: Let p be an odd prime. Show that there exist primitive roots modulo $2p^\alpha$.

The regular polygon of 17 sides (the regular heptadecagon or in brief, 17-gon).

We are going to apply the theory of primitive congruence roots to the problem of cyclotomy: the division of the circumference of the circle in equal parts, as it was first done by Gauss. The most spectacular case is that of the regular 17-gon, which we shall discuss in detail. The resulting formulae imply the surprising fact that the regular 17-gon can be constructed by means of ruler and compass. In the next chapter we shall investigate some more general problems of cyclotomy.

Let us now consider our regular 17-gon. We consider it as a figure in the complex plane, its vertices being complex numbers, and we assume that it is inscribed in the unit circle about the origin with one of its vertices at the point of the complex number 1. Let $z = x + iy$ be any other vertex of the polygon. Since $|z| = 1$ we can write $z_j = \cos\theta_j + i\sin\theta_j$ $j = 1, 2, \cdots, 16$. The vertices of the regular 17-gon divide the circumference into equal parts, so that we have $\theta_j = j\theta_1$, where $\theta_1 = (2\pi/17)$. The 17th vertex is again 1:

$$1 = \cos 17\theta_1 + i\sin 17\theta_1 = (\cos\theta_1 + i\sin\theta_1)^{17} = z_1^{17}.$$

Thus z_1 is a root of the equation $z^{17} - 1 = 0$. Since $z^{17} - 1 = (z-1)(z^{16} + z^{15} + \cdots + z + 1)$ and $z_1 \neq 1$, it follows that z_1 is a root of the "cyclotomic equation"

$$z^{16} + z^{15} + \cdots + z + 1 = 0\,. \tag{7.2}$$

Our problem, therefore, is to find an explicit formula for a solution of this equation, or at least an explicit formula for its real part $x = \cos\theta$ or for its imaginary part $y = \sin\theta$.

Since 17 is a prime, there exist primitive roots modulo 17. For example, 10 is a primitive root modulo 17. We set up the following table modulo 17:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
10^n	1	10	15	14	4	6	9	5	16	7	2	3	13	11	8	12

(7.3)

As we have seen, each residue which is prime to 17 occurs once on the bottom row, since 10 is a primitive root. Let us order the exponents of (7.2) according to the order of (7.3). We can rewrite (7.2) as

$$z + z^{10} + z^{15} + z^{14} + \cdots + z^8 + z^{12} = -1\,. \tag{7.4}$$

We see that the exponents are successive powers of 10 modulo 17. Now we shall follow the method of Gauss and break (7.4) into two sums which he called "periods": namely,

$$\eta_1 = z + z^{15} + z^4 + z^9 + z^{16} + z^2 + z^{13} + z^8$$
$$\eta_2 = z^{10} + z^{14} + z^6 + z^5 + z^7 + z^3 + z^{11} + z^{12}\,.$$

Here we have taken in η_1 and in η_2 every second summand of (7.4): in η_1 those which are listed in (7.3) below an even n (including 0), and in η_2 those below an odd n. We have immediately

$$\eta_1 + \eta_2 = -1\,.$$

We would like to find the product $\eta_1\eta_2$, which contains 64 terms. This is not as bad as it looks if we multiply in a special way. Let us try multiplying each element in η_1 with the element in η_2 that is directly below it, and summing. The successive products are $zz^{10} = z^{11}$, $z^{15}z^{14} = z^{29} = z^{12}$, $z^4z^6 = z^{10}$, and so on. We finally get the sum, after we reduce exponents modulo 17:

$$z^{11} + z^{12} + z^{10} + z^{14} + z^6 + z^5 + z^7 + z^3\,.$$

But this is just η_2! Why is this so? The reason is that, as we have said, the successive exponents of z in (7.4) differ by a factor of 10 modulo 17. It follows that the successive exponents which occur in η_1 and η_2 differ by a factor of 100 modulo 17. Say z^a and z^b are members of η_1 and η_2, respectively, and their product is z^{a+b} in either η_1 or η_2. Now the next terms of η_1 and η_2 are z^{100a} and z^{100b}, respectively, and their product is just the next member of the period in which z^{a+b} belongs, namely $z^{100(a+b)}$. In this way, we always will get in the product all the members of η_1 or of η_2 if we get one of them. Which of the η's we get depends only on one element, say the first. Similarly, if we were to multiply each member of η_1 by the member of η_2 below and one step to the right of it, then all of the products would be in the same period of η. Thus, since $z \cdot z^{14} = z^{15}$ is in η_1, it follows that the seven remaining products $z^{15} \cdot z^6 = z^4, \cdots, z^8 \cdot z^{10} = z$ appear in η_1, and the sum of all eight products is η_1. Now we can do the same thing again, taking each member of η_1 with the member of η_2 that is two, three, four, and so on steps to the right below it; we will always get a period η as a sum of eight products. The period we get will be just the period containing the first product. Thus

we have very little computation to do and we find that $\eta_1\eta_2$ is the sum of the terms:

$$\begin{aligned} z^{11} + \quad \cdots \quad &= \eta_2 \\ z^{15} + \quad \cdots \quad &= \eta_1 \\ z^{7} + \quad \cdots \quad &= \eta_2 \\ z^{6} + \quad \cdots \quad &= \eta_2 \\ z^{8} + \quad \cdots \quad &= \eta_1 \\ z^{4} + \quad \cdots \quad &= \eta_1 \\ z^{12} + \quad \cdots \quad &= \eta_2 \\ z^{13} + \quad \cdots \quad &= \eta_1 \end{aligned}$$

Thus $\eta_1\eta_2 = 4\eta_1 + 4\eta_2 = -4$. Since we know the sum and product of η_1 and η_2, we may form the quadratic equation which has η_1 and η_2 for its roots. The equation is $y^2 + y - 4 = 0$. We can solve this by the quadratic formula, and the roots are seen to be

$$\eta_1, \eta_2 = \tfrac{1}{2}(-1 \pm \sqrt{1 + 16}) = \tfrac{1}{2}(-1 \pm \sqrt{17}) . \tag{7.5}$$

We see that 17 appears under the radical. In general, if we start with a p-gon, we can show that η_1 and η_2 are quadratic irrationalities and that $\pm p$ appears under the radical.† Now let us take new periods, η_1', η_2', η_3', and η_4', forming η_1', η_2', from η_1, and η_3', η_4', from η_2, in the same way that we formed η_1, η_2 from η. That means taking into each new sum only every second summand of the old sum:

$$\begin{aligned} z \ + z^4 + z^{16} + z^{13} &= \eta_1' \\ z^{15} + z^9 + z^2 \ + z^8 \ &= \eta_2' \\ z^{10} + z^6 + z^7 \ + z^{11} &= \eta_3' \\ z^{14} + z^5 + z^3 \ + z^{12} &= \eta_4' . \end{aligned} \tag{7.6}$$

We see immediately that

$$\begin{aligned} \eta_1' + \eta_2' &= \eta_1 \\ \eta_3' + \eta_4' &= \eta_2 . \end{aligned}$$

Now let us form $\eta_1' \cdot \eta_2'$. The same trick that we used before works now, and we find $\eta_1'\eta_2'$ is the sum of the terms

$$\begin{aligned} z^{16} + \quad \cdots \quad &= \eta_1' \\ z^{10} + \quad \cdots \quad &= \eta_3' \\ z^{3} + \quad \cdots \quad &= \eta_4' \\ z^{9} + \quad \cdots \quad &= \eta_2' . \end{aligned}$$

The sum is:

$$\eta_1'\eta_2' = \eta_1' + \eta_3' + \eta_4' + \eta_2' = \eta_1 + \eta_2 = -1 .$$

Thus η_1' and η_2' satisfy the equation

$$w^2 - \eta_1 w - 1 = 0 .$$

† We show later that $\eta_1 - \eta_2$ is a Gaussian sum; see (9.3), (10.1).

From the quadratic formula we find

$$\eta_1', \eta_2' = \tfrac{1}{2}(\eta_1 \pm \sqrt{\eta_1^2 + 4}) . \tag{7.7}$$

Similarly, we can write down

$$\eta_3', \eta_4' = \tfrac{1}{2}(\eta_2 \pm \sqrt{\eta_2^2 + 4}) . \tag{7.8}$$

Now we claim that the choice of the distribution of signs in (7.5) and (7.7) determines the distribution of signs for (7.8). To show this, we will expand,

$$(\eta_1' - \eta_2')(\eta_3' - \eta_4') = \eta_1'\eta_3' - \eta_1'\eta_4' - \eta_2'\eta_3' + \eta_2'\eta_4' .$$

Here, again, the same trick works, and we have $\eta_1'\eta_3'$ the sum of

$$\begin{aligned} z^{11} + \quad \cdots \quad &= \eta_3' \\ z^{7} + \quad \cdots \quad &= \eta_3' \\ z^{8} + \quad \cdots \quad &= \eta_3' \\ z^{12} + \quad \cdots \quad &= \eta_4' , \end{aligned}$$

so that

$$\eta_1'\eta_3' = 2\eta_3' + \eta_2' + \eta_4' ;$$

$\eta_1'\eta_4'$ is the sum of

$$\begin{aligned} z^{15} + \quad \cdots \quad &= \eta_2' \\ z^{6} + \quad \cdots \quad &= \eta_3' \\ z^{4} + \quad \cdots \quad &= \eta_1' \\ z^{13} + \quad \cdots \quad &= \eta_1' , \end{aligned}$$

$\eta_1'\eta_4' = 2\eta_1' + \eta_2' + \eta_3'$; $\eta_2'\eta_3'$ is the sum of

$$\begin{aligned} z^{8} + \quad \cdots \quad &= \eta_2' \\ z^{4} + \quad \cdots \quad &= \eta_1' \\ z^{5} + \quad \cdots \quad &= \eta_4' \\ z^{9} + \quad \cdots \quad &= \eta_2' \end{aligned}$$

$\eta_2'\eta_3' = 2\eta_2' + \eta_1' + \eta_4'$; and finally, $\eta_2'\eta_4'$ is the sum of

$$\begin{aligned} z^{12} + \quad \cdots \quad &= \eta_4' \\ z^{3} + \quad \cdots \quad &= \eta_4' \\ z \ + \quad \cdots \quad &= \eta_1' \\ z^{10} + \quad \cdots \quad &= \eta_3' , \end{aligned}$$

$\eta_2'\eta_4' = 2\eta_4' + \eta_1' + \eta_3'$.

Thus we have obtained

$$(\eta_1' - \eta_2')(\eta_3' - \eta_4') =$$
$$2\eta_3' + \eta_2' + \eta_4' - 2\eta_1' - \eta_3' - \eta_2' - 2\eta_3' - \eta_1' - \eta_4' + 2\eta_4' + \eta_1' + \eta_2' .$$

Or, canceling and contracting, we have

$$(\eta_1' - \eta_2')(\eta_3' - \eta_4') = -2\eta_1' - 2\eta_2' + 2\eta_3' + 2\eta_4' = -2(\eta_1 - \eta_2) .$$

This is a very important result, and from it we have

$$\eta_3' - \eta_4' = \frac{-2(\eta_1 - \eta_2)}{\eta_1' - \eta_2'} . \tag{7.9}$$

Now the left-hand side certainly depends upon which of η_3' or η_4' gets the plus sign and which the minus sign in (7.8). But the right-hand side depends only on the arbitrary choice of the signs among η_1, η_2, η_1', η_2', and thus we see that the distribution of signs in η_3', η_4' is dependent upon the distribution among the other four.

Now, let us go on with this breaking up into periods. We can break up η_1' into periods as follows:

$$z + z^{16} = \eta_1''$$
$$z^4 + z^{13} = \eta_2'' .$$

Now we have immediately that

$$\eta_1'' + \eta_2'' = \eta_1'$$
$$\eta_1''\eta_2'' = z^5 + z^{14} + z^3 + z^{12} = \eta_4' .$$

Thus η_1'' and η_2'' are roots of the equation

$$u^2 - \eta_1' u + \eta_4' = 0$$

and we see that

$$\eta_1'', \eta_2'' = \tfrac{1}{2}(\eta_1' \pm \sqrt{\eta_1'^2 - 4\eta_4'}) .$$

Finally, we have

$$z + z^{16} = \eta_1''$$
$$z \cdot z^{16} = 1 .$$

Thus z and z^{16} satisfy the equation

$$z^2 - \eta_1'' z + 1 = 0 . \tag{7.10}$$

We now assemble all the information we have gained about the η's. In (7.5) we make a choice of the signs and set

$$\eta_1 = \tfrac{1}{2}(-1 + \sqrt{17}), \qquad \eta_2 = \tfrac{1}{2}(-1 - \sqrt{17}) ,$$

which implies

$$\eta_1 - \eta_2 > 0 .$$

If we again choose the upper sign in (7.7) for η_1', we have

$$\eta_1' = \tfrac{1}{2}(\eta_1 + \sqrt{\eta_1^2 + 4})$$

and have herewith decided

$$\eta_1' - \eta_2' > 0 .$$

Then (7.9) shows

$$\eta_4' - \eta_3' > 0 .$$

Hence in (7.8) the choice of the sign is determined for η_4':

$$\eta_4' = \tfrac{1}{2}(\eta_2 + \sqrt{\eta_2^2 + 4}) .$$

Carrying out these computations we get

$$\eta_1^2 = \tfrac{1}{4}(-1 + \sqrt{17})^2 = \tfrac{1}{4}(18 - 2\sqrt{17})$$
$$\eta_2^2 = \tfrac{1}{4}(-1 - \sqrt{17})^2 = \tfrac{1}{4}(18 + 2\sqrt{17})$$

and, hence,

$$\eta_1' = \tfrac{1}{2}\{-1 + \sqrt{17} + \sqrt{2(17 - \sqrt{17})}\}$$
$$\eta_4' = \tfrac{1}{2}\{-1 - \sqrt{17} + \sqrt{2(17 + \sqrt{17})}\} .$$

These values are needed for η_1'', η_2'' where the choice of signs is again quite free. We set

$$\eta_1'' = \tfrac{1}{2}(\eta_1' + \sqrt{\eta_1'^2 - 4\eta_4'})$$

and obtain

$$\eta_1'' = \tfrac{1}{2}\{\tfrac{1}{2}(-1 + \sqrt{17} + \sqrt{2(17 - \sqrt{17})})$$
$$+ \sqrt{\tfrac{1}{16}(-1 + \sqrt{17} + \sqrt{2(17 - \sqrt{17})})^2 + 1 + \sqrt{17} - \sqrt{2(17 + \sqrt{17})}}\}$$
$$= \tfrac{1}{8}\{-1 + \sqrt{17} + \sqrt{2(17 - \sqrt{17})} + \sqrt{R}\} ,$$

where

$$R = (-1 + \sqrt{17} + \sqrt{2(17 - \sqrt{17})})^2 + 16 + 16\sqrt{17} - 16\sqrt{2(17 + \sqrt{17})}$$
$$= 4 \cdot 17 + 12\sqrt{17} + 2(-1 + \sqrt{17})\sqrt{2(17 - \sqrt{17})} - 16\sqrt{2(17 + \sqrt{17})} .$$

But we have

$$(1 + \sqrt{17})\sqrt{2(17 - \sqrt{17})} = \sqrt{2\sqrt{17}(\sqrt{17} - 1)(1 + \sqrt{17})^2}$$
$$= \sqrt{2\sqrt{17} \cdot 16(1 + \sqrt{17})}$$
$$= 4\sqrt{2(17 + \sqrt{17})} .$$

This yields

$$R = 4 \cdot 17 + 12\sqrt{17} - 4\sqrt{2(17 - \sqrt{17})} - 8\sqrt{2(17 + \sqrt{17})} ,$$

and thus

$$\eta_1'' = \tfrac{1}{8}\{-1 + \sqrt{17} + \sqrt{2(17 - \sqrt{17})}$$
$$+ 2\sqrt{17 + 3\sqrt{17} - \sqrt{2(17 - \sqrt{17})} - 2\sqrt{2(17 + \sqrt{17})}}\} ,$$

a quantity which is *real*. We could obtain z and z^{16} from (7.10) but are satisfied with

$$\eta_1'' = z + z^{16} = z + z^{-1} = 2\cos\theta ,$$
$$\cos\theta = \tfrac{1}{2}\eta_1'' ,$$

for the central angle of the regular 17-gon.

The surprising feature of the expression for η_1'' is that, in addition to the operations of addition, subtraction, and multiplication, it contains only one further operation, the extraction of square roots starting from rational integers. The unit length being given as the radius of the circumcircle of the 17-gon, the expression for η_1'' can therefore be constructed by means of ruler and compass as elementary theorems of geometry show. It is possible therefore to construct the regular 17-gon by means of ruler and compass alone. This is Gauss's discovery published in his *Disquisitions Arithmeticae* (1801), article 365. With justifiable pride, Gauss there points out that in Euclid's time the division of the circle in 3 and 5 equal parts was known and therefore the constructibility of the regular polygons of 2^μ, $3 \cdot 2^\mu$, $5 \cdot 2^\mu$, $15 \cdot 2^\mu$ sides,† but that for 2000 years nothing had been added to this knowledge.

It is fairly clear that Gauss's construction depends on the fact that $16 = 17 - 1$ is a power of 2. For it is just this fact that allows us to halve the periods again and again, and so to reduce the solution of the cyclotomic equation of degree 16 to the solution of a sequence of quadratic equations. If p is a prime of the form $2^k + 1$, then Gauss's method may be used in this manner to construct the regular p-gon with ruler and compass. If k contains an odd factor u, then $2^k + 1$ cannot be prime. For if $k = ul$, then with $2^l = A$ we have

$$2^k + 1 = 2^{ul} + 1 = A^u + 1 = (A + 1)(A^{u-1} - A^{u-2} + \cdots + 1),$$

and $2^k + 1$ is not a prime. Thus k must be a power of 2, say $k = 2^n$, and we must look for primes of the form $2^{2^n} + 1$. Let us see some examples:

$$\begin{aligned} 3 &= 2^1 + 1 \\ 5 &= 2^2 + 1 \\ 17 &= 2^4 + 1 \\ 257 &= 2^8 + 1 \\ 65537 &= 2^{16} + 1 . \end{aligned}$$

All these numbers are primes, called Fermat primes, and the corresponding regular polygons may be constructed with ruler and compass. Fermat studied the numbers $2^{2^n} + 1$ in a different connection and conjectured that they were all primes. But Euler showed that $2^{32} + 1$ is divisible by 641.

If we try to apply Gauss's method to the solution of the cyclotomic equation for $p = 7$,

$$z^6 + z^5 + \cdots + z + 1 = 0 ,$$

we may construct three periods of two terms or two periods of three terms, and we are led to a cubic equation which cannot be solved by rational operations and the extraction of square roots. But, using Cardano's solution of the cubic equation, we see that solutions of this equation may be expressed

† The latter because $\frac{1}{15} = \frac{1}{6} - \frac{1}{10}$.

in terms of square roots *and* cube roots. It is a remarkable fact that all cyclotomic equations

$$z^{p-1} + z^{p-2} + \cdots + z + 1 = 0 , \quad p \text{ prime},$$

can be solved by rational operations and successive extraction of roots. This is not at all the case for the general equation of the nth degree. In the sixteenth century Cardano constructed an explicit solution for the general cubic, and Ferrari an explicit solution for the general biquadratic equation, both solutions in terms of rational operations and successive extraction of roots. But Abel showed, early in the nineteenth century, that the general equation of degree n, where $n \geqq 5$, is not solvable in terms of radicals (roots).

Exercise: Solve the cyclotomic equation $z^6 + z^5 + \cdots + z + 1 = 0$.

8

Solution of Cyclotomic Equations

Primitive roots of unity. A solution of the algebraic equation

$$x^n - 1 = 0 \tag{8.1}$$

is called an nth *root of unity* or a *root of unity of order* n. The number 1 is a root of unity of any order, a trivial root we may say. Of importance will be those roots of (8.1) which are not also roots of $x^k - 1 = 0$ with $k < n$. We call such a root (the existence of which we shall have to show) a *primitive root of unity*. Let ζ be any root of (8.1). Then it will be the *primitive root* of some equation

$$x^k - 1 = 0, \qquad 0 < k \leqq n\,, \tag{8.2}$$

where k is chosen as the smallest positive integer for which $\zeta^k = 1$. Now put

$$d = (k, n)\,.$$

We can then find a and b so that

$$d = ka + nb\,.$$

Since, moreover,

$$d = k(a + tn) + n(b - tk)\,,$$

we can assume $a > 0$ without loss of generality. Then, for any positive c

$$d + nc = ka + n(b + c)\,.$$

We can take c so large that $b + c > 0$.

Then from (8.1) and (8.2) we obtain

$$\zeta^{ka} = 1, \qquad \zeta^{n(b+c)} = 1$$

so that

$$\zeta^{ka+n(b+c)} = \zeta^{d+nc} = \zeta^d = 1\,.$$

Therefore, since k was minimal, $k = d$, and k is a divisor of n. Any root of (8.1) is therefore a primitive root of unity of some order d, where $d \mid n$.

Let $\zeta_j, j = 1, 2, \cdots, v$, now be all the *primitive* roots of $x^n - 1 = 0$. We define the *cyclotomic polynomial* $F_n(x)$ of *order* n as

$$F_n(x) = \prod_j (x - \zeta_j)\,.$$

We see that $F_n(x)$ is a monic polynomial (i.e., one of highest coefficient 1). If no primitive roots of order n should exist, we might set $F_n(x) = 1$ but we shall see presently that this will not occur. We now have evidently

$$x^n - 1 = \prod_{d \mid n} F_d(x)\,. \tag{8.3}$$

For instance: $F_1(x) = x - 1$,

$$x^2 - 1 = F_1(x) \cdot F_2(x)\,;$$

thus

$$F_2(x) = x + 1\,.$$

THEOREM 29: *The cyclotomic polynomial $F_n(x)$ of order n is a monic polynomial of degree $\varphi(n)$ with integer coefficients.*

Proof: We employ induction. The theorem is true for $n = 1, 2$. Assume it to be true for all $F_k(x)$, $k < n$. Now

$$x^n - 1 = F_n(x) \cdot \prod_{\substack{d \mid n \\ d < n}} F_d(x) = F_n(x) \cdot G_n(x)\,, \tag{8.4}$$

say. But here, because of $d < n$, $G_n(x)$ is a product of monic polynomials with integer coefficients, hence it is also monic with integer coefficients. Then

$$F_n(x) = \frac{x^n - 1}{G_n(x)}\,.$$

Long division produces only integer coefficients here, because the divisor has highest coefficient 1. Now as to the degree of $F_n(x)$, if we assume the degree $\varphi(d)$ for $F_d(x)$, $d < n$, we have from (8.4), if v is the degree of $F_n(x)$:

$$n = v + \sum_{\substack{d \mid n \\ d < n}} \varphi(d) = v - \varphi(n) + \sum_{d \mid n} \varphi(d)\,.$$

Thus $v = \varphi(n)$, in view of (3.8), Theorem 11. (This proof is completely analogous to that of Theorem 27.)

Long division, used as a tool in this proof, provides through (8.4) a construction for consecutive $F_n(x)$. Besides the F_1 and F_2 already mentioned we find the following examples:

$F_3(x) = x^2 + x + 1$ $\qquad$ $F_8(x) = x^4 + 1$

$F_4(x) = x^2 + 1$ $\qquad$ $F_9(x) = x^6 + x^3 + 1$

$F_5(x) = x^4 + x^3 + x^2 + x + 1$ $\qquad$ $F_{10}(x) = x^4 - x^3 + x^2 - x + 1$

$F_6(x) = x^2 - x + 1$ $\qquad$ $F_{11}(x) = x^{10} + x^9 + \cdots + x + 1$

$F_7(x) = x^6 + x^5 + \cdots + x + 1$ $\qquad$ $F_{12}(x) = x^4 - x^2 + 1$.†

† In all these examples only the coefficients 1, −1, 0 appear. This is not so in all $F_n(x)$. Erdös (Bull. Amer. Math. Soc. **52** (1946), pp. 179–184) has proved that there exist cyclotomic polynomials which have some arbitrarily large coefficients.

A remark at the end of Chapter 7 can now be stated precisely as

THEOREM 30: *The cyclotomic equation can be solved by radicals.*

We break the proof down into several steps. Suppose first that $n = n_1 \cdot n_2$ is composite with $(n_1, n_2) = 1$. If ζ_1 is a primitive root of unity of order n_1 and ζ_2 is a primitive root of unity of order n_2, then we assert that $\zeta_1\zeta_2$ is a primitive root of unity of order n_1n_2. For let us assume that $(\zeta_1\zeta_2)^k = 1$. Then also

$$1 = (\zeta_1\zeta_2)^{kn_2} = \zeta_1^{kn_2} \cdot \zeta_2^{kn_2} = \zeta_1^{kn_2} .$$

We can solve the congruence $n_2a \equiv 1 \pmod{n_1}$ for a and have then

$$1 = \zeta_1^{kn_2a} = \zeta_1^k .$$

But ζ_1 is a primitive root of unity of order n_1, and therefore $n_1 \mid k$. In the same way we see that $n_2 \mid k$, and, since $(n_1, n_2) = 1$, also $n_1n_2 \mid k$. But this means indeed that $\zeta_1\zeta_2$ is a primitive root of unity of order n_1n_2.

Thus if $n = p_1^{\beta_1}p_2^{\beta_2} \cdots p_r^{\beta_r}$ is the decomposition of n into primes, we see that, in order to show that $F_n(x) = 0$ may be solved by consecutive extraction of roots, it suffices to show that all cyclotomic equations $F_{p^\beta}(x) = 0$ may be solved by radicals. The case $p = 2$ and its powers can be dealt with directly. We have on the one hand

$$x^{2^\beta} - 1 = (x^{2^{\beta-1}} - 1)(x^{2^{\beta-1}} + 1) .$$

On the other hand by (8.3)

$$x^{2^\beta} - 1 = \prod_{d \mid 2^\beta} F_d(x) = \prod_{\alpha=0}^{\beta} F_{2^\alpha}(x) = F_{2^\beta}(x) \cdot \prod_{\alpha=0}^{\beta-1} F_{2^\alpha}(x)$$
$$= F_{2^\beta}(x) \cdot (x^{2^{\beta-1}} - 1)$$

and thus

$$F_{2^\beta}(x) = x^{2^{\beta-1}} + 1 .$$

But $x^{2^{\beta-1}} + 1 = 0$ is already an equation of the binomial form. We have

$$-1, \qquad \sqrt{-1} = \pm i, \qquad \sqrt{i} = \pm\frac{1+i}{\sqrt{2}}$$

as primitive roots of unity of orders 2, 4, and 8, respectively, and can continue this list by successive extraction of square roots. In virtue of the formula

$$\sqrt{A + Bi} = \sqrt{\frac{A + \sqrt{A^2 + B^2}}{2}} + i\sqrt{\frac{-A + \sqrt{A^2 + B^2}}{2}},$$

we can even express a primitive root of order 2^β in the form $R_1 + iR_2$, where R_1 and R_2 are real and contain only repeated square roots of positive radicands, for $\sqrt{A^2 + B^2} \geqq |A|$.

So let p be odd. The important case is $F_p(x) = 0$. We make a remark about $n = p^\beta$ at the end of this chapter.

The Lagrange resolvent. Since

$$x^p - 1 = F_1(x) \cdot F_p(x) = (x - 1)F_p(x) ,$$

we have for any prime p

$$F_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1 .$$

Let ζ be a primitive root of unity of order p. Then $F_p(\zeta) = 0$ or

$$\zeta + \zeta^2 + \cdots + \zeta^r = -1$$

where here and subsequently we write

$$p - 1 = r .$$

We now take a primitive congruence root g modulo p. Then $g^r \equiv 1 \pmod{p}$ and no lower power of g can be congruent to 1 modulo p. The $r = p - 1$ numbers

$$g^0, g^1, \cdots, g^{r-1}$$

are congruent modulo p to some permutation of the numbers $1, 2, \cdots, r$, and since an exponent of ζ counts only modulo p, we may replace the preceding equation by

$$\zeta^{g^0} + \zeta^{g^1} + \zeta^{g^2} + \cdots + \zeta^{g^{r-1}} = -1 . \tag{8.5}$$

In addition to the sum on the left side of this equation, we introduce, following Lagrange, some other linear combinations of the powers of ζ. Let ρ be an rth root of unity. (Note that ρ is a root of unity of lower order than ζ.) The root ρ does not have to be primitive. We consider now the sum

$$(\rho, \zeta) = \zeta + \rho\zeta^g + \rho^2\zeta^{g^2} + \cdots + \rho^{r-1}\zeta^{g^{r-1}} , \tag{8.6}$$

called a *Lagrange resolvent*. In this notation we may write (8.5) simply as

$$(1, \zeta) = -1 . \tag{8.7}$$

If we knew all (ρ, ζ) for all ρ, as we know $(1, \zeta)$, then we would also know ζ expressed in terms of ρ. Indeed we have

$$\sum_\rho (\rho, \zeta) = \sum_\rho \zeta + \sum_\rho \rho\zeta^g + \sum_\rho \rho^2\zeta^{g^2} + \cdots + \sum_\rho \rho^{r-1}\zeta^{g^{r-1}} . \tag{8.8}$$

In order to evaluate the individual sums on the right side, we take ρ_0, a *primitive* root of unity of order r. Then all the ρ's form the set

$$\rho_0^0, \rho_0^1, \cdots, \rho_0^{r-1}$$

and thus

$$\sum_\rho \rho^k = \sum_{l=0}^{r-1} \rho_0^{lk} = \begin{cases} r \text{ if } r \mid k \\ \dfrac{\rho_0^{rk} - 1}{\rho_0^k - 1} = 0 \text{ if } r \nmid k . \end{cases}$$

We obtain therefore from (8.8)

$$\zeta = \frac{1}{r}\sum_{\rho}(\rho, \zeta)\,.$$

Our problem will be solved if we can obtain the Lagrange resolvents (ρ, ζ) by rational operations and repeated extraction of roots. We need the the following lemma.

LEMMA: $(\rho, \zeta) = \rho(\rho, \zeta^g)$.

To prove this we only have to realize that we can write $\zeta = \rho^r\zeta^{g^r}$, and thus, by a cyclic shift

$$\begin{aligned}(\rho, \zeta) &= \rho\zeta^g + \rho^2\zeta^{g^2} + \cdots + \rho^{r-1}\zeta^{g^{r-1}} + \rho^r\zeta^{g^r}\\ &= \rho\{\zeta^g + \rho\zeta^{g^2} + \cdots + \rho^{r-1}\zeta^{g^r}\}\,,\end{aligned}$$

which yields the assertion of the lemma. Repeated application of the lemma also shows

$$(\rho, \zeta) = \rho^h(\rho, \zeta^{g^h})\,. \tag{8.9}$$

In order to find an equation for (ρ, ζ), we need the product of two Lagrange resolvents:

$$\begin{aligned}(\rho^k, \zeta) &= \zeta + \rho^k\zeta^g + \rho^{2k}\zeta^{g^2} + \cdots + \rho^{(r-1)k}\zeta^{g^{r-1}}\\ (\rho^l, \zeta) &= \zeta + \rho^l\zeta^g + \rho^{2l}\zeta^{g^2} + \cdots + \rho^{(r-1)l}\zeta^{g^{r-1}}\,.\end{aligned}$$

We carry out the multiplication in the manner in which we multiplied two "periods" occurring in the theory of the 17-gon. We start by multiplying terms in the same column and add their products (which will yield here (ρ^{k+l}, ζ^2)); then after shifting the lower row cyclically, we again multiply terms above each other, and so on. It is better to express this process more formally. We have

$$\begin{aligned}(\rho^k, \zeta) &= \sum_{h=0}^{r-1}\rho^{kh}\zeta^{g^h}\,,\\ (\rho^l, \zeta) &= \sum_{j=0}^{r-1}\rho^{lj}\zeta^{g^j} = \sum_{j=0}^{r-1}\rho^{l(j+h)}\zeta^{g^{(j+h)}}\end{aligned}$$

for any h, in view of (8.9). Thus

$$\begin{aligned}(\rho^k, \zeta)\cdot(\rho^l, \zeta) &= \sum_{h=0}^{r-1}\sum_{j=0}^{r-1}\rho^{kh}\zeta^{g^h}\rho^{l(j+h)}\zeta^{g^{j+h}}\\ &= \sum_{j=0}^{r-1}\rho^{lj}\sum_{h=0}^{r-1}\rho^{(k+l)h}\zeta^{g^h(1+g^j)}\\ &= \sum_{j=0}^{r-1}\rho^{lj}(\rho^{k+l}, \zeta^{1+g^j})\,.\end{aligned} \tag{8.10}$$

Now as j runs from 0 to $r-1$, g^j takes on the values $1, 2, \cdots, p-1$ modulo p in some order, and thus $1 + g^j$ takes on the values

$$2, 3, \cdots, p-1, 0 \text{ modulo } p$$

in some order. Here

$$1 + g^j \equiv 0 \pmod p$$

occurs for $j = r/2$, because

$$0 \equiv g^r - 1 \equiv (g^{r/2} - 1)(g^{r/2} + 1) \pmod p\,,$$

and, therefore,

$$g^{r/2} + 1 \equiv 0 \pmod p\,,$$

since g is a primitive congruence root modulo p. For $j \neq (r/2)$, therefore, $1 + g^j$ runs through $2, 3, \cdots, p-1$. Now let $\lambda = \lambda(j)$ be that exponent† for which

$$g^{\lambda(j)} \equiv 1 + g^j \pmod p \qquad j = 1, 2, \cdots, r\,; \qquad j \neq \frac{r}{2}\,.$$

Incidentally, we see that $\lambda(j)$ runs in some order through the set $1, 2, \cdots, r-1$. If we now return to (8.10), we find that we have obtained

$$(\rho^k, \zeta)\cdot(\rho^l, \zeta) = \sum_{\substack{j=0\\ j\neq r/2}}^{r-1}\rho^{lj}\cdot(\rho^{k+l}, \zeta^{g^{\lambda(j)}}) + \rho^{(lr/2)}(\rho^{k+l}, \zeta^0)\,. \tag{8.11}$$

Although we have defined (ρ, ζ) in (8.6) only for a *primitive* root ζ, it is clear from the context that (ρ^{k+l}, ζ^0) here stands for

$$\begin{aligned}(\rho^{k+l}, 1) &= 1 + \rho^{k+l} + \rho^{2(k+l)} + \cdots + \rho^{(r-1)(k+l)}\\ &= \begin{cases}0 \text{ for } \rho^{k+l} \neq 1\\ r \text{ for } \rho^{k+l} = 1\,.\end{cases}\end{aligned} \tag{8.12}$$

If we replace ρ by ρ^{k+l} and h by $\lambda(j)$ in (8.9), we obtain

$$(\rho^{k+l}, \zeta) = \rho^{(k+l)\lambda(j)}(\rho^{k+l}, \zeta^{g^{\lambda(j)}})$$

or

$$(\rho^{k+l}, \zeta^{g^{\lambda(j)}}) = \rho^{-(k+l)\lambda(j)}(\rho^{k+l}, \zeta)\,.$$

With this and (8.12), equation (8.11) becomes

$$(\rho^k, \zeta)(\rho^l, \zeta) = (\rho^{k+l}, \zeta)\sum_{\substack{j=0\\ j\neq r/2}}^{r-1}\rho^{lj-(k+l)\lambda(j)} + R \tag{8.13}$$

where

$$R = \begin{cases}0 \text{ for } \rho^{k+l} \neq 1\\ r\rho^{(r/2)l} \text{ for } \rho^{k+l} = 1\,.\end{cases} \tag{8.14}$$

† If $g^\mu = m \pmod p$, then μ is called the "index" of m. In our case, therefore, $\lambda(j)$ is the index of $1 + g^j$.

(Since $\rho^r = 1$, we can have only $\rho^{r/2} = \pm 1$, and thus also $\rho^{(r/2)l} = \pm 1$.) We write for abbreviation

$$(\rho^k, \zeta) \cdot (\rho^l, \zeta) = (\rho^{k+l}, \zeta)\psi_{k,l}(\rho) + R\,, \tag{8.15}$$

where $\psi_{k,l}(\rho)$ is a polynomial in ρ of degree $< r$, which is obtained from

$$\sum_{\substack{j=0 \\ j \neq r/2}}^{r-1} \rho^{lj-(k+l)\lambda(j)} \tag{8.16}$$

through the application of $\rho^r = 1$. It follows that $\psi_{k,l}(\rho)$ has integer coefficients.

In the case $l = -k$ we can easily find the polynomial $\psi_{k,-k}(\rho)$ from (8.16). We have here, for a later application,

$$\psi_{k,-k}(\rho) = \sum_{\substack{j=0 \\ j \neq r/2}}^{r-1} \rho^{-kj} = \sum_{j=0}^{r-1} \rho^{-kj} - \rho^{-kr/2}$$
$$= \begin{cases} -\rho^{-kr/2} \text{ for } \rho^k \neq 1 \\ r-1 \text{ for } \rho^k = 1\,. \end{cases}$$

Therefore from (8.15), (8.14), and (8.7) after some simplification, we have

$$(\rho^k, \zeta)(\rho^{-k}, \zeta) = \begin{cases} p\rho^{(r/2)k} \text{ for } \rho^k \neq 1 \\ 1 \text{ for } \rho^k = 1\,. \end{cases} \tag{8.17}$$

(The statement for $\rho^k = 1$, namely $(1, \zeta)(1, \zeta) = 1$, is, of course, already implied by (8.7).) Formula (8.17) implies that $(\rho^k, \zeta) \neq 0$ for all k.

We can now use (8.15) for the computation of (ρ, ζ). Let d be the smallest natural number such that $\rho^d = 1$. Certainly $d \mid r$, and we can assume $d > 1$ since $(1, \zeta) = -1$ is known. We write (8.15) for $k = 1$ and successively for $l = 1, 2, \cdots, d-1$

$$(\rho, \zeta)(\rho, \zeta) = (\rho^2, \zeta)\psi_{1,1}(\rho)$$
$$(\rho, \zeta)(\rho^2, \zeta) = (\rho^3, \zeta)\psi_{1,2}(\rho)$$
$$\cdots\cdots\cdots\cdots$$
$$(\rho, \zeta)(\rho^{d-2}, \zeta) = (\rho^{d-1}, \zeta)\psi_{1,d-2}(\rho)$$
$$(\rho, \zeta)(\rho^{d-1}, \zeta) = (1, \zeta)\psi_{1,d-1}(\rho) + r\rho^{r/2} = -\psi_{1,d-1}(\rho) + r\rho^{r/2}\,.$$

Multiply these equations with each other and cancel a nonvanishing factor $(\rho^2, \zeta) \cdot (\rho^3, \zeta) \cdots (\rho^{d-1}, \zeta)$ on both sides. We then obtain a formula of the sort

$$(\rho, \zeta)^d = \Psi_d(\rho)$$

where $\Psi_d(\rho)$ is a polynomial in ρ with integer coefficients. Thus any Lagrange resolvent (ρ, ζ) can be obtained by root extraction from some polynomial in ρ.

The (ρ, ζ) then in turn furnish ζ, as we have seen. Since ρ is of lower order than ζ, we can assume by induction that the ρ's can be obtained in the same way, and we have then proved our theorem up to one gap, namely the discussion of $n = p^\alpha$, $\alpha > 1$. This, however, we reduce easily to the case $n = p$.

The cyclotomic equation for a prime power as index. We have, after (8.3),

$$x^{p^k} - 1 = \prod_{d \mid p^k} F_d(x) = \prod_{t=0}^{k} F_{p^t}(x)$$
$$= \prod_{t=0}^{k-1} F_{p^t}(x) \cdot F_{p^k}(x)$$
$$= (x^{p^{k-1}} - 1)F_{p^k}(x)\,, \tag{8.18}$$

so that

$$F_{p^k}(x) = \frac{y^p - 1}{y - 1} = F_p(y)\,,$$

where

$$y = x^{p^{k-1}}$$

is used as an abbreviation. The solution of

$$F_{p_k}(x) = 0$$

is therefore achieved in two steps.

(1) $$F_p(y) = 0\,,$$

can be solved by radicals, as we have proved; and

(2) $$x^{p^{k-1}} = y$$

can also be solved by radicals:

$$x = \sqrt[p^{k-1}]{y}\,,$$

which involves $k-1$ successive extractions of pth roots.

These observations now make the induction complete for all roots of unity of order less than p^k and thus finish the proof of Theorem 30.

9

Gaussian Sums as Special Lagrange Resolvents

Some applications of the Lagrange resolvents. The formulas of the previous chapter contain a wealth of arithmetical information. Take a prime number $p \equiv 1 \pmod 4$ so that $r = p - 1 \equiv 0 \pmod 4$. The numbers ± 1, $\pm i$ are 4th roots of unity and therefore also rth roots of unity. We then write the following Lagrange resolvents according to (8.6) and (8.7), where ζ is a primitive root of unity of order p:

$$\begin{aligned}(1, \zeta) &= \zeta + \zeta^g + \zeta^{g^2} + \cdots + \zeta^{g^{r-1}} = -1\\ (-1, \zeta) &= \zeta - \zeta^g + \zeta^{g^2} - \cdots - \zeta^{g^{r-1}}\\ (i, \zeta) &= \zeta + i\zeta^g - \zeta^{g^2} - \cdots - i\zeta^{g^{r-1}}\\ (-i, \zeta) &= \zeta - i\zeta^g - \zeta^{g^2} + \cdots + i\zeta^{g^{r-1}}\end{aligned}$$

Now (8.15) in conjunction with (8.14) shows that

$$\begin{aligned}(i, \zeta)(i, \zeta) &= (-1, \zeta)\psi_{1,1}(i)\\ (-i, \zeta)(-i, \zeta) &= (-1, \zeta)\psi_{1,1}(-i)\,,\end{aligned}$$

and from (8.17) we infer, since $r/2$ is even in our case,

$$(-1, \zeta)(-1, \zeta) = p\,.$$

Multiplication of the three last equations yields, after cancellation of $(-1, \zeta)^2$,

$$(i, \zeta)^2(-i, \zeta)^2 = \psi_{1,1}(i)\psi_{1,1}(-i)p\,.$$

Then (8.17) with $k = 1$, $\rho = i$, $\rho^{r/2} = (-1)^{r/4}$ shows that

$$(i, \zeta)(-i, \zeta) = (-1)^{r/4}p\,,$$

so that, in view of the previous formula,

$$p^2 = \psi_{1,1}(i)\psi_{1,1}(-i)p$$

or

$$p = \psi_{1,1}(i)\psi_{1,1}(-i)\,.$$

Now $\psi_{1,1}$ is a polynomial with integer coefficients, and therefore we may write $\psi_{1,1}(i) = a + bi$, where a and b are integers. Then, however, $\psi_{1,1}(-i) = a - bi$ and

$$p = (a + bi)(a - bi) = a^2 + b^2\,.$$

Thus we have proved the famous theorem of Fermat.

THEOREM 31: *If p is a prime congruent to 1 modulo 4, then p is the sum of 2 squares.*

We shall be able shortly to give another proof of this theorem, based on Theorem 18.

Exercise: The procedure in Chapter 8 makes it possible to compute $\psi_{1,1}$ for each special odd p. Determine $\psi_{1,1}$ for $p = 13$ and verify through it that $13 = 2^2 + 3^2$.

We can prove a theorem similar to the last one for primes $p \equiv 1$ modulo 3.

THEOREM 32: *If p is a prime congruent to 1 modulo 3, then $p = a^2 + 3b^2$ where a and b are integers.*

To show this, let ζ be a primitive pth root of unity, and ρ be a primitive cube root of unity; that is, a root of the polynomial $F_3(x) = x^2 + x + 1$. Then we have from (8.15), for $k = l = 1$

$$\begin{aligned}(\rho, \zeta) \cdot (\rho, \zeta) &= (\rho^2, \zeta)\psi_{1,1}(\rho)\\ (\rho^2, \zeta) \cdot (\rho^2, \zeta) &= (\rho^4, \zeta) \cdot \psi_{1,1}(\rho^2) = (\rho, \zeta)\psi_{1,1}(\rho^2)\,.\end{aligned} \tag{9.1}$$

On the other hand, $\rho^2 = \rho^{-1}$, and (8.17) shows that

$$(\rho, \zeta) \cdot (\rho^2, \zeta) = (\rho, \zeta) \cdot (\rho^{-1}, \zeta) = p\rho^{r/2} = p\,, \tag{9.2}$$

since $\dfrac{r}{2} = \dfrac{p-1}{2}$ is divisible by 3 under our assumption. Multiplying equations (9.1) with each other and canceling a factor $(\rho, \zeta) \cdot (\rho^2, \zeta)$, which, as (9.2) shows, does not vanish, we obtain by means of (9.2)

$$p = \psi_{1,1}(\rho) \cdot \psi_{1,1}(\rho^2)\,.$$

But $\psi_{1,1}(\rho)$, a polynomial with integer coefficients, can always be written as $A + B\rho$, since higher powers of ρ can be eliminated through $\rho^2 = -\rho - 1$. Hence $\psi_{1,1}(\rho^2) = A + B\rho^2$ and

$$p = (A + B\rho)(A + B\rho^2) = A^2 - AB + B^2\,.$$

To obtain the theorem in the form stated we note that A and B cannot both be even. If A is even and B odd, then the last formula can be rewritten as

$$p = \left(\frac{A}{2} - B\right)^2 + 3\left(\frac{A}{2}\right)^2,$$

while if A and B are both odd, then

$$p = \left(\frac{A + B}{2}\right)^2 + 3\left(\frac{A - B}{2}\right)^2,$$

in both cases showing integer squares.

Gaussian sums. Let p be any odd prime. In the theory of the regular 17-gon we took every second term of a sum of primitive roots of unity arranged by means of powers of a congruence root g:

$$(1, \zeta) = \zeta^{g^0} + \zeta^{g^1} + \zeta^{g^2} + \cdots + \zeta^{g^{r-1}} = -1 .$$

This was broken into the two sums

$$\eta_1 = \zeta^{g^0} + \zeta^{g^2} + \cdots + \zeta^{g^{r-2}}$$
$$\eta_2 = \zeta^{g^1} + \zeta^{g^3} + \cdots + \zeta^{g^{r-1}} ,$$

where as before $r = p - 1$, an even number. In the notation of the Lagrange resolvents, we have then

$$\begin{aligned} \eta_1 + \eta_2 &= (1, \zeta) \\ \eta_1 - \eta_2 &= (-1, \zeta) . \end{aligned} \tag{9.3}$$

Now η_1 only shows exponents which are squares:

$$1 = g^0, g^2, g^4, \cdots, g^{r-2} .$$

The summands in η_2 show exponents which are not squares and not congruent to squares modulo p. Indeed η_1 has all the summands which are congruent to a square modulo p (except 0). If we have a square t^2, $p \nmid t$, then there exists u so that

$$g^u \equiv t^2 \pmod{p} .$$

We can then find an exponent v so that

$$t \equiv g^v \pmod{p} ,$$
$$g^u \equiv g^{2v} \pmod{p} .$$

But this congruence implies, since g is a primitive congruence root modulo p

$$u \equiv 2v \pmod{r} ,$$

which shows that u is even since r is even.

The numbers that are congruent to a square modulo p are called *quadratic residues modulo p*, i.e., the numbers

$$1, g^2, g^4, \cdots, g^{r-2} \pmod{p} ;$$

the others, i.e.,

$$g, g^3, g^5, \cdots, g^{r-1} \pmod{p} ,$$

are called *quadratic nonresidues modulo p*. The number of quadratic residues is equal to the number of quadratic nonresidues, and that number is $\frac{r}{2} = \frac{p-1}{2}$.

In $(-1, \zeta)$ the powers with quadratic residues as exponents have a plus sign, those with quadratic nonresidues a minus sign. If we therefore introduce the *Legendre symbol* (m/p) with the following definition

$$\left(\frac{m}{p}\right) = \begin{cases} +1 \text{ if } m \text{ quadratic residue modulo } p \\ -1 \text{ if } m \text{ quadratic nonresidue modulo } p \end{cases} \tag{9.4}$$

with $p \nmid m$, we can then simply write

$$(-1, \zeta) = \sum_{m=1}^{p-1} \left(\frac{m}{p}\right) \zeta^m . \tag{9.5}$$

This special Lagrange resolvent is called a *Gaussian sum*, which we shall write simply as $G(\zeta)$. It is convenient to supplement the definition of the Legendre symbol (9.4) by the additional definition

$$\left(\frac{m}{p}\right) = 0 \text{ for } p \mid m . \tag{9.6}$$

We can then write

$$G(\zeta) = (-1, \zeta) = \sum_{m=0}^{p-1} \left(\frac{m}{p}\right) \zeta^m . \tag{9.7}$$

The Gaussian sum can also be written without the use of the Legendre symbol. Consider the squares modulo p

$$0, 1^2, 2^2, \cdots, (p-1)^2 .$$

Any number a which occurs in this array, that is, any quadratic residue, appears *twice*, since if

$$x^2 \equiv a \pmod{p} ,$$

then also $(p - x)^2 = p^2 - 2px + x^2 \equiv a \pmod{p}$. Therefore, we have

$$\zeta^0 + \sum_{m=1}^{p-1} \zeta^{m^2} = 1 + 2\eta_1 ,$$

and in view of (9.3),

$$\sum_{m=0}^{p-1} \zeta^{m^2} = 1 + (1, \zeta) + (-1, \zeta) = (-1, \zeta) ,$$

so that we have also

$$G(\zeta) = \sum_{m=0}^{p-1} \zeta^{m^2} . \tag{9.8}$$

The Legendre symbol. Because

even + even = even
odd + odd = even
odd + even = odd

we can read off from

$$g^u \cdot g^v = g^{u+v}$$

the facts

$$R \times R = R$$
$$R \times N = N$$
$$N \times N = R\,,$$

where R stands for quadratic residue, and N for quadratic nonresidue. The Legendre symbol permits the condensed expression:

$$\left(\frac{m}{p}\right) \cdot \left(\frac{n}{p}\right) = \left(\frac{mn}{p}\right). \tag{9.9}$$

From our definition it is clear that the upper number in the symbol represents only its congruence class, so that of course

$$\left(\frac{m_1}{p}\right) = \left(\frac{m_2}{p}\right) \text{ for } m_1 \equiv m_2 \pmod{p}. \tag{9.10}$$

We express the multiplicative property (9.9) of the Legendre symbol by saying that it is a "character of the residue group modulo p."† Since there are as many quadratic residues as nonresidues, *viz.* $(p-1)/2$, we also note the property

$$\sum_{m=0}^{p-1} \left(\frac{m}{p}\right) = 0\,. \tag{9.11}$$

Now if a is a quadratic residue, then the congruence

$$y^2 \equiv a \pmod{p}$$

is solvable. If we raise both sides to the power $(p-1)/2$, we obtain

$$a^{(p-1)/2} \equiv y^{p-1} \equiv 1 \pmod{p} \tag{9.12}$$

in view of Fermat's theorem. Formula (9.12) gives a necessary condition for a quadratic residue a modulo p. This criterion, found by Euler, is also sufficient, since the congruence

$$x^{(p-1)/2} \equiv 1 \pmod{p}\,, \tag{9.13}$$

which is satisfied by the $(p-1)/2$ quadratic residues, cannot have further solutions (Theorem 13). Now we have

$$0 \equiv x^{p-1} - 1 \equiv (x^{(p-1)/2} - 1)(x^{(p-1)/2} + 1) \pmod{p}\,.$$

† The multiplicative property of the Legendre symbol settles a question left open in the proof of Theorem 23. The two equations mentioned in the theorem imply $n \equiv x_1^2 \pmod 5$, and $2n \equiv x_2^2 \pmod 5$, respectively. The solvability of *both* these congruences can be expressed by $(n/5) = (2n/5) = 1$, which is impossible since $(2/5) = -1$.

Therefore those numbers b which do not satisfy (9.13), i.e., the quadratic nonresidues, must fulfill

$$b^{(p-1)/2} \equiv -1 \pmod{p}\,.$$

We now compare these statements with the definition of the Legendre symbol.

THEOREM 33: *The Legendre symbol satisfies the congruence*

$$\left(\frac{m}{p}\right) \equiv m^{(p-1)/2} \pmod{p}\,. \tag{9.14}$$

The statement (9.14) is evidently also fulfilled for the supplementary definition (9.6) of (m/p) in the case $p \mid m$.

From (9.14) follow immediately again the statements (9.9) and (9.10). In the case $m = p - 1 \equiv -1$, Theorem 33 gives rise to a corollary.

COROLLARY:

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}\,, \tag{9.15}$$

or explicitly:
-1 is a quadratic residue of the prime numbers $p \equiv 1$ (mod 4) and a quadratic nonresidue of the prime numbers $p \equiv 3$ (mod 4).

Therefore, if $p \equiv 1 \pmod 4$, the congruence

$$A^2 + 1 \equiv 0 \pmod{p}$$

is solvable. But the fact $p \mid (A^2 + 1)$ has the consequence $p = a^2 + b^2$ in view of Theorem 18. Thus we have arrived at a new proof of Fermat's Theorem 31.

The Corollary shows that those primes for which -1 is a quadratic residue lie in an arithmetic progression. We may ask: For which primes is a given number a quadratic residue? The surprising answer is that those primes always lie in certain arithmetic progressions. This fact is a consequence of Gauss's famous *law of quadratic reciprocity*.

THEOREM 34: *If p and q are (different) odd primes, then*

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{(p-1)/2 \,\cdot\, (q-1)/2}\,. \tag{9.16}$$

Thus $(p/q) = (q/p)$, unless both p and q are congruent to 3 modulo 4, in which case

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

It is clear that if p is fixed, then (q/p) has the same value for all $q' \equiv q \pmod{p}$. This is implied by the definition of the Legendre symbol. But Theorem 34 implies that, if the prime number q is fixed, then (q/p) has the same value for all $p' \equiv p \pmod{4q}$. For then

$$\left(\frac{p'}{q}\right) = \left(\frac{p}{q}\right),$$

since $p' \equiv p \pmod{4q}$ implies $p' \equiv p \pmod{q}$, and

$$\frac{p'-1}{2} \equiv \frac{p-1}{2} \pmod{2},$$

since $p' \equiv p \pmod{4q}$ implies $p' \equiv p \pmod{4}$,

$$(-1)^{(p'-1)/2 \cdot (q-1)/2} = (-1)^{(p-1)/2 \cdot (q-1)/2}.$$

We reserve the proof of Theorem 34 for the next chapter.

10

The Law of Quadratic Reciprocity

The Gaussian sums as periodic functions. The theorem mentioned in the title of this chapter is Theorem 34 of the previous chapter, which we shall now prove by means of the theory of Gaussian sums, drawing on our studies of cyclotomy.

Let p and q be different odd primes, and let ζ be a primitive pth root of unity. We have from (9.7) $G(\zeta) = (-1, \zeta)$, and from (8.17) for $\rho = -1$

$$G(\zeta)^2 = (-1)^{(p-1)/2} p. \tag{10.1}$$

It follows that

$$\begin{aligned} G(\zeta)^{q-1} = (G(\zeta)^2)^{(q-1)/2} &= (-1)^{(p-1)/2 \cdot (q-1)/2} p^{(q-1)/2} \\ &\equiv (-1)^{(p-1)/2 \cdot (q-1)/2} \left(\frac{p}{q}\right) \pmod{q} \end{aligned} \tag{10.2}$$

by Theorem 33. All we have to show now for the proof of Theorem 34 is that

$$\left(\frac{q}{p}\right) \equiv G(\zeta)^{q-1} \pmod{q}. \tag{10.3}$$

We now generalize the definition (9.7) by writing for any integer t

$$G(\zeta^t) = \sum_{m=0}^{p-1} \left(\frac{m}{p}\right) \zeta^{tm}. \tag{10.4}$$

This actually is new only for $t \equiv 0 \pmod{p}$, since ζ^t for the other values of t is a primitive pth root of unity together with ζ itself. For $t \equiv 0 \pmod{p}$ we have, however, by (9.11),

$$G(\zeta^0) = \sum_{m=0}^{p-1} \left(\frac{m}{p}\right) = 0. \dagger \tag{10.5}$$

With this definition we have now

$$G(\zeta^t) = \left(\frac{t}{p}\right) G(\zeta), \tag{10.6}$$

† Note that the use of (9.8) for a definition of $G(\zeta^t)$ would have given a different definition of $G(\zeta^0)$.

which is true for $t \equiv 0 \pmod p$ and is proved for other t as follows. For $t \not\equiv 0 \pmod p$ we choose t' such that

$$tt' \equiv 1 \pmod p .$$

Then we put

$$tm \equiv m' \pmod p ;$$

thus

$$m \equiv m't' .$$

We have from (10.4), since m' runs with m through a full residue system modulo p,

$$G(\zeta^t) \equiv \sum_{m' \bmod p} \left(\frac{m't'}{p}\right) \zeta^{m'} = \left(\frac{t'}{p}\right) \sum_{m'=0}^{p-1} \left(\frac{m'}{p}\right) \zeta^{m'}$$

or

$$G(\zeta^t) = \left(\frac{t}{p}\right) G(\zeta) ,$$

where we have observed that $(t'/p) = (t/p)$ because of $(tt'/p) = (1/p) = 1$. This proves (10.6).

Definition (10.4) as well as equation (10.6) shows that $G(\zeta^t)$ is a periodic function in t of period p.

Finite Fourier series. Such periodic arithmetic functions can now be expanded in a *finite Fourier series*, in complete analogy to Fourier series in analysis. Indeed the following theorem, which is of interest beyond our present purpose, is valid.

THEOREM 35: *Let $F(t)$ be a function defined for all integers t with the period m, and let η be a primitive mth root of unity.*

Then

$$F(t) = \sum_{u=0}^{m-1} a(u)\eta^{ut} \tag{10.7}$$

with

$$a(u) = \frac{1}{m} \sum_{t=0}^{m-1} F(t)\eta^{-tu} . \tag{10.8}$$

Proof: Formula (10.7) represents for $t = 0, 1, \cdots, m-1$ a system of m linear equations for the m unknowns $a(u)$. Assume now for a moment that it can be solved.† Then the $a(u)$ obtained must fulfill certain conditions. To

† That it can be solved is of course seen immediately, since its determinant is a nonvanishing Vandermonde determinant. However, we do not need this remark.

see this we multiply both sides of (10.7) by η^{-vt} and sum over t:

$$\sum_{t=0}^{m-1} F(t)\eta^{-vt} = \sum_{t=0}^{m-1} \left(\sum_{u=0}^{m-1} a(u)\eta^{ut}\right)\eta^{-vt}$$
$$= \sum_{u=0}^{m-1} a(u) \sum_{t=0}^{m-1} \eta^{(u-v)t} = m \cdot a(v) .$$

This shows that $a(u)$, if it exists as a solution, is unique and can have no other form than (10.8). But this $a(u)$ indeed satisfies (10.7) as can be seen by direct substitution:

$$\sum_{u=0}^{m-1} a(u)\eta^{ut} = \frac{1}{m} \sum_{u=0}^{m-1} \left(\sum_{s=0}^{m-1} F(s)\eta^{-us}\right)\eta^{ut}$$
$$= \frac{1}{m} \sum_{s=0}^{m-1} F(s) \sum_{u=0}^{m-1} \eta^{u(t-s)} = F(t) ,$$

which proves the theorem.

We may call the $a(u)$ the "Fourier coefficients" of the finite Fourier series (10.7).

NOTE: Because of the periodicity of $F(t)$ and $a(u)$ with the period m, the series in (10.7) and (10.8) need not be extended over the particular residue system $0, 1, 2, \cdots, m-1$, but may just as well be taken over any complete residue system modulo m, a remark which we shall presently put to use.

Exercise: The function $F(t) = \zeta^{g^t}$, where ζ is a pth root of unity and g a primitive congruence root modulo p, is periodic of period $r = p - 1$. Express the Lagrange resolvents in terms of Fourier coefficients of $F(t)$.

Proof of the quadratic reciprocity theorem. We now apply Theorem 35 to the Gaussian sums. Since $G(\zeta^t)$ is periodic modulo p, so is $G(\zeta^t)^k$ for any positive integer k. The modulus m of the theorem is to be replaced by p, and the root of unity η by the pth root of unity ζ. We then have

$$G(\zeta^t)^k = \sum_{u \bmod p} a_k(u)\zeta^{ut} \tag{10.9}$$

with

$$a_k(u) = \frac{1}{p} \sum_{v \bmod p} G(\zeta^v)^k \zeta^{-vu}$$
$$= \frac{1}{p} \sum_{v \bmod p} \sum_{m_1 \bmod p} \left(\frac{m_1}{p}\right) \zeta^{vm_1} \cdots \sum_{m_k \bmod p} \left(\frac{m_k}{p}\right) \zeta^{vm_k} \zeta^{-vu} \quad \text{by (10.4)}$$
$$= \frac{1}{p} \sum_{v \bmod p} \sum_{\substack{m_1, m_2, \cdots, m_k \\ \bmod p}} \left(\frac{m_1 m_2 \cdots m_k}{p}\right) \zeta^{v(m_1 + m_2 + \cdots + m_k - u)}$$
$$= \frac{1}{p} \sum_{\substack{m_1, m_2, \cdots, m_k \\ v \bmod p}} \left(\frac{m_1 m_2 \cdots m_k}{p}\right) \sum_{v \bmod p} \zeta^{v(m_1 + m_2 + \cdots + m_k - u)} ; \tag{10.10}$$

thus

$$a_k(u) = \sum_{\substack{m_j \bmod p \\ m_1+m_2+\cdots+m_k \equiv u \pmod p}} \left(\frac{m_1 m_2 \cdots m_k}{p}\right). \tag{10.11}$$

But $a_k(u)$ also has another determination which we express in the following lemma.

LEMMA: *For odd k the Fourier coefficients $a_k(u)$ defined in* (10.10) *have the property*

$$a_k(u) = \left(\frac{u}{p}\right) a_k(1). \tag{10.12}$$

Proof: For odd k we have from (10.6) that

$$G(\zeta^v)^k = \left(\frac{v}{p}\right) G(\zeta)^k.$$

Hence

$$\begin{aligned} a_k(u) &= \frac{1}{p} G(\zeta)^k \sum_{v \bmod p} \left(\frac{v}{p}\right) \zeta^{-vu} \qquad \text{by (10.8)} \\ &= \frac{1}{p} G(\zeta)^k G(\zeta^{-u}) \qquad \text{by (10.4)} \\ &= \frac{1}{p} G(\zeta)^k \left(\frac{u}{p}\right) G(\zeta^{-1}), \end{aligned}$$

where (10.6) is again applied, this time with ζ^{-1} instead of ζ. Comparing the last formula with its own special case $u = 1$, we obtain (10.12).

We now insert (10.12) in (10.9), with the result (*only for k odd*)

$$G(\zeta^t)^k = a_k(1) \sum_{u \bmod p} \left(\frac{u}{p}\right) \zeta^{ut} = a_k(1) G(\zeta^t).$$

For $t = 1$ we know $G(\zeta) \neq 0$ (see (10.1)) and therefore have

$$G(\zeta)^{k-1} = a_k(1).$$

In view of (10.3) we now have to study the case $k = q$ for an odd prime $q \neq p$. We need $G(\zeta)^{q-1}$ only modulo q. Since it turns out that $a_q(q)$ is easier to treat than $a_q(1)$, we again apply (10.12) and obtain

$$G(\zeta)^{q-1} = a_q(1) = \left(\frac{q}{p}\right) a_q(q);$$

then from (10.11)

$$G(\zeta)^{q-1} = \left(\frac{q}{p}\right) \sum_{m_j \bmod p} \left(\frac{m_1 m_2 \cdots m_q}{p}\right), \tag{10.13}$$

where the m_j are restricted to $m_1 + m_2 + \cdots + m_q \equiv q \pmod p$. This sum has to be evaluated only modulo q. Among the admitted values of the summation variables, let us first consider the possibility

$$m_1 \equiv m_2 \equiv \cdots \equiv m_q \pmod p.$$

This would require $m_1 + m_2 + \cdots + m_q \equiv q m_j \equiv q \pmod p$ and therefore by Theorem 13 $m_j \equiv 1 \pmod p$, yielding only the one summand

$$\left(\frac{1}{p}\right) = 1. \tag{10.14}$$

All other solutions of $m_1 + m_2 + \cdots + m_q \equiv q \pmod p$ must contain some incongruent elements. Then a cyclic permutation of $m_1, m_2, \cdots, m_q$ will give a new solution. Indeed, a cyclic permutation of

$$m_1, m_2, \cdots, m_q$$

can be expressed as

$$m_{1+s}, m_{2+s}, \cdots, m_{q+s}$$

for a certain s, $1 \leq s < q$, the subscripts of the m's being taken modulo q. If this set were the same as the previous one, we would have

$$m_j \equiv m_{j+s} \pmod p.$$

Therefore, by letting $j = s, 2s, \cdots, (q-1)s$, successively,

$$m_s \equiv m_{2s} \equiv \cdots \equiv m_{qs} \pmod p,$$

where the subscripts form a complete residue system modulo q. This possibility we have however already treated separately. Therefore, those solutions of $m_1 + m_2 + \cdots + m_q \equiv q \pmod p$ in which some incongruent elements appear produce q times the same summand

$$\left(\frac{m_1 m_2 \cdots m_q}{p}\right)$$

and thus form a sum $\equiv 0 \pmod q$. Consequently, modulo q only the single summand (10.14) counts, and we have from (10.13)

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

which finally settles (10.3). From (10.2) we then have

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

But this congruence must actually be an equality, since both sides are ± 1, and thus must either be equal or differ by 2, whereas the congruence shows that they could differ only by a multiple of the odd prime number q. In other words, we have proved (9.16), the reciprocity theorem.

A supplementary theorem. So far we have mentioned only odd primes. Now the prime 2 cannot appear in the "denominator" of a Legendre symbol, but it makes sense to ask for the value $(2/p)$. It is plausible that here also the prime numbers p for which 2 is a quadratic residue will lie in a certain arithmetic progression. By means of Euler's criterion, we first prepare a small list of $(2/p)$:

$$\left(\tfrac{2}{7}\right) = \left(\tfrac{2}{17}\right) = \left(\tfrac{2}{23}\right) = \left(\tfrac{2}{31}\right) = \cdots = +1$$

$$\left(\tfrac{2}{3}\right) = \left(\tfrac{2}{5}\right) = \left(\tfrac{2}{11}\right) = \left(\tfrac{2}{13}\right) = \left(\tfrac{2}{19}\right) = \cdots = -1 .$$

It seems that $(2/p) = +1$ for $p \equiv \pm 1 \pmod 8$ and $(2/p) = -1$ for $p \equiv \pm 3 \pmod 8$. Indeed we shall prove this conjecture in the following concise form.

THEOREM 36:

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

The *proof* will run very much like the proof for Theorem 34, although, of course, there is no question of reciprocity. We first define an arithmetic function

$$\chi(n) = \begin{cases} (-1)^{(n^2-1)/8} & n \text{ odd} \\ 0 & n \text{ even} \end{cases} . \tag{10.15}$$

This function is a character of the congruence group modulo 8:

(1) $\chi(n_1) = \chi(n_2)$ for $n_1 \equiv n_2 \pmod 8$, from definition (10.15).

(2) $\chi(n) \cdot \chi(m) = \chi(nm)$.

This is certainly true if $m \cdot n$ is even. But if m, n are both odd, we have

$$\frac{n^2-1}{8} + \frac{m^2-1}{8} - \frac{(mn)^2-1}{8} = -\frac{(n^2-1)(m^2-1)}{8} \equiv 0 \pmod 2$$

which implies (2) in this case. Moreover, we have

(3) $\sum_{n=1}^{8} \chi(n) = 0 .$

We now define a sort of Gaussian sum. Let ζ be a primitive 8th root of unity. We then set

$$\begin{aligned} H(\zeta) &= \sum_{n=1}^{8} \chi(n)\zeta^n \qquad (10.16) \\ &= \zeta - \zeta^3 - \zeta^5 + \zeta^7 = 2\zeta + 2\zeta^7 \\ &= 2(\zeta + \bar{\zeta}) = \pm 2\sqrt{2} , \end{aligned}$$

as can be seen from the values of the primitive 8th root of unity given on page 62. Thus

$$H(\zeta)^2 = 8 . \tag{10.17}$$

For an odd prime number q we therefore obtain

$$H(\zeta)^{q-1} = (H(\zeta)^2)^{(q-1)/2} = 8^{(q-1)/2} = 2^{q-1} \cdot 2^{(q-1)/2} \equiv \left(\frac{2}{q}\right) \pmod q \tag{10.18}$$

in view of Fermat's theorem and Euler's criterion for quadratic residues. We define now, generalizing (10.16),

$$H(\zeta^t) = \sum_{n=1}^{8} \chi(n)\zeta^{tn} .$$

Here $H(\zeta^t)$, and therefore also $(H(\zeta^t))^k$, is a periodic function of period 8. It possesses a finite Fourier expansion

$$H(\zeta^t)^k = \sum_{u=1}^{8} b_k(u)\zeta^{tu} \tag{10.19}$$

with

$$b_k(u) = \tfrac{1}{8} \sum_{t=1}^{8} H(\zeta^t)^k \zeta^{-tu} .$$

In analogy with the lemma on page 78 we have here the following lemma.

LEMMA: *For k odd*

$$b_k(u) = \chi(u) b_k(1) . \tag{10.20}$$

Indeed:

$$\begin{aligned} b_k(u) &= \tfrac{1}{8} \sum_{t=1}^{8} \left(\sum_{n=1}^{8} \chi(n)\zeta^{tn} \right)^k \zeta^{-tu} \\ &= \tfrac{1}{8} \sum_{t=1}^{8} \left(\sum_{n_1=1}^{8} \chi(n_1)\zeta^{tn_1} \cdots \sum_{n_k=1}^{8} \chi(n_k)\zeta^{tn_k} \right) \zeta^{-tu} \\ &= \tfrac{1}{8} \sum_{n_1,\ldots,n_k \bmod 8} \chi(n_1 n_2 \cdots n_k) \sum_{t=1}^{8} \zeta^{(n_1+n_2+\cdots+n_k-u)t} \\ &= \sum_{\substack{n_j \bmod 8 \\ n_1+n_2+\cdots+n_k \equiv u \pmod 8}} \chi(n_1 n_2 \cdots n_k) . \end{aligned}$$

First, for u odd we determine u', so that

$$uu' \equiv 1 \pmod 8 .$$

Putting then $u' \cdot n_j = n_j'$, we conclude

$$b_k(u) = \chi(u)^k \sum_{\substack{n_j' \bmod 8 \\ n_1'+\cdots+n_k' \equiv 1 \pmod 8}} \chi(n_1' n_2' \cdots n_k') = \chi(u) b_k(1) .$$

Secondly, for u even, the requirement

$$n_1 + n_2 + \cdots + n_k \equiv u \pmod 8$$

demands, since k is odd, that at least one of the n_j be even, which makes $\chi(n_1 n_2 \cdots) = 0$. This proves the lemma for u even.

The use of the lemma in (10.19) produces

$$H(\zeta^t)^k = b_k(1)\sum_{u=1}^{8}\chi(u)\zeta^{tu} = b_k(1)H(\zeta^t)$$

and, since after (10.17) $H(\zeta) \neq 0$,

$$H(\zeta)^{k-1} = b_k(1)\,.$$

In particular, for $k = q$ we obtain

$$H(\zeta)^{q-1} = b_q(1) = \chi(q)b_q(q)\,,$$

the latter from (10.20). The sum

$$b_q(q) = \sum_{\substack{n_j \bmod 8 \\ n_1+n_2+\cdots+n_q \equiv q \pmod 8}} \chi(n_1 n_2 \cdots n_q)$$

is now treated exactly as $a_q(q)$ following formula (10.13).

We also have here

$$b_q(q) \equiv \chi(1) \equiv 1 \pmod q\,,$$

and thus

$$H(\zeta)^{q-1} \equiv \chi(q) \pmod q\,.$$

This together with (10.18) shows

$$\chi(q) \equiv \left(\frac{2}{q}\right) \pmod q\,,$$

and therefore

$$\chi(q) = \left(\frac{2}{q}\right)\,,$$

which is Theorem 36, in view of the definition (10.15).

The reciprocity theorem and the "supplementary theorem" about $(2/p)$ can be used to compute the Legendre symbol for large primes.

EXAMPLE:

$$\begin{aligned}\left(\tfrac{12703}{18361}\right) &= \left(\tfrac{18361}{12703}\right) = \left(\tfrac{5658}{12703}\right) = \left(\tfrac{2}{12703}\right)\left(\tfrac{1829}{12703}\right)\\ &= \left(\tfrac{31}{12703}\right)\left(\tfrac{59}{12703}\right) = -\left(\tfrac{12703}{31}\right)\cdot -\left(\tfrac{12703}{59}\right)\\ &= \left(\tfrac{24}{31}\right)\left(\tfrac{18}{59}\right) = \left(\tfrac{2}{31}\right)\left(\tfrac{3}{31}\right)\left(\tfrac{2}{59}\right) = -\left(\tfrac{31}{3}\right)\cdot(-1) = \left(\tfrac{1}{3}\right) = +1\,.\end{aligned}$$

11

The Product Formula for the Gaussian Sums

The problem of the sign of a Gaussian sum. The formula (10.1) of the previous chapter gives only the square of the Gaussian sum $G(\zeta)$ and yields merely

$$G(\zeta) = \pm i^{(p-1)/2}\sqrt{p}\,.$$

Now, as the formula stands, the ambiguity of the sign is unavoidable because we have

$$G(\zeta^t) = \left(\frac{t}{p}\right)G(\zeta) = \pm G(\zeta)\,.$$

But *algebraically*, for $t \not\equiv 0 \pmod p$, the roots of unity ζ and ζ^t are indistinguishable, both being primitive roots. However, if we use a *transcendental* characterization of ζ, namely

$$\zeta = e^{2\pi i h/p}, \quad p \nmid h\,,$$

then the definition (9.7) as well as the definition (9.8) give specific complex numbers for $G(\zeta)$ with no alternative in the $\pm$ signs. The problem of determining the sign of the Gaussian sums for transcendentally specified primitive roots has become famous since the time of Gauss, who devoted a beautiful paper to it ("Summatio quarumdam serierum singularium"). Since then a number of other quite different methods have been invented to deal with the sign of the Gaussian sums.

Whereas our previous discussion of Gaussian sums was all based on the definition (9.7), we turn now, following Gauss, to the definition (9.8). The main result will be a remarkable product expansion of the Gaussian sums. We shall at the same time generalize the Gaussian sums from a prime number order to any odd order of the primitive root ζ which appears in the definition. This generalization will also lead to a generalization of the Legendre symbol.

The Gaussian polynomials. Following Gauss we introduce the rational functions of an indeterminate x

$$\begin{bmatrix} n \\ m \end{bmatrix} = \frac{(1-x^n)(1-x^{n-1})\cdots(1-x^{n-m+1})}{(1-x)(1-x^2)\cdots(1-x^m)}\,. \tag{11.1}$$

Here n and m are positive integers. It will turn out that these expressions are actually polynomials in x, as we shall prove presently.

A slight algebraic manipulation gives a relation between certain pairs of these rational functions:

$$\begin{bmatrix} n \\ m \end{bmatrix} = \frac{(1-x^n)(1-x^{n-1})\cdots(1-x^{n-m+1})}{(1-x)(1-x^2)\cdots(1-x^m)} \cdot \frac{(1-x^{n-m})\cdots(1-x)}{(1-x)\cdots(1-x^{n-m})}$$
$$= \begin{bmatrix} n \\ n-m \end{bmatrix}.$$

(Here we observe that the numerator of the rational expression depends only on n, while the denominator is symmetrical in m and $n-m$, and thus the equality follows.)

This relation reminds us of the equality of binomial coefficients: $\binom{n}{m} = \binom{n}{n-m}$. In order to complete the analogy, we define

$$\begin{bmatrix} n \\ 0 \end{bmatrix} = 1, \text{ for } n = 0, 1, 2, \cdots$$

so that now $\begin{bmatrix} n \\ m \end{bmatrix} = \begin{bmatrix} n \\ n-m \end{bmatrix}$ for all $n = 0, 1, 2, \cdots$ and $m = 0, 1, \cdots, n$.

We also note that, in consequence of the original definition (11.1), $\begin{bmatrix} n \\ m \end{bmatrix} = 0$ for all $m > n$. (If we wish, we might also let $\begin{bmatrix} n \\ m \end{bmatrix} = 0$ for all $m < 0$.)

If m and n are positive we have the recursion formula

$$\begin{bmatrix} n \\ m \end{bmatrix} = \frac{1-x^n}{1-x^m}\begin{bmatrix} n-1 \\ m-1 \end{bmatrix}$$
$$= \left(1 + \frac{x^m - x^n}{1-x^m}\right)\begin{bmatrix} n-1 \\ m-1 \end{bmatrix}$$
$$= \begin{bmatrix} n-1 \\ m-1 \end{bmatrix} + x^m \begin{bmatrix} n-1 \\ m \end{bmatrix}.$$

Since

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1; \quad \begin{bmatrix} 1 \\ j \end{bmatrix} = 0, \qquad j \neq 0, 1,$$

this recursion formula shows inductively that the $\begin{bmatrix} n \\ m \end{bmatrix}$ are indeed polynomials, which we call the *Gaussian polynomials*. If we factor out $1-x$ in all the terms occuring in the numerator and denominator of $\begin{bmatrix} n \\ m \end{bmatrix}$, we find

$$\begin{bmatrix} n \\ m \end{bmatrix} =$$

$$\frac{(1+x+\cdots+x^{n-1})\,(1+x+\cdots+x^{n-2})\cdots(1+x+\cdots+x^{n-m})}{1 \qquad\qquad (1+x) \qquad\qquad \cdots\ (1+x+\cdots x^{m-1})}.$$

Setting $x = 1$ we see that the value of the polynomial $\begin{bmatrix} n \\ m \end{bmatrix}$ is the binomial coefficient

$$\frac{n(n-1)\cdots(n-m+1)}{1\cdot 2\cdots m} = \binom{n}{m}.$$

For $x = 1$ our recursion formula for $\begin{bmatrix} n \\ m \end{bmatrix}$ becomes a familiar formula for the binomial coefficients. Note also that we have a second recurrence relation:

$$\begin{bmatrix} n \\ m \end{bmatrix} = \begin{bmatrix} n \\ n-m \end{bmatrix} = \begin{bmatrix} n-1 \\ n-m-1 \end{bmatrix} + x^{n-m}\begin{bmatrix} n-1 \\ n-m \end{bmatrix}$$
$$= \begin{bmatrix} n-1 \\ m \end{bmatrix} + x^{n-m}\begin{bmatrix} n-1 \\ m-1 \end{bmatrix}.$$

Exercise: Prove the identity

$$(1+y)(1+xy)(1+x^2y)\cdots(1+x^{n-1}y)$$
$$= 1 + \begin{bmatrix} n \\ 1 \end{bmatrix}y + \begin{bmatrix} n \\ 2 \end{bmatrix}xy^2 + \begin{bmatrix} n \\ 3 \end{bmatrix}x^3y^3 + \cdots + \begin{bmatrix} n \\ \nu \end{bmatrix}x^{\nu(\nu-1)/2}y^\nu + \cdots.$$

Setting $x = 1$ gives the binomial expansion for $(1+y)^n$.

A sum of Gaussian polynomials. Now let us follow Gauss and form the sum

$$f(x, m) = 1 - \begin{bmatrix} m \\ 1 \end{bmatrix} + \begin{bmatrix} m \\ 2 \end{bmatrix} - \cdots + (-1)^m \begin{bmatrix} m \\ m \end{bmatrix}.$$

We use our second recurrence relation and have

$$1 = 1,$$
$$-\begin{bmatrix} m \\ 1 \end{bmatrix} = -\begin{bmatrix} m-1 \\ 1 \end{bmatrix} - x^{m-1},$$
$$\begin{bmatrix} m \\ 2 \end{bmatrix} = \begin{bmatrix} m-1 \\ 2 \end{bmatrix} + x^{m-2}\begin{bmatrix} m-1 \\ 1 \end{bmatrix},$$
$$-\begin{bmatrix} m \\ 3 \end{bmatrix} = -\begin{bmatrix} m-1 \\ 3 \end{bmatrix} - x^{m-3}\begin{bmatrix} m-1 \\ 2 \end{bmatrix},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

Summing these equations we find

$$f(x, m) = (1 - x^{m-1}) - (1 - x^{m-2})\begin{bmatrix} m-1 \\ 1 \end{bmatrix} + (1 - x^{m-3})\begin{bmatrix} m-1 \\ 2 \end{bmatrix} - \cdots.$$

But from definition (11.1)

$$(1 - x^{m-j})\begin{bmatrix} m-1 \\ j-1 \end{bmatrix} = (1 - x^{m-1})\begin{bmatrix} m-2 \\ j-1 \end{bmatrix}$$

and hence

$$f(x, m) = (1 - x^{m-1})\{1 - \begin{bmatrix} m-2 \\ 1 \end{bmatrix} + \begin{bmatrix} m-2 \\ 2 \end{bmatrix} - \cdots\}$$
$$= (1 - x^{m-1}) f(x, m-2) .$$

Now $f(x, 0) = 1$, so that we find recursively

$$\begin{aligned} f(x, 2) &= 1 - x \\ f(x, 4) &= (1 - x)(1 - x^3) \\ &\cdots\cdots\cdots\cdots \\ f(x, 2n) &= (1 - x)(1 - x^3)\cdots(1 - x^{2n-1}) . \end{aligned} \tag{11.2}$$

Since $f(x, 1) = 1 - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0$, it follows that $f(x, 2n-1) = 0$. This is also evident from the symmetry $\begin{bmatrix} 2n-1 \\ k \end{bmatrix} = \begin{bmatrix} 2n-1 \\ 2n-1-k \end{bmatrix}$.

Application to Gaussian sums. Now we apply these notions to the Gaussian sums. Let k be odd and let α be a primitive kth root of unity. By definition (11.1) we have

$$f(\alpha, k-1) = 1 - \frac{1-\alpha^{k-1}}{1-\alpha} + \frac{(1-\alpha^{k-1})(1-\alpha^{k-2})}{(1-\alpha)(1-\alpha^2)} - \cdots + \frac{(1-\alpha^{k-1})(1-\alpha^{k-2})\cdots(1-\alpha)}{(1-\alpha)(1-\alpha^2)\cdots(1-\alpha^{k-1})} .$$

Since

$$\frac{1-\alpha^{k-j}}{1-\alpha^j} = \alpha^{-j}\frac{\alpha^j - \alpha^k}{1-\alpha^j} = \alpha^{-j}\frac{\alpha^j - 1}{1-\alpha^j} = -\alpha^{-j} ,$$

we find that for $j = 1, 2, \cdots, k-1$

$$f(\alpha, k-1) = 1 + \alpha^{-1} + \alpha^{-1-2} + \cdots + \alpha^{-1-2-\cdots-(k-1)} = \sum_{j=0}^{k-1} \alpha^{-j(j+1)/2} .$$

Now k is odd and therefore α^{-2} is also a primitive kth root of unity. If we therefore replace α by α^{-2}, we obtain

$$f(\alpha^{-2}, k-1) = \sum_{j=0}^{k-1} \alpha^{j(j+1)} .$$

But $\alpha^k = 1$, and thus

$$\alpha^{j(j+1)} = \alpha^{[j+(k+1)/2]^2 - [(k+1)/2]^2} ;$$

hence

$$f(\alpha^{-2}, k-1) = \alpha^{-[(k+1)/2]^2} \sum_{j=0}^{k-1} \alpha^{[j+(k+1)/2]^2} .$$

Now if j runs from 0 to $k-1$, then $j + (k+1)/2$ runs through all residue classes modulo k. Since the exponents of α count only modulo k, we obtain

$$f(\alpha^{-2}, k-1) = \alpha^{-[(k+1)/2]^2} G(\alpha) , \tag{11.3}$$

where

$$G(\alpha) = \sum_{j \bmod k} \alpha^{j^2} \tag{11.4}$$

is a *Gaussian sum* for the modulus k, in generalization of (9.8) for a prime modulus. The application of (11.2) to (11.3) yields

$$G(\alpha) = \alpha^{[(k+1)/2]^2}(1 - \alpha^{-2})(1 - \alpha^{-4})\cdots(1 - \alpha^{-2(k-2)})$$
$$= \alpha^{[(k+1)/2]^2 - [(k-1)/2]^2}(\alpha - \alpha^{-1})(\alpha^3 - \alpha^{-3})\cdots(\alpha^{k-2} - \alpha^{-(k-2)}) ,$$

so that we have the remarkable product formula

$$G(\alpha) = (\alpha - \alpha^{-1})(\alpha^3 - \alpha^{-3})\cdots(\alpha^{k-2} - \alpha^{-(k-2)}) . \tag{11.5}$$

Here only the odd exponents appear. But we have

$$\begin{aligned} \alpha^{k-2} - \alpha^{-(k-2)} &= -(\alpha^2 - \alpha^{-2}) \\ \alpha^{k-4} - \alpha^{-(k-4)} &= -(\alpha^4 - \alpha^{-4}) \\ &\cdots\cdots\cdots\cdots \\ \alpha^1 - \alpha^{-1} &= -(\alpha^{k-1} - \alpha^{-(k-1)}) , \end{aligned}$$

so that we obtain

$$G(\alpha) = (-1)^{(k-1)/2}(\alpha^2 - \alpha^{-2})(\alpha^4 - \alpha^{-4})\cdots(\alpha^{k-1} - \alpha^{-(k-1)}) , \tag{11.6}$$

which shows only even exponents.

Formulae (11.5) and (11.6) together yield

$$G(\alpha)^2 = (-1)^{(k-1)/2}\prod_{j=1}^{k-1}(\alpha^j - \alpha^{-j})$$
$$= (-1)^{(k-1)/2}\alpha^{1+2+\cdots+(k-1)}\prod_{j=1}^{k-1}(1 - \alpha^{-2j}) .$$

But k is odd; therefore $-2j$ runs with j through the nonzero residue classes modulo k, so that

$$G(\alpha)^2 = (-1)^{(k-1)/2}\alpha^{k(k-1)/2}\prod_{j=1}^{k-1}(1 - \alpha^j) .$$

Now since α^j here represents all roots of $x^k - 1$ with the exception of 1, we see that

$$\prod_{j=1}^{k-1}(x - \alpha^j) = \frac{x^k - 1}{x - 1} = 1 + x + \cdots + x^{k-1} ,$$

which gives for $x = 1$

$$\prod_{j=1}^{k-1}(1 - \alpha^j) = k .$$

Thus we have finally

$$G(\alpha)^2 = (-1)^{(k-1)/2}k\,, \tag{11.7}$$

in agreement with our previous (10.1) where k was a prime p.

The Jacobi symbol. If $k = p$ is a prime and $(h, p) = 1$, we have seen in (10.6) that

$$\left(\frac{h}{p}\right) = \frac{G(\alpha^h)}{G(\alpha)}\,.$$

Thus for any odd k it is natural to consider the quotient

$$Q_h(\alpha) = \frac{G(\alpha^h)}{G(\alpha)}\,. \tag{11.8}$$

Here, as always, α is a primitive kth root, and we consider only the case $(h, k) = 1$. It is a remarkable fact that $Q_h(\alpha)$ is independent of α and will depend on, besides h, only the order k of the primitive root of unity α.

To prove this we make use of the product formula (11.6) for $G(\alpha)$ and $G(\alpha^h)$, which is permissible since α^h is also a primitive kth root of unity. We have

$$Q_h(\alpha) = \frac{\prod_{j=1}^{(k-1)/2} (\alpha^{2hj} - \alpha^{-2hj})}{\prod_{j=1}^{(k-1)/2} (\alpha^{2j} - \alpha^{-2j})}\,. \tag{11.9}$$

We shall see that the factors in the numerator and denominator are equal in pairs, except for a $\pm$ sign which depends only on h and k and not on α. To see this we put

$$hj \equiv r \pmod{k}, \qquad j = 1, 2, \cdots, \frac{k-1}{2}\,, \tag{11.10}$$

with the specification

$$-\frac{k}{2} < r < \frac{k}{2}\,. \tag{11.11}$$

The integer r is uniquely defined by j and clearly $r \neq 0$, since $(h, k) = 1$. We then have

$$\alpha^{2hj} - \alpha^{-2hj} = \alpha^{2r} - \alpha^{-2r} = (\alpha^{2s} - \alpha^{-2s}) \cdot \operatorname{sign} r\,, \tag{11.12}$$

where $s = |r|$. Now r takes $(k-1)/2$ different values; but we shall now show that s also takes as many different values. Indeed if $s' = s''$, which means $|r'| = |r''|$, then we would have $r' = \pm r''$. Hence

$$hj' \equiv \pm hj'' \pmod{k}$$

or

$$h(j' \pm j'') \equiv 0 \pmod{k}\,,$$

and consequently

$$j' \pm j'' \equiv 0 \pmod{k}\,,$$

which is, however, impossible since all j are different and

$$1 \leqq |j' \pm j''| \leqq k - 1\,.$$

Formulae (11.9) and (11.12) together then yield

$$Q_h(\alpha) = \prod_{j=1}^{(k-1)/2} \operatorname{sign} r\,, \tag{11.13}$$

and this is independent of α, since each r is found uniquely through (11.10) and (11.11) where only h and k appear. In order to emphasize this independence of α, we introduce a new symbol and write

$$\left(\frac{h}{k}\right) = Q_h(\alpha)\,. \tag{11.14}$$

This symbol, defined only for any odd number k and any positive or negative integer h prime to k, is called the *Jacobi symbol*. Definition (11.8) shows that the Legendre symbol is a special case of it. Equation (11.13) shows that (h/k) takes only the values ± 1.

The Jacobi symbol as a character of the multiplicative group modulo k. The definition (11.8) shows that

$$\left(\frac{h}{k}\right) = \left(\frac{h'}{k}\right) \text{ if } h \equiv h' \pmod{k}\,. \tag{11.15}$$

If h_1 and h_2 are both prime to k, we may use the fact that $Q_h(\alpha)$ is independent of α to conclude that

$$\frac{G(\alpha^{h_1 h_2})}{G(\alpha^{h_1})} = \frac{G(\alpha^{h_2})}{G(\alpha)}$$

because α^{h_1} is also a primitive kth root of unity. Hence

$$\frac{G(\alpha^{h_1 h_2})}{G(\alpha)} = \frac{G(\alpha^{h_1})G(\alpha^{h_2})}{G(\alpha)G(\alpha)}\,,$$

which we can write

$$\left(\frac{h_1 h_2}{k}\right) = \left(\frac{h_1}{k}\right)\left(\frac{h_2}{k}\right)\,. \tag{11.16}$$

This shows that the Jacobi symbol is a character of the multiplicative group of residues modulo k. The Jacobi symbol enjoys a reciprocity property just as the Legendre symbol does, as we shall show later.

We can use (11.13) to get an explicit expression for (h/k). In view of (11.10) we can write

$$hj = mk + r$$

and therefore

$$\frac{2hj}{k} = 2m + \frac{2r}{k}$$

with a certain integer m. Since, according to (11.11),

$$-1 < \frac{2r}{k} < 1\,,$$

and thus†

$$\left[\frac{2r}{k}\right] = 0 \text{ for } r > 0$$

$$\left[\frac{2r}{k}\right] = -1 \text{ for } r < 0\,,$$

we conclude

$$\left[\frac{2hj}{k}\right] \text{ is even for } r > 0$$

$$\left[\frac{2hj}{k}\right] \text{ is odd for } r < 0\,.$$

Therefore,

$$\text{sign } r = (-1)^{[2hj/k]}\,,$$

and from (11.13) we infer then the following theorem.

THEOREM 37: *The Jacobi symbol for coprime h, k, k being odd, is given by*

$$\left(\frac{h}{k}\right) = (-1)^{\sum_{j=1}^{(k-1)/2} [2hj/k]} \tag{11.17}$$

Let us compute some values of (h/k). Definition (11.14) together with (11.8) shows immediately

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

as can also be seen from (11.17).

Since, furthermore,

$$\left[\frac{-2j}{k}\right] = -1 \text{ for } j = 1, 2, \cdots, \frac{k-1}{2}\,,$$

† The symbol $[x]$ is explained on p. 31.

we immediately have from (11.17)

$$\left(\frac{-1}{k}\right) = (-1)^{(k-1)/2}\,.$$

For $h = 2$ we observe

$$\left[\frac{4j}{k}\right] = 0 \text{ for } 0 < j < \frac{k}{4}\,,$$

$$\left[\frac{4j}{k}\right] = 1 \text{ for } \frac{k}{4} < j < \frac{k}{2}\,,$$

so that

$$\sum_{j=1}^{(k-1)/2} \left[\frac{4j}{k}\right] = \left[\frac{k}{2}\right] - \left[\frac{k}{4}\right] = \frac{k-1}{2} - \left[\frac{k}{4}\right]\,,$$

and therefore

$$\left(\frac{2}{k}\right) = (-1)^{(k-1)/2 - [k/4]}\,.$$

This expression can be put into a more elegant form if we observe that here k matters obviously only modulo 8. We therefore prepare a list for the cases $k \equiv 1, 3, 5, 7$ modulo 8 and obtain

$$\left(\frac{2}{k}\right) = \begin{cases} 1 & k \equiv 1 \\ -1 & k \equiv 3 \\ -1 & k \equiv 5 \\ 1 & k \equiv 7 \end{cases} \text{ (modulo 8)}$$

or

$$\left(\frac{2}{k}\right) = \begin{cases} 1 \text{ for } k \equiv \pm 1 \pmod 8 \\ -1 \text{ for } k \equiv \pm 3 \pmod 8\,. \end{cases}$$

But we have already found an expression for this arithmetic function where $k = p$ is a prime. We thus have

$$\left(\frac{2}{k}\right) = (-1)^{(k^2-1)/8}\,. \tag{11.18}$$

The formulae (11.17) and (11.18) will furnish a proof of the reciprocity law for the Jacobi symbol by means of the concept of lattice points, which we shall give in the next chapter.

Right now we give a proof of the reciprocity law by means of the properties of Gaussian sums. For this we need the following lemma about Gaussian sums of different order.

LEMMA: *Let k and l be odd, $(k, l) = 1$. Let α be a primitive kth root of unity and β a primitive lth root of unity. Then $\alpha\beta$ is a primitive klth root of unity and*

$$G(\alpha\beta) = G(\alpha)G(\beta)\,. \tag{11.19}$$

Proof: That $\alpha\beta$ is a primitive klth root of unity has already been observed in the beginning of the proof of Theorem 30 of Chapter 8. We put now

$$j = lt + ku\,.$$

It is easily seen, since $(k, l) = 1$, that if t runs through a complete residue system modulo k and u through a complete residue system modulo l, then j runs through a complete residue system modulo kl. We need now only the definition of the Gaussian sums:

$$\begin{aligned} G(\alpha\beta) &= \sum_{j \bmod kl} (\alpha\beta)^{j^2} \\ &= \sum_{t \bmod k} \sum_{u \bmod l} (\alpha\beta)^{(lt+ku)^2} \\ &= \sum_{t \bmod k} (\alpha\beta)^{l^2t^2} \sum_{u \bmod l} (\alpha\beta)^{k^2u^2} \end{aligned}$$

But along with t, lt also runs through a full residue system modulo k, and analogously since $\beta^{l^2} = \alpha^{k^2} = 1$, for ku. Thus we obtain,

$$G(\alpha\beta) = \sum_{t \bmod k} \alpha^{t^2} \sum_{u \bmod l} \beta^{u^2}\,,$$

which is (11.19).

The sign of the Gaussian sums. Gauss used the product formula (11.6) for the determination of the sign of $G(\alpha)$, which the formula (11.7) necessarily leaves open. As we have seen, the problem of the sign of a Gaussian sum becomes meaningful only if we specify the kth root of unity by transcendental (nonalgebraic) means. For this purpose let us put

$$\alpha = e^{2\pi i h/k}, \qquad (h, k) = 1\,.$$

For the sake of brevity we now write

$$G(h, k) = G(e^{2\pi i h/k})\,. \tag{11.20}$$

Then, in particular for $h = 1$, formula (11.6) becomes

$$\begin{aligned} G(1, k) &= (-1)^{(k-1)/2} \prod_{j=1}^{(k-1)/2} (e^{4\pi ij/k} - e^{-4\pi ij/k}) \\ &= (-1)^{(k-1)/2} (2i)^{(k-1)/2} \prod_{j=1}^{(k-1)/2} \sin \frac{4\pi j}{k} \end{aligned}$$

since

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}\,.$$

The product sign here runs over real quantities. The absolute value of $G(1, k)$ is already determined as $\sqrt{k}$ by (11.7). Thus

$$G(1, k) = \sqrt{k}(-i)^{(k-1)/2} \cdot \operatorname{sign} \prod_{j=1}^{(k-1)/2} \sin \frac{4\pi j}{k}\,.$$

Now we have here

$$0 < \frac{4\pi j}{k} < 2\pi\,.$$

In this range the sine function has only one change of sign. We have

$$\sin \frac{4\pi j}{k} > 0 \text{ for } 0 < j < \frac{k}{4}$$

and

$$\sin \frac{4\pi j}{k} < 0 \text{ for } \frac{k}{4} < j < \frac{k}{2}\,;$$

therefore

$$\operatorname{sign} \prod_{j=1}^{(k-1)/2} \sin \frac{4\pi j}{k} = (-1)^{[k/2]-[k/4]}\,.$$

This power of -1 has just been discussed in the determination of $(2/k)$ and we therefore obtain

$$G(1, k) = \sqrt{k}(-i)^{(k-1)/2}(-1)^{(k^2-1)/8}\,.$$

But

$$\begin{aligned} (-1)^{(k^2-1)/8}(-i)^{(k-1)/2} &= i^{(k^2-1)/4} \cdot (-i)^{(k-1)/2} = i^{(k^2-1)/4-(k-1)/2} \\ &= i^{(k-1)/2[(k+1)/2]-1} = i^{[(k-1)/2]^2}\,, \end{aligned}$$

which finally yields the following theorem.

THEOREM 38: *The explicit value of the Gaussian sum for $\alpha = e^{2\pi i/k}$ is*

$$G(e^{2\pi i/k}) = G(1, k) = i^{[(k-1)/2]^2}\sqrt{k}\,. \tag{11.21}$$

Since (h/k) is explicitly known from Theorem 37 and since

$$G(h, k) = G(e^{2\pi ih/k}) = \left(\frac{h}{k}\right) G(e^{2\pi i/k}) = \left(\frac{h}{k}\right) G(1, k)\,,$$

the value of $G(h, k)$ is also known.

The reciprocity law for the Jacobi symbol. It is now a simple matter to establish the reciprocity property of the Jacobi symbol.

Let h and k be two odd natural numbers, $(h, k) = 1$. We then choose a and b so that

$$ha + kb = 1\,.$$

Then clearly

$$e^{2\pi i/hk} = e^{2\pi ib/h} e^{2\pi ia/k}\,.$$

We now put

$$\alpha = e^{2\pi ia/k}, \qquad \beta = e^{2\pi ib/h}\,,$$

which are both primitive roots of order k, h respectively. With these roots of unity, equation (11.19) of the lemma becomes

$$G(1, hk) = G(a, k)G(b, h)$$

in the new notation (11.20), and therefore

$$G(1, hk) = \left(\frac{a}{k}\right)G(1, k) \cdot \left(\frac{b}{h}\right)G(1, h) .$$

Now we have by (11.8) and (11.14)

$$1 = \left(\frac{ha + kb}{k}\right) = \left(\frac{ha}{k}\right) = \left(\frac{h}{k}\right)\left(\frac{a}{k}\right) \quad \text{by (11.16)},$$

so that

$$\left(\frac{a}{k}\right) = \left(\frac{h}{k}\right)$$

and similarly

$$\left(\frac{b}{h}\right) = \left(\frac{k}{h}\right) .$$

We have thus found

$$G(1, hk) = \left(\frac{h}{k}\right)\left(\frac{k}{h}\right) \cdot G(1, k)G(1, h) .$$

Here we simply insert (11.21) and obtain

$$\left(\frac{h}{k}\right)\left(\frac{k}{h}\right) = i^{[(hk-1)/2]^2-[(k-1)/2]^2-[(h-1)/2]^2} .$$

Now

$$\begin{aligned}\left(\frac{hk-1}{2}\right)^2 - \left(\frac{k-1}{2}\right)^2 - \left(\frac{h-1}{2}\right)^2 \\ = \tfrac{1}{4}[(k^2-1)(h^2-1) - 2(kh-k-h+1)] \\ = \frac{(k-1)(h-1)}{2} \cdot \left(\frac{(k+1)(h+1)}{2} - 1\right) \\ = 2 \cdot \frac{(k-1)(h-1)}{4} \cdot u ,\end{aligned}$$

where u is an odd number as $h+1$ and $k+1$ are even. If we put this into the previous result we obtain

$$\left(\frac{h}{k}\right)\left(\frac{k}{h}\right) = i^{2[(k-1)(h-1)/4]u} = (-1)^{[(k-1)/2\cdot(h-1)/2]}$$

(since $2u$ even implies $i^{2u} = -1$) and thus the reciprocity theorem.

THEOREM 39: *The Jacobi symbol has the reciprocity property for two coprime odd natural numbers h, k*

$$\left(\frac{h}{k}\right)\left(\frac{k}{h}\right) = (-1)^{(h-1)/2\cdot(k-1)/2} . \tag{11.22}$$

This is the full generalization of Theorem 34.

The Jacobi symbol is even useful in the computation of the Legendre symbol, since there is no need in the intervening steps always to find out whether the newly appearing odd numbers are prime or not.

In the *Example* at the end of Chapter 10 we would not have to find the decomposition $1829 = 31 \cdot 59$ but could proceed as follows:

$$\begin{aligned}\left(\frac{12703}{16361}\right) = \left(\frac{2}{12703}\right)\left(\frac{1829}{12703}\right) = \left(\frac{12703}{1829}\right) \\ = \left(\frac{1729}{1829}\right) = \left(\frac{-100}{1829}\right) = \left(\frac{-1}{1829}\right) = +1 .\end{aligned}$$

Some further properties of the Jacobi symbol. The Jacobi symbol forms a character, i.e., its "numerator" is multiplicative as (11.16) explicitly states. It turns out that its modulus or "denominator" shares this property. Indeed, let h, k, l be odd natural numbers with $(h, k) = (h, l) = 1$. Then we have, by means of (11.22) and (11.16),

$$\begin{aligned}\left(\frac{h}{k}\right)\left(\frac{h}{l}\right) &= \left(\frac{k}{h}\right)\left(\frac{l}{h}\right) \cdot (-1)^{[(h-1)/4](k-1+l-1)} \\ &= \left(\frac{kl}{h}\right) \cdot (-1)^{[(h-1)/4](k-1+l-1)} \\ &= \left(\frac{h}{kl}\right) (-1)^{[(h-1)/4](k-1+l-1-kl+1)}\end{aligned}$$

where we have observed $(-1)^a = (-1)^{-a}$ for an integer a. Now

$$k - 1 + l - 1 - kl + 1 = -(k-1)(l-1) ,$$

and

$$\tfrac{1}{4}(h-1)(k-1)(l-1)$$

is *even*. Therefore we have obtained in this case

$$\left(\frac{h}{k}\right)\left(\frac{h}{l}\right) = \left(\frac{h}{kl}\right) . \tag{11.23}$$

There remains the case $h = 2$.

We have, according to (11.18),

$$\left(\frac{2}{k}\right)\left(\frac{2}{l}\right) = \left(\frac{2}{kl}\right)(-1)^{(k^2-1)/8+(l^2-1)/8-(k^2l^2-1)/8} .$$

But

$$k^2 - 1 + l^2 - 1 - k^2l^2 + 1 = -(k^2 - 1)(l^2 - 1) \equiv 0 \pmod{64}$$

since k and l are odd and thus

$$\left(\frac{2}{k}\right)\left(\frac{2}{l}\right) = \left(\frac{2}{kl}\right). \tag{11.24}$$

We now take (11.23) and (11.24) together.

THEOREM 40. *If k, l are odd natural numbers and $(m, kl) = 1$ then the Jacobi symbol satisfies*

$$\left(\frac{m}{k}\right)\left(\frac{m}{l}\right) = \left(\frac{m}{kl}\right). \tag{11.25}$$

Remark: Because of (11.15) the "numerator" does not have to be positive.

This theorem now permits us to reduce the Jacobi symbol to a product of Legendre symbols. Let $k = p_1p_2 \cdots p_s$ be the prime number decomposition of k; then the repeated application of (11.25) yields

$$\left(\frac{m}{k}\right) = \left(\frac{m}{p_1p_2 \cdots p_s}\right) = \left(\frac{m}{p_1}\right)\left(\frac{m}{p_2}\right) \cdots \left(\frac{m}{p_s}\right). \tag{11.26}$$

(As a matter of fact, this is the usual way to define the Jacobi symbol.)

This decomposition shows that $(m/k) = +1$ is only necessary, but not sufficient, for the solvability of

$$x^2 \equiv m \pmod{k}.$$

If, for instance, $m = p_1p_2$ with $p_1 \neq p_2$, then the congruence implies

$$x^2 \equiv m \pmod{p_1}, \qquad x^2 \equiv m \pmod{p_2}$$

or

$$\left(\frac{m}{p_1}\right) = \left(\frac{m}{p_2}\right) = 1,$$

and this, by the Chinese remainder theorem, is then also sufficient for m to be a quadratic residue modulo k. However, $(m/k) = 1$ will also take place if $(m/p_1) = (m/p_2) = -1$, in which case $x^2 \equiv m \pmod{k}$ will not have a solution.

It is also clear from the decomposition (11.25) that $(m/k) = 1$ for *any* m prime to k if k is a perfect square. The symbol (m/k) is in this case the "principal character" modulo k. But this takes place *only* if k is a perfect square. We leave this statement to the reader as an exercise.

Exercise: Prove that for k odd and not a perfect square there exists always a number b such that $(b/k) = -1$.

12

Lattice Points

Introduction and a lemma. We call *lattice points* in n-dimensional space those points of which all n coordinates are integers. We shall restrict ourselves to lattice points in the plane. The problem will be: Given an area, how many lattice points are in it? This question will in general have a number-theoretical aspect because *integers* are involved as the coordinates of the lattice points.

Problems of the enumeration of lattice points arise quite naturally in the discussion of some arithmetical functions. Many of these functions are so irregular that it is advisable to consider averages of their values, in order to obtain smoother functions in which individual peculiarities of the summands are suppressed. We shall deal with three such functions and their averages, which can be interpreted as numbers of lattice points in certain domains.

(1) The function $r(n)$, the number of solutions in integers of the equation $x^2 + y^2 = n$. Here we study the smoother function

$$R(N) = \sum_{n \leq N} r(n),$$

which is related to the number of lattice points in a circle.

(2) The function $\sigma(n)$, the number of divisors of an integer n. We have $\sigma(n) = 2$ infinitely often, namely for n a prime. On the other hand $\sigma(n)$ can evidently increase beyond any bound for composite numbers. Here we are interested in

$$T(n) = \sum_{n \leq N} \sigma(n),$$

which can be interpreted as the number of lattice points under a certain equilateral hyperbola.

(3) Euler's function $\varphi(n)$, the number of numbers less than n and prime to n. The summatory function

$$\Phi(N) = \sum_{n \leq N} \varphi(n)$$

can here be related to the number of those lattice points in a square whose coordinates are coprime.

Lastly we shall use the device of lattice points to complete another proof of the reciprocity law of the Jacobi symbol, begun in Chapter 11.

It seems advisable to begin with a lemma on which we will have to depend later.

LEMMA: *If $g(t)$ is a monotonically decreasing function, $g(t) > 0$ for all $t > 0$, then*

$$\sum_{1 \leq n \leq X} g(n) = \int_1^X g(t)\,dt + C + O(g(X)).\dagger \tag{12.1}$$

Here n runs through integers only; X can be any real number, $X \geqq 1$; and C is a constant depending only on the function $g(t)$.

Proof: Since $g(t)$ is decreasing in the interval $[n, n+1]$, we have

$$g(n+1) \leqq \int_n^{n+1} g(t)\,dt \leqq g(n),$$

and thus

$$0 \leqq d_n = g(n) - \int_n^{n+1} g(t)\,dt \leqq g(n) - g(n+1).$$

Therefore we have for any positive integers $M < N$

$$\sum_{n=M}^{N} d_n \leqq \sum_{n=M}^{N} \{g(n) - g(n+1)\} = g(M) - g(N+1) < g(M).$$

This shows that the series

$$\sum_{n=1}^{\infty} d_n$$

converges. In particular, we have

$$\sum_{n=M}^{\infty} d_n \leqq g(M).$$

If we put

$$C = \sum_{n=1}^{\infty} d_n$$

we have

$$C = \sum_{n=1}^{N} d_n + \sum_{n=N+1}^{\infty} d_n = \sum_{n=1}^{N} \{g(n) - \int_n^{n+1} g(t)\,dt\} + O(g(N+1)).$$

It follows that

$$\sum_{n=1}^{N} g(n) = \int_1^{N+1} g(t)\,dt + C + O(g(N+1)).$$

† The statement $f(X) = O(g(X))$ means that there exists a certain constant K such that

$$|f(X)| < Kg(X) \text{ for all } X.$$

For $N = [X]$ this becomes

$$\sum_{1 \leq n \leq X} g(n) = \int_1^{[X]+1} g(t)\,dt + C + O(g([X]+1))$$

$$= \int_1^X g(t)\,dt + C + O(g(X)),$$

since

$$\int_X^{[X]+1} g(t)\,dt \leqq g(X) \text{ and}$$

$$0 < g([X]+1) \leqq g(X).$$

This proves the lemma.

COROLLARY 1:

$$\sum_{1 \leq n \leq X} \frac{1}{n} = \log X + \gamma + O\left(\frac{1}{X}\right). \tag{12.2}$$

The constant γ here is called the Euler-Mascheroni constant. Its value is approximately 0.5772157 ····. It is defined according to (12.2) as

$$\gamma = \lim_{N \to \infty} \left\{ \sum_{n=1}^{N} \frac{1}{n} - \log N \right\}.$$

COROLLARY 2:

$$\sum_{2 \leq n \leq X} \frac{1}{n \log n} = \log\log X + C + O\left(\frac{1}{X \log X}\right). \tag{12.3}$$

We have to observe here only

$$\int_2^X \frac{dt}{t \log t} = \log\log X - \log\log 2.$$

Lattice points in a circle. We have defined $r(n)$ and $R(N)$ above and find

$$R(N) = \sum_{n \leq N} r(n) = \sum_{n \leq N} \sum_{x^2+y^2=n} 1 = \sum_{x^2+y^2 \leq N} 1,$$

where x and y run only through integers. This formula shows that $R(N)$ is the number of lattice points in the interior or on the boundary of a circle of radius $\sqrt{N}$. From this point of view we shall now proceed to obtain an approximation to $R(N)$.

To each lattice point in and on the circle we attach a square with unit sides, in such a way, let us say, that the lattice point forms the "southwest" corner of the square. Then the area of these squares together is equal to $R(N)$. This is, however, not quite equal to the area of the circle. Some squares protrude beyond the circle and there is, on the other hand, some

unfilled area in the circle. However, all the chosen squares are contained in a circle of radius $\sqrt{N} + \sqrt{2}$, since the diagonal of each square has the length $\sqrt{2}$. Thus, comparing areas, we obtain

$$R(N) < \pi(\sqrt{N} + \sqrt{2})^2.$$

Similarly the circle of radius $\sqrt{N} - \sqrt{2}$ is entirely covered by those squares so that

$$R(N) > \pi(\sqrt{N} - \sqrt{2})^2.$$

We have therefore

$$\pi(N - 2\sqrt{2N} + 2) < R(N) \leqq \pi(N + 2\sqrt{2N} + 2),$$

or, briefly, the following theorem.

THEOREM 41:

$$R(N) = \pi N + O(\sqrt{N}). \tag{12.4}$$

We may consider the analogue of $R(N)$ in k dimensions. Let $R_k(N)$ be the number of lattice points inside a k-dimensional sphere of radius $\sqrt{N}$. In this notation we have $R(N) = R_2(N)$. An argument similar to the one we have given shows

$$R_3(N) = \tfrac{4}{3}\pi N^{\frac{3}{2}} + O(N).$$

Exercise: Prove this statement about $R_3(N)$ in detail and find a similar statement concerning $R_4(N)$.

Our result about $R(N)$ was known to Gauss about 1800, and it was not improved until 1906 when the Polish mathematician W. Sierpinski proved the surprising result

$$R(N) = \pi N + O(N^{\frac{1}{2}}).$$

But this is not all: In 1923 van der Corput proved that the exponent $\frac{1}{2}$ is not the best possible, but can be replaced by a certain number $\theta < \frac{1}{2}$. In particular, his pupil I. W. Nielandt computed a suitable number θ, obtaining

$$R(N) = \pi N + O(N^{27/82}),$$

where $\frac{27}{82} < \frac{27}{81} = \frac{1}{3}$. How far this exponent θ can be lowered is not known at present. But through the work of G. H. Hardy and E. Landau we know that the formula

$$R(N) = \pi N + O(N^{\frac{1}{4}})$$

is *certainly false*.

The summatory function of the number of divisors. The number $\sigma(n)$ of divisors of a natural number n can be expressed as

$$\sigma(n) = \sum_{x|n} 1 = \sum_{xy=n} 1$$

where x runs through all divisors of n. We define the summatory function

$$T(N) = \sum_{1 \leq n \leq N} \sigma(n)$$

and then have

$$T(N) = \sum_{1 \leq n \leq N} \sum_{xy=n} 1 = \sum_{1 \leq xy \leq N} 1. \tag{12.5}$$

This shows that $T(N)$ is the number of lattice points (x, y) in the first quadrant which lie under or on the hyperbola $xy = N$. (The points on the axes would give $xy = 0$ and are thus excluded.)

We count the lattice points on each vertical line $x =$ integer, on or below the hyperbola. Since the lattice points are spaced one unit apart, we have $[N/x]$ lattice points on the ordinate of length N/x and altogether

$$T(N) = \sum_{x=1}^{N} \left[\frac{N}{x}\right].$$

If we put

$$\left[\frac{N}{x}\right] = \frac{N}{x} - \theta_x, \qquad 0 \leq \theta_x < 1,$$

the above becomes

$$\begin{aligned} T(N) &= N \sum_{x=1}^{N} \frac{1}{x} - \sum_{x=1}^{N} \theta_x \\ &= N \sum_{x=1}^{N} \frac{1}{x} + O(N) \\ &= N \log N + O(N) \end{aligned} \tag{12.6}$$

in virtue of Corollary 1. Dirichlet showed by a simple device that this asymptotic formula can be improved considerably. See Fig. 4.

Since the hyperbola is symmetric about the line $x = y$, the number of lattice points in the area $OBAEG$ is equal to the number of lattice points in the area $OCADF$. Of course there are only a finite number of lattice points below the hyperbola, since we exclude the lattice points of the axes. Thus the number of lattice points under the hyperbola is twice the number in one of the aforementioned areas, minus those in $OCAB$, which we have counted twice. We can therefore replace (12.6) by

$$\begin{aligned} T(N) &= 2 \sum_{\substack{1 \leq x \leq \sqrt{N} \\ 1 \leq xy \leq N}} 1 - [\sqrt{N}]^2 \\ &= 2 \sum_{1 \leq x \leq \sqrt{N}} \sum_{1 \leq y \leq N/x} 1 - [\sqrt{N}]^2 \\ &= 2 \sum_{1 \leq x \leq \sqrt{N}} \left[\frac{N}{x}\right] - [\sqrt{N}]^2. \end{aligned}$$

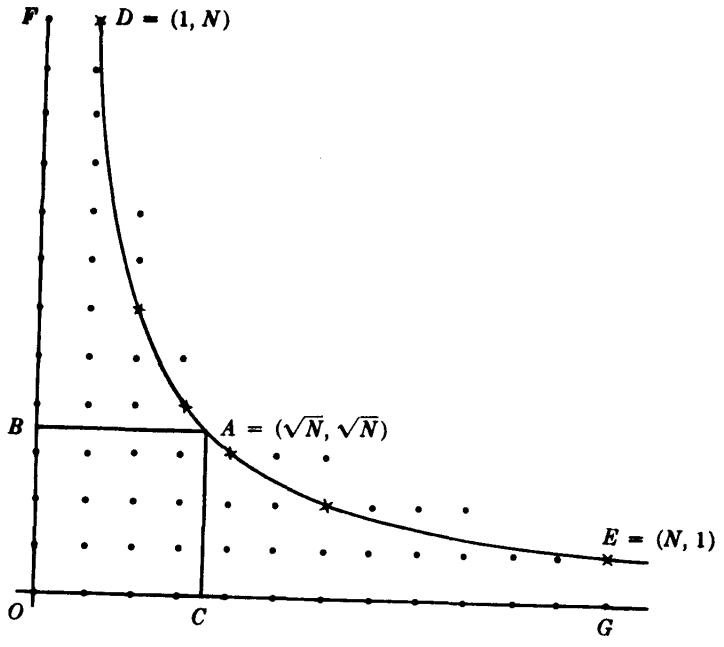

Figure 4

Now since for any real number z,

$$[z] = z - \theta, \qquad 0 \leqq \theta < 1,$$

we obtain

$$T(N) = 2 \sum_{1 \leqq x \leqq \sqrt{N}} \frac{N}{x} - 2 \sum_{1 \leqq x \leqq \sqrt{N}} \theta - (\sqrt{N} - \theta)^2,$$

with the θ in different meanings but always $0 \leqq \theta < 1$. This leads to

$$T(N) = 2N \sum_{1 \leqq x \leqq \sqrt{N}} \frac{1}{x} + O(\sqrt{N}) - N + O(\sqrt{N}) + O(1),$$

and in consequence of Corollary 1, to the next theorem.

THEOREM 42: *The summatory function $T(N)$ of $\sigma(n)$ fulfills*

$$T(N) = N \log N + (2\gamma - 1)N + O(\sqrt{N}). \tag{12.7}$$

This is an improvement over (12.6), since the error term $O(N)$ there is now decomposed into a precise term $(2\gamma - 1)N$ and an error term $O(\sqrt{N})$ of lower order.

The formula (12.7) like (12.4) has been the subject of intensive investigation in this century. First the Russian mathematician Voronoi showed in

1903 that the error term $O(\sqrt{N})$ can be replaced by $O(N^{\frac{1}{2}} \log N)$. Again van der Corput could show that the error term is of the order $O(N^{\theta})$ with a certain $\theta < \frac{1}{2}$. The lowest permissible value of θ is unknown here also.

In the two examples which we have discussed so far, $R(x)$ and $T(x)$, the summands $r(n)$ and $\sigma(n)$ take on small values infinitely often as can easily be seen: e.g., $r(n) = 0$ for $n \equiv 3 \pmod 4$ and $\sigma(n) = 2$ for n a prime number. Both functions also attain arbitrarily large values. For $\sigma(n)$ this is clear: We need only take n as a number with k different primes, and then have $\sigma(n) = 2^k$. For $r(n)$, it can be derived from our discussions of numbers which are divisors of the sum of two coprime squares. In the next example we treat the arithmetical function $\varphi(n)$, which increases with increasing n, but rather irregularly, so that again it is advisable to investigate its summatory function.

A digression: the Moebius function. Presently we shall need an arithmetical function, the Moebius function $\mu(n)$, defined for natural numbers n, which will be most useful for the solution of certain systems of linear equations.

DEFINITION: *The function $\mu(n)$, $n \geqq 1$, satisfies*

(I) $$\mu(1) = 1$$

(II) $$\sum_{d \mid n} \mu(d) = 0 \quad \text{for} \quad n > 1.$$

This is evidently a recursive definition. Indeed, (II) shows

$$\mu(1) + \mu(2) = 0, \text{ so that } \mu(2) = -\mu(1) = -1;$$
$$\mu(1) + \mu(3) = 0, \text{ so that } \mu(3) = -\mu(1) = -1;$$
$$\mu(1) + \mu(2) + \mu(4) = 0, \text{ so that } \mu(4) = 0;$$

and so on.

THEOREM 43: *The function $\mu(n)$, defined by* (I) *and* (II), *has the following properties:*

(a) $\mu(p) = -1$ *if p is prime*
(b) $\mu(p^k) = 0$ *for $k > 1$*
(c) $\mu(n_1 n_2) = \mu(n_1) \cdot \mu(n_2)$, *if* $(n_1, n_2) = 1$ (12.8)
(d) $\mu(n) = 0$ *if $p^2 \mid n$ for some p*
(e) $\mu(n) = (-1)^r$ *if $n = p_1 p_2 \cdots p_r$ is a product of r distinct primes.*

Proof: Since the only divisors of p are 1 and p itself, we have $\mu(1) + \mu(p) = 0$ and so $\mu(p) = -1$. Then for $k \geqq 2$ we have

$$0 = \sum_{d \mid p^k} \mu(d) = \mu(1) + \mu(p) + \mu(p^2) + \cdots + \mu(p^k),$$

and because of the foregoing

$$\mu(p^2) + \cdots + \mu(p^k) = 0\,,\ k = 2, 3, 4, \cdots.$$

This shows (b) by setting k successively equal to 2, 3, ·····.

As far as (c) is concerned, it is trivially true for $n_1 = 1$ or $n_2 = 1$. Now let $(n_1, n_2) = 1$, and without loss of generality $n_1 > 1$, $n_2 > 1$. Let us assume (c) for all l_1, l_2 with $(l_1, l_2) = 1$ and $l_1 l_2 < n_1 n_2$. If $d \mid n_1 n_2$, then it follows from $(n_1, n_2) = 1$ that there exist unique d_1, d_2 such that $d = d_1 d_2$, $d_1 \mid n_1$, and $d_2 \mid n_2$. Thus

$$0 = \sum_{d \mid n_1 n_2} \mu(d) = \sum_{\substack{d_1 \mid n_1 \\ d_2 \mid n_2}} \mu(d_1 d_2).$$

Here we can apply the induction assumption for $d_1 d_2 < n_1 n_2$ and have

$$0 = \sum_{\substack{d_1 \mid n_1 \\ d_1 d_2 < n_1 n_2}} \sum_{d_2 \mid n_2} \mu(d_1)\mu(d_2) + \mu(n_1 n_2).$$

Extending the sum now over *all* divisors of $n_1 n_2$ we obtain

$$\begin{aligned} 0 &= \sum_{d_1 \mid n_1} \sum_{d_2 \mid n_2} \mu(d_1)\mu(d_2) - \mu(n_1)\mu(n_2) + \mu(n_1 n_2) \\ &= \sum_{d_1 \mid n_1} \mu(d_1) \sum_{d_2 \mid n_2} \mu(d_2) - \mu(n_1)\mu(n_2) + \mu(n_1 n_2) \\ &= -\mu(n_1)\mu(n_2) + \mu(n_1 n_2), \end{aligned}$$

which proves (c). The statements (d) and (e) are now simple consequences of (a), (b), and (c).

We now use the Moebius function $\mu(n)$ for a general "inversion" formula.

THEOREM 44: *Let $f(t)$ and $F(t)$ be functions of the real variable $t \geqq 1$. If these functions have the relation*

$$F(t) = \sum_{1 \leq n \leq t} f\left(\frac{t}{n}\right), \tag{12.9}$$

then they satisfy the "inverse" relation

$$f(t) = \sum_{1 \leq m \leq t} \mu(m) F\left(\frac{t}{m}\right). \tag{12.10}$$

Conversely, (12.9) *follows from* (12.10).

Proof: We have from (12.9)

$$\begin{aligned} \sum_{1 \leq m \leq t} \mu(m) F\left(\frac{t}{m}\right) &= \sum_{1 \leq m \leq t} \mu(m) \sum_{1 \leq n \leq t/m} f\left(\frac{t}{mn}\right) \\ &= \sum_{\substack{m,n \\ 1 \leq mn \leq t}} \mu(m) f\left(\frac{t}{mn}\right). \end{aligned}$$

Here we sum over all lattice points (m, n), $m \geqq 1$, $n \geqq 1$, which lie under or on the hyperbola $mn = t$. We now rearrange the sum in such a way that we assemble terms with $mn = r$, $1 \leqq r \leqq t$. Then the last sum becomes

$$\sum_{1 \leq r \leq t} \sum_{m \mid r} \mu(m) f\left(\frac{t}{r}\right) = \sum_{1 \leq r \leq t} f\left(\frac{t}{r}\right) \sum_{m \mid r} \mu(m) = f(t),$$

according to (I) and (II) of the definition of $\mu(m)$. This derives (12.10) from (12.9). The converse is proved by a similar argument.

Exercise 1: Prove the following inversion formula, which is different from that of Theorem 44.

Theorem 45: Let $g(n)$, $G(n)$ be arithmetical functions, i.e., defined for all natural numbers n. If they satisfy the relation

$$G(n) = \sum_{d \mid n} g(d), \qquad n = 1, 2, 3, \cdots, \tag{12.11}$$

then they fulfill also

$$g(n) = \sum_{d \mid n} \mu(d) G\left(\frac{n}{d}\right). \tag{12.12}$$

Remark: If we write down (12.11) for the values $n = 1, 2, \cdots, N$, we get a system of N linear equations for $g(1), g(2), \cdots, g(N)$. Then (12.12) can be looked upon as the solution for $g(m)$ in terms of $G(n)$.

Exercise 2: Use Theorem 45 to deduce Euler's formula (3.6) for $\varphi(n)$ from the property (3.8)

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

The summatory function $\Phi(t)$ of the Euler function $\varphi(n)$. We define, for real $t \geqq 1$,

$$\begin{aligned} \Phi(t) &= \sum_{1 \leq n \leq t} \varphi(n) = \sum_{1 \leq n \leq t} \sum_{\substack{1 \leq m \leq n \\ (m,n) = 1}} 1 \\ &= \sum_{\substack{1 \leq m \leq n \leq t \\ (m,n) = 1}} 1. \end{aligned}$$

The sum can be interpreted as the number of lattice points with coprime integer coordinates m, n in the right triangle $0 < y \leqq x \leqq t$. Of these lattice points only 1,1 lies *on* the line $x = y$.

Let $\Psi(t)$ be the number of lattice points with coprime coordinates in the square

$$0 < x \leqq t$$
$$0 < y \leqq t.$$

See Fig. 5. Then, because of symmetry with respect to the line $x = y$, we have

$$\Psi(t) = 2\Phi(t) - 1, \tag{12.13}$$

where the subtraction of 1 arises from the fact that in $\Phi(t)$ as well as in $\Psi(t)$ the point 1,1 is counted once. The function $\Psi(t)$ is somewhat easier to discuss than $\Phi(t)$, to which we shall return in the end.

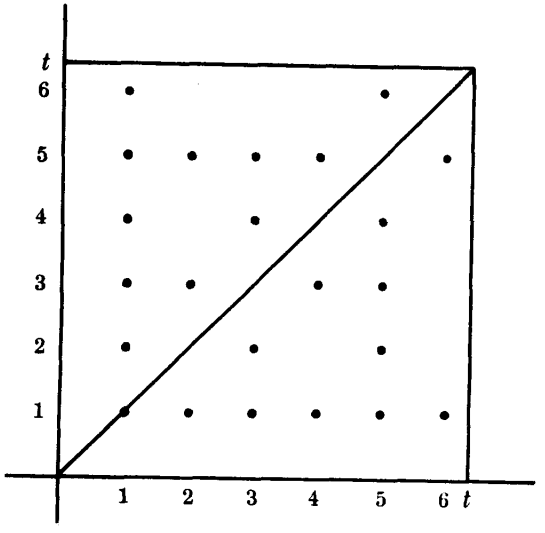

Figure 5

The *total number* of lattice points in the square

$$0 < x \leqq t, \qquad 0 < y \leqq t$$

is $[t]^2$:

$$[t]^2 = \sum_{\substack{0<m\leqq t\\ 0<n\leqq t}} 1.$$

Here we can sort the lattice points according to the greatest common divisor d of their coordinates m, n:

$$[t]^2 = \sum_{1\leqq d\leqq t} \sum_{\substack{0<m\leqq t\\ 0<n\leqq t\\ (m,n)=d}} 1. \tag{12.14}$$

Now $(m, n) = d$ if and only if $(m/d, n/d) = 1$. Thus there is a 1–1 correspondence between points m, n with

$$0 < m \leqq t, \qquad 0 < n \leqq t, \qquad (m, n) = d,$$

and pairs m', n' with

$$0 < m' \leqq \frac{t}{d}, \qquad 0 < n' \leqq \frac{t}{d}, \qquad (m', n') = 1.$$

There are, however, $\Psi(t/d)$ of the latter. This enables us to rewrite (12.14) as

$$[t]^2 = \sum_{1\leqq d\leqq t} \Psi\left(\frac{t}{d}\right).$$

We can now apply Theorem 44 to this formula and obtain

$$\begin{aligned}\Psi(t) &= \sum_{1\leqq d\leqq t} \mu(d)\left[\frac{t}{d}\right]^2 \\ &= \sum_{1\leqq d\leqq t} \mu(d)\left(\frac{t}{d} + O(1)\right)^2 \\ &= t^2 \sum_{1\leqq d\leqq t} \frac{\mu(d)}{d^2} + 2t\,O\left(\sum_{1\leqq d\leqq t} \frac{1}{d}\right) + O\left(\sum_{1\leqq d\leqq t} 1\right).\end{aligned}$$

We consider these three terms separately. We have

$$\sum_{1\leqq d\leqq t} \frac{\mu(d)}{d^2} = \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - \sum_{d=[t]+1}^{\infty} \frac{\mu(d)}{d^2}.$$

Now

$$\left|\sum_{d=[t]+1}^{\infty} \frac{\mu(d)}{d^2}\right| < \sum_{d=[t]+1}^{\infty} \frac{1}{d^2} < \int_{[t]}^{\infty} \frac{du}{u^2} = \frac{1}{[t]} = O\left(\frac{1}{t}\right).$$

To the second term we apply Corollary 1 of this chapter, and the third term is $O(t)$. Thus we obtain altogether

$$\Psi(t) = t^2 \cdot S + O(t \log t), \tag{12.15}$$

where

$$S = \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2}. \tag{12.16}$$

Euler's product expansion of the ζ-function. The sum S can be evaluated explicitly. Let us consider for $s > 1$ the absolutely convergent series

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

(The symbol $\zeta(s)$ was used for this series by Riemann in 1859 and has been adopted universally.) Euler, making use of the uniqueness of prime

factorization, found the equality

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \left(1 + \frac{1}{2^s} + \frac{1}{2^{2s}} + \cdots\right)\left(1 + \frac{1}{3^s} + \frac{1}{3^{2s}} + \cdots\right) \times \left(1 + \frac{1}{5^s} + \frac{1}{5^{2s}} + \cdots\right)\cdots = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots\right),$$

where the product is extended over all prime numbers p. Indeed, because of absolute convergence of the series (which contains only positive terms anyway for $s > 1$), we can rearrange the terms in all infinite series appearing in the left- and right-hand members. Now in each term of the left-hand member, the prime factorization of n can be performed uniquely, $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$, so that

$$\frac{1}{n^s} = \frac{1}{p_1^{a_1 s}} \cdot \frac{1}{p_2^{a_2 s}} \cdots \frac{1}{p_k^{a_k s}},$$

and thus each term $1/n^s$ is found once and only once among the products obtained by multiplying out the right-hand member of the formula. This proves the equality. We realize, moreover, that each infinite series on the right is a geometric series, which can be summed, thus obtaining

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - (1/p^s)}. \tag{12.17}$$

This is Euler's product. (Riemann's investigations of the ζ-function are concerned with complex values of s, which Euler did not consider. The ζ-function is the paramount tool of modern research on the distribution of primes.)

The reciprocal of (12.17) is of interest for our problem. We have

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) = \left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right)\left(1 - \frac{1}{5^s}\right)\cdots = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

We understand this identity if we realize that in the last sum only those n appear in the denominator which have different prime factors (because of $\mu(n)$); and $\mu(n) = (-1)^r$, where r is the number of prime factors in n, according to (12.8).

For $s = 2$ we obtain

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

It is now known† that

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6},$$

a result which was also obtained by Euler. Using these results, we now obtain from (12.13), (12.15), and (12.16) our final result.

THEOREM 46:

$$\Phi(t) = \sum_{1 \le n \le t} \varphi(n) = \frac{3}{\pi^2} t^2 + O(t \log t). \tag{12.18}$$

Again the reciprocity of the Jacobi symbol. In Theorem 37 the Jacobi symbol is evaluated explicitly. Here $(h, k) = 1$ and k is an *odd* natural number. We first transform (11.17) slightly for the case that h is also an odd natural number.

Since $h + k$ is now even, we have

$$h + k = 2 \cdot \frac{h + k}{2}$$

in integers and thus

$$\left(\frac{h}{k}\right) = \left(\frac{h + k}{k}\right) = \left(\frac{2}{k}\right) \cdot \left(\frac{(h + k)/2}{k}\right).$$

Therefore from (11.17)

$$\left(\frac{h}{k}\right) = \left(\frac{2}{k}\right)(-1)^{\sum_{j=1}^{(k-1)/2} \left[\frac{(h+k)j}{k}\right]}.$$

Now

$$\sum_{j=1}^{(k-1)/2} \left[\frac{(h + k)j}{k}\right] = \sum_{j=1}^{(k-1)/2} \left[\frac{hj}{k}\right] + \sum_{j=1}^{(k-1)/2} j = \sum_{j=1}^{(k-1)/2} \left[\frac{hj}{k}\right] + \frac{k^2 - 1}{8}.$$

If we compare this result and our previous equation with (11.18), we obtain

$$\left(\frac{h}{k}\right) = (-1)^{\sum_{j=1}^{(k-1)/2} \left[\frac{hj}{k}\right]}, \tag{12.19}$$

valid for coprime odd positive h and k. One sees that the sum in the last formula does not have the factor 2 in the numerator which the formula (11.17) shows. We interchange the role of h and k in (12.19) and find through multiplication

$$\left(\frac{h}{k}\right)\left(\frac{k}{h}\right) = (-1)^M$$

† See, e.g., K. Knopp, *Theory and Application of Infinite Series*, Blackie & Son, London, 1951, pages 237, 323, and 376.

where

$$M = \sum_{j=1}^{(k-1)/2} \left[\frac{hj}{k}\right] + \sum_{i=1}^{(h-1)/2} \left[\frac{ki}{h}\right].$$

This sum now can be interpreted as the number of lattice points in the rectangle with the vertices $(0, 0)$, $(h/2, 0)$, $(h/2, k/2)$, $(0, k/2)$. Indeed, let us

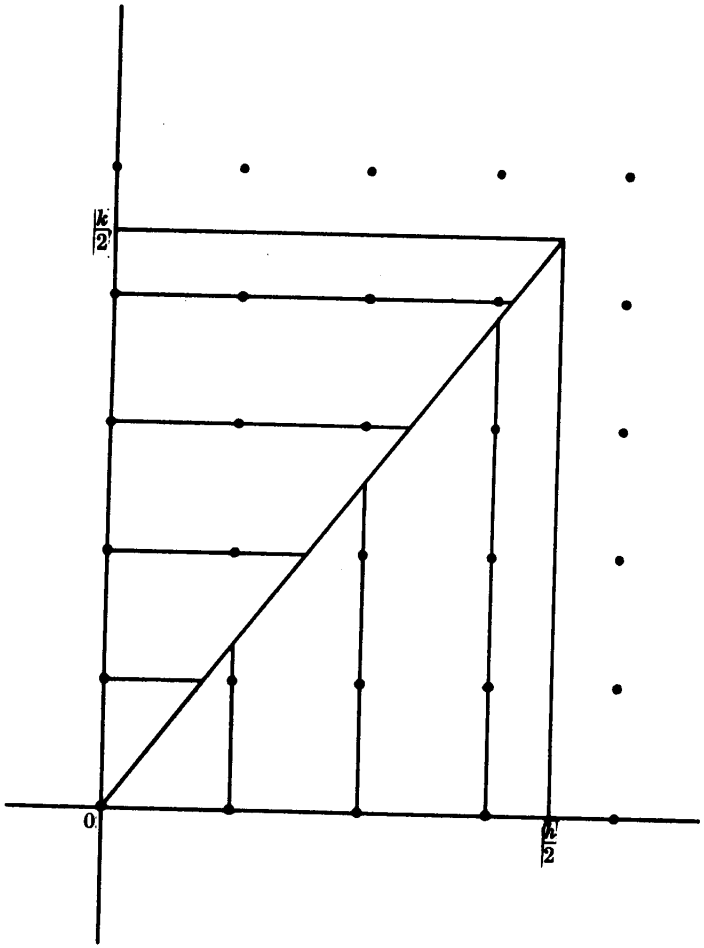

Figure 6

draw the diagonal in this rectangle issuing from 0, 0. See Fig. 6. It has slope k/h. Since h and k are coprime, the point (h, k) is the lattice point on it which is nearest to the origin. Inside the rectangle the diagonal does not meet any lattice point. We count the lattice points inside the rectangle by rows:

Below the diagonal along ordinates, above the diagonal along abscissas. Then

$$\sum_{1 \le i < h/2} \left[\frac{k}{h} i\right]$$

is the number of lattice points below the diagonal inside the rectangle, and

$$\sum_{1 \le j < k/2} \left[\frac{h}{k} j\right]$$

is the number of lattice points above the diagonal. Thus M is the total number of lattice points in the rectangle under discussion. But this number is evidently

$$M = \frac{h-1}{2} \cdot \frac{k-1}{2},$$

so that we obtain

$$\left(\frac{h}{k}\right)\left(\frac{k}{h}\right) = (-1)^{((h-1)(k-1))/4}, \tag{12.20}$$

the reciprocity law for the Jacobi symbol (Theorem 39). This is the proof which we announced on page 91.

It is worthwhile to note that our previous proof depends on the explicit evaluation (11.21) of the Gaussian sum for the primitive root $\alpha = e^{2\pi i/k}$ and therefore implies a transcendental element in the reasoning. The present proof is purely algebraic, since it does not specify the primitive kth root α. (The use of geometry is only an apparent one, since all arguments can be written by means of the symbol $[x]$.)

Exercise: Let p be a prime $\equiv 1 \pmod 4$. Prove the formula of Buniakovski,

$$\sum_{\lambda=1}^{(p-1)/4} \left[\sqrt{p\lambda}\right] = \frac{p^2-1}{12}.$$

Hint: Count the lattice points *above* the parabola $y = \sqrt{px}$ in a suitable rectangle.

13

About the Distribution of Prime Numbers

Historical remarks. We know that the sequence of prime numbers has no end. We have also seen that the primes are rather irregularly distributed among the natural numbers. Still it is obviously of interest to have some idea about the occurrence of primes among the natural numbers. How many primes are there less than or equal to a given number x? Let us call $\pi(x)$ this number of primes not exceeding x. Gauss conjectured by inspecting a table of primes that

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

The sign $\sim$ is read "is asymptotic to" and the above formula means

$$\frac{\pi(x)}{x/\log x} \to 1 \text{ as } x \to \infty.$$

Gauss made his conjecture in 1792. The conjecture, known today as the *prime number theorem*, was proved for the first time more than a hundred years later by Hadamard and de la Vallée Poussin.

These two mathematicians continued investigations begun by Riemann about the zeta function $\zeta(s)$ for the complex variable s. But before this achievement was attained, the Russian mathematician Chebyshev was able to compare in some way the functions $\pi(x)$ and $x/\log x$ which appear in Gauss's conjecture. Chebyshev proved by elementary methods that two positive constants c and C exist such that

$$c\frac{x}{\log x} < \pi(x) < C\frac{x}{\log x}, \quad x \geq 2.$$

His work was supplemented by Mertens in the 1880's. But in this century the opinion still prevailed among mathematicians that for the proof of the prime number theorem the theory of Riemann's function $\zeta(s)$ was indispensable. It was a great surprise in the world of mathematics, therefore, when in 1950 A. Selberg and P. Erdös succeeded in proving the prime number theorem by arguments which start from Chebyshev's research and use only "elementary" methods in the sense that no complex function theory is used. We shall prove here only Chebyshev's result and some remarks of Mertens.

A divergent series involving prime numbers. We start with Chebyshev with the factorial function

$$n! = 1 \cdot 2 \cdot \cdots \cdot n.$$

With Legendre we ask for the highest power of a given prime p dividing $n!$ There are $[n/p]$ multiples of p not exceeding n, which would contribute $p^{[n/p]}$ to factors of $n!$ But some multiples of p are also multiples of $p^2, p^3, \cdots$, which furnish more p factors. There are $[n/p^2]$ numbers less than or equal to n divisible by p^2, which would give 2 factors of p each. However, one of them has already been counted in the multiples of p itself, so that we now have $[n/p] + [n/p^2]$ factors p. In the same manner we have to consider the $[n/p^3]$ multiples of p^3, each furnishing 3 factors p, of which however, 2 have already been counted. Progressing in this way we find that p appears in $n!$ to the power

$$\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots,$$

where we may continue this sum formally to infinity, since $[n/p^k] = 0$ for $p^k > n$. This determination of the power of p in $n!$ goes for any prime number p. Taking all primes together we thus obtain Legendre's formula

$$n! = \prod_p p^{\sum_k [n/p^k]}. \tag{13.1}$$

We put this into more convenient forms at the price of losing some precision. We have

$$\frac{n}{p} - 1 < \sum_{k=1}^{\infty} \left[\frac{n}{p^k}\right] < \frac{n}{p} + \sum_{k=2}^{\infty} \frac{n}{p^k} = \frac{n}{p} + \frac{n}{p(p-1)},$$

and thus from (13.1),

$$\prod_{p \leq n} p^{((n/p)-1)} < n! < \prod_{p \leq n} p^{n/p + n/(p(p-1))},$$

and taking logarithms

$$\sum_{p \leq n} \left(\frac{n}{p} - 1\right) \log p < \log n! < \sum_{p \leq n} \left(\frac{n}{p} + \frac{n}{p(p-1)}\right) \log p. \tag{13.2}$$

Since

$$\sum_{p \leq n} \frac{\log p}{p(p-1)} < \sum_{m=2}^{\infty} \frac{\log m}{m(m-1)} = C_1$$

(the series being convergent), it follows that

$$\log n! < n\left\{\sum_{p \leq n} \frac{\log p}{p} + C_1\right\}.$$

We shall use this to get a statement about the right-hand member which shows the primes explicitly. For the estimation of the factorial on the left

side we could use Stirling's formula, but for our purpose the following argument suffices:

$$e^n = 1 + \frac{n}{1!} + \frac{n^2}{2!} + \cdots + \frac{n^n}{n!} + \cdots > \frac{n^n}{n!},$$

and thus

$$n! > \left(\frac{n}{e}\right)^n$$

$$\log n! > n(\log n - 1),$$

which leads to

$$\sum_{p \leq n} \frac{\log p}{p} > \log n - 1 - C_1. \tag{13.3}$$

For an upper estimate of the same sum we use the left-hand inequality in (13.2):

$$n \sum_{p \leq n} \frac{\log p}{p} - \sum_{p \leq n} \log p < \log n! < \log n^n = n \log n,$$

and obtain

$$\sum_{p \leq n} \frac{\log p}{p} < \log n + \frac{1}{n} \sum_{p \leq n} \log p. \tag{13.4}$$

Now we have to estimate the latter sum, for which we use a device due to Landau. If $m < p \leqq 2m$ then the prime number p divides $(2m)!$ but not $m!$. Therefore such a p divides the binomial coefficient

$$\frac{(2m)!}{m!\, m!} = \binom{2m}{m},$$

and this is true for all primes p with $m < p \leqq 2m$:

$$\prod_{m < p \leq 2m} p \,\Big|\, \binom{2m}{m}$$

and therefore

$$\prod_{m < p \leq 2m} p \leqq \binom{2m}{m}.$$

Since

$$2^{2m} = (1+1)^{2m} = \sum_{j=0}^{2m} \binom{2m}{j} > \binom{2m}{m},$$

we have

$$\prod_{m < p \leq 2m} p < 2^{2m} = 4^m < 4^{2m}. \tag{13.5}$$

For $p > 2$ all primes are odd and the equality sign under the product is therefore useless for an integer $m > 1$:

$$\prod_{\frac{2m+1}{2} < p \leq 2m-1} p < 4^m$$

and thus, admitting possibly one more prime factor,

$$\prod_{\frac{2m-1}{2} < p \leq 2m-1} p \leqq m \prod_{\frac{2m+1}{2} < p \leq 2m-1} p < m \cdot 4^m \leqq 4^{2m-1} \tag{13.6}$$

since $m \leqq 4^{m-1}$ for all integers $m > 0$. Formulae (13.5) and (13.6) show that for odd *and* even integers $r > 1$ we have

$$\prod_{r/2 < p \leq r} p < 4^r.$$

But this remains true if we introduce instead of the integer r any real number $x \geqq 2$, since the condition

$$\frac{x}{2} < p \leqq x$$

puts the same restriction on the integer p as

$$\frac{[x]}{2} < p \leqq [x].$$

This gives us

$$\prod_{x/2 < p \leq x} p < 4^{[x]} \leqq 4^x.$$

Writing here $x/2$ for x we have, provided $(x/4) \geqq 1$,

$$\prod_{x/4 < p \leq x/2} p < 4^{x/2},$$

and we can continue

$$\prod_{x/8 < p \leq x/4} p < 4^{x/4},$$

and so on. A finite number of such inequalities will include all prime numbers $p \leqq x$. We multiply these inequalities and obtain

$$\prod_{p \leq x} p < 4^{x + x/2 + x/4 + \cdots} < 4^{2x},$$

and thus

$$\sum_{p \leq x} \log p < x \log 16.$$

This we apply in (13.4) with the result

$$\sum_{p \leq n} \frac{\log p}{p} < \log n + \log 16.$$

If we also observe (13.3), we find that we have proved the theorem of Mertens.

THEOREM 47:

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

Since $\log x \to \infty$ with $x \to \infty$, the theorem states that the sum taken over *all* primes must diverge.

Another sum concerning primes. This theorem will enable us to prove Chebyshev's result about $\pi(x)$. However, before studying $\pi(x)$ we consider the function

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

which is a little simpler to handle. We have already found

$$\theta(x) < x \log 16 . \tag{13.7}$$

Looking for a lower bound, we have clearly

$$\theta(x) \geqq \sum_{x/A < p \leq x} \log p$$

where $A > 1$ is a constant that remains to be chosen. Then

$$\theta(x) \geqq \frac{x}{A} \sum_{x/A < p \leq x} \frac{\log p}{p} = \frac{x}{A} \left\{ \sum_{p \leq x} \frac{\log p}{p} - \sum_{p \leq x/A} \frac{\log p}{p} \right\} .$$

Theorem 47 now states that there exists a positive constant B such that for every $x \geqq 1$

$$-B < \sum_{p \leq x} \frac{\log p}{p} - \log x < B .$$

If we insert this in the previous inequality, we obtain for $x \geqq A$,

$$\theta(x) \geqq \frac{x}{A} \left\{ \log x - \log \frac{x}{A} - 2B \right\}$$

$$= \frac{x}{A} \{ \log A - 2B \} .$$

Here we choose A so that $\log A - 2B \geqq 1$, and have

$$\theta(x) \geqq \frac{x}{A} .$$

This, however, is proved only for $x \geqq A$, but for $2 \leqq x \leqq A$ there exists a lower bound of the positive function $(\theta(x)/x)$. With a suitable positive k we have

$$\theta(x) \geqq kx \quad \text{for} \quad x \geqq 2 .$$

We have thus proved the following theorem.

THEOREM 48: *There exist two positive constants k, K such that for $x \geqq 2$*

$$kx \leqq \theta(x) \leqq Kx . \tag{13.8}$$

Chebyshev's theorem. From the preceding theorem we now deduce the theorem about $\pi(x)$ by (Abel's) partial summation, a device which is very useful in the treatment of series and sums. We start with the identity

$$\pi(x) = \sum_{p \leq x} 1 = \sum_{2 \leq n \leq x} \frac{\theta(n) - \theta(n-1)}{\log n} .$$

The last sum needs some attention. Indeed, if $n = p$ is a prime number, then $\theta(n)$ exceeds $\theta(n-1)$ by $\log n$, so that a summand 1 appears. If, however, n is not a prime number, then $\theta(n)$ is equal to $\theta(n-1)$ and the summand zero appears, as it should. The above equation also involves $\theta(1) = 0$. The method of partial summation now calls for a rearrangement of the sum so that $\theta(n)$ is always collected from two consecutive terms, with the result

$$\pi(x) = \sum_{2 \leq n \leq x} \theta(n) \left(\frac{1}{\log n} - \frac{1}{\log (n+1)} \right) + \frac{\theta([x])}{\log ([x]+1)} , \tag{13.9}$$

where the last single summand is a compensation for a subtrahendus in excess in the sum. Now we have

$$\theta([x]) = \theta(x) .$$

Moreover

$$\left| \frac{1}{\log ([x]+1)} - \frac{1}{\log x} \right| = \left| \frac{-\log \dfrac{[x]+1}{x}}{\log x \cdot \log ([x]+1)} \right| < \frac{1}{x \log^2 x} .$$

In view of (13.8) we have then

$$\frac{\theta([x])}{\log ([x]+1)} = \frac{\theta(x)}{\log x} + O\left(\frac{1}{\log^2 x} \right) . \tag{13.10}$$

We shall see now that the sum in (13.9) is of lower order of magnitude than $\theta(x)/\log x$.
We have indeed

$$\sum_{2 \leq n \leq x} \theta(n) \left(\frac{1}{\log n} - \frac{1}{\log (n+1)} \right)$$

$$\leqq \sum_{2 \leq n \leq x} \theta(n) \frac{\log (1 + (1/n))}{(\log n)^2} < \sum_{2 \leq n \leq x} \frac{\theta(n)}{n} \cdot \frac{1}{(\log n)^2}$$

$$< K \sum_{2 \leq n \leq x} \frac{1}{(\log n)^2}$$

according to (13.8). Now

$$\sum_{2 \leq n \leq x} \frac{1}{(\log n)^2} = \sum_{2 \leq n \leq \sqrt{x}} \frac{1}{(\log n)^2} + \sum_{\sqrt{x} < n \leq x} \frac{1}{(\log n)^2}$$

$$\leqq \frac{\sqrt{x}}{(\log 2)^2} + \frac{1}{(\log \sqrt{x})^2} \sum_{n \leq x} 1 < 4\sqrt{x} + \frac{4x}{(\log x)^2} .$$

But

$$4\sqrt{x} + \frac{4x}{(\log x)^2} = O\left(\frac{x}{(\log x)^2}\right),$$

so that (13.9), (13.10) furnish

$$\pi(x) = \frac{\theta(x)}{\log x} + O\left(\frac{x}{(\log x)^2}\right).$$

In virtue of (13.8), the second summand on the right side divided by the first summand tends to zero if $x \to \infty$. Thus we may write

$$\pi(x) \sim \frac{\theta(x)}{\log x},$$

and from Theorem 48 we deduce Chebyshev's theorem:

THEOREM 49: *There exist two positive constants c, C such that for $x \geqq 2$*

$$c\frac{x}{\log x} \leqq \pi(x) \leqq C\frac{x}{\log x}.$$

Actually Chebyshev was able to give some fairly good numerical values for c and C. But because of the later developments in prime number theory these computations are no longer of great importance.

A further sum concerning primes. We conclude this chapter with an estimate of the sum

$$\sum_{p \leq x} \frac{1}{p},$$

firstly because we will need it later, and secondly because it will give us an opportunity to apply again Abel's method of partial summation. We set

$$G(x) = \sum_{p \leq x} \frac{\log p}{p}, \qquad G(1) = 0,$$
$$H(x) = G(x) - \log x.$$

According to Theorem 47 we have

$$H(x) = O(1).$$

Now we write, following a procedure similar to that in a previous paragraph,

$$\sum_{p \leq x} \frac{1}{p} = \sum_{2 \leq n \leq x} \frac{G(n) - G(n-1)}{\log n},$$

which is valid because $G(n) - G(n-1) = (\log n)/n$ when n is a prime, while $G(n) - G(n-1) = 0$ when n is not a prime. We obtain thus

$$\sum_{p \leq x} \frac{1}{p} = \sum_{2 \leq n \leq x} \frac{H(n) - H(n-1)}{\log n} + \sum_{2 \leq n \leq x} \frac{\log n - \log (n-1)}{\log n}.$$

For the first sum we see that

$$\begin{aligned}
&\left|\sum_{n=N}^{N+M} \frac{H(n) - H(n-1)}{\log n}\right| \\
&= \left|\sum_{n=N}^{N+M} H(n)\left(\frac{1}{\log n} - \frac{1}{\log (n+1)}\right) + \frac{H(N+M)}{\log (N+M+1)} - \frac{H(N-1)}{\log N}\right| \\
&\leqq \sum_{n=N}^{N+M} |H(n)| \left(\frac{1}{\log n} - \frac{1}{\log (n+1)}\right) + \left|\frac{H(N+M)}{\log (N+M+1)}\right| + \left|\frac{H(N-1)}{\log N}\right| \\
&< B\left\{\sum_{n=N}^{N+M}\left(\frac{1}{\log n} - \frac{1}{\log (n+1)}\right) + \frac{1}{\log (N+M+1)} + \frac{1}{\log N}\right\} \\
&< B\left\{\frac{1}{\log N} + \frac{1}{\log N}\right\} \\
&= \frac{2B}{\log N}.
\end{aligned} \tag{13.11}$$

Thus $\sum_{n=2}^{\infty} \frac{H(n) - H(n-1)}{\log n}$ converges. Let K be its sum. Then

$$\begin{aligned}
\sum_{n=2}^{x} \frac{H(n) - H(n-1)}{\log n} &= \sum_{n=2}^{\infty} \frac{H(n) - H(n-1)}{\log n} - \sum_{n=x+1}^{\infty} \frac{H(n) - H(n-1)}{\log n} \\
&= K + O\left(\frac{1}{\log x}\right),
\end{aligned}$$

according to the estimation (13.11). Thus we have

$$\begin{aligned}
\sum_{p \leq x} \frac{1}{p} &= \sum_{n=2}^{x} \frac{\log n - \log (n-1)}{\log n} + K + O\left(\frac{1}{\log x}\right) \\
&= -\sum_{n=2}^{x} \frac{\log (1 - (1/n))}{\log n} + K + O\left(\frac{1}{\log x}\right).
\end{aligned}$$

Now Taylor's theorem for the function $\log (1-x)$ shows

$$\log\left(1 - \frac{1}{n}\right) = -\frac{1}{n} - \frac{\theta_n}{n^2}, \qquad 0 < \theta_n < 1,$$

and therefore we obtain

$$\sum_{p \leq x} \frac{1}{p} = \sum_{n=2}^{x} \frac{1}{n \log n} + \sum_{n=2}^{x} \frac{\theta_n}{n^2 \log n} + K + O\left(\frac{1}{\log x}\right).$$

But

$$\sum_{n=2}^{\infty} \frac{1}{n^2 \log n}$$

converges and hence so does

$$\sum_{n=2}^{\infty} \frac{\theta_n}{n^2 \log n}.$$

Let K_1 be the sum of this last series. Then

$$\sum_{n=2}^{x} \frac{\theta_n}{n^2 \log n} = K_1 - \sum_{n=x+1}^{\infty} \frac{\theta_n}{n^2 \log n}$$
$$= K_1 + O\left(\frac{1}{x \log x}\right),$$

since

$$\sum_{n=x+1}^{\infty} \frac{\theta_n}{n^2 \log n} < \int_x^{\infty} \frac{dt}{t^2 \log t} < \frac{1}{\log x} \int_x^{\infty} \frac{dt}{t^2} = \frac{1}{x \log x}.$$

Thus

$$\sum_{p \le x} \frac{1}{p} = \sum_{n=2}^{x} \frac{1}{n \log n} + K + K_1 + O\left(\frac{1}{\log x}\right).$$

Finally, according to (12.3),

$$\sum_{n=2}^{x} \frac{1}{n \log n} = \log \log x + K_2 + O\left(\frac{1}{x \log x}\right)$$

when K_2 is a constant. This proves

THEOREM 50:

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

In particular, $\sum_p (1/p)$ diverges, where the sum is taken over all the primes. It would be quite easy to give a direct proof of this remark from Euler's identity (12.17). We shall see this and much more in the next chapter.

14

Primes in an Arithmetical Progression

Euler's proof of the infinity of primes. In 1837 Dirichlet proved his famous theorem on primes in arithmetic progressions: Every arithmetic progression $kn + a$, where a and k are relatively prime integers and n runs through all positive integers, contains an infinite number of primes. We have already seen examples of this theorem in the progressions $4n + 1$ and $4n + 3$ (Theorem 19 and page 2). The condition that a and k be relatively prime is clearly necessary, since any common factor of a and k would be a common factor of all the numbers $kn + a$.

Dirichlet used an idea of Euler, who proved the existence of an infinite number of primes in an essentially new way. Consider

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

We showed in a previous lecture (Chapter 12) that

$$\zeta(s) = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\right) = \prod_p \frac{1}{1 - (1/p^s)}.$$

If $s = 1 + \varepsilon$ with $\varepsilon > 0$, then the series for $\zeta(s)$ converges, but

$$\zeta(1 + \varepsilon) = \sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}} \to \infty \quad \text{as} \quad \varepsilon \to 0.$$

If there were only a finite number of primes $p_1, \cdots, p_n$, then we would have

$$\zeta(1 + \varepsilon) \to \left(\frac{1}{1 - (1/p_1)}\right)\left(\frac{1}{1 - (1/p_2)}\right) \cdots \left(\frac{1}{1 - (1/p_n)}\right),$$

which is a contradiction.

For the study of prime numbers modulo k, Dirichlet invented an expansion, in analogy to Euler's product,

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} = \prod_p \frac{1}{1 - (a_p/p^s)}$$

where the a_n are periodic modulo k and have the multiplicative property $a_{mn} = a_m a_n$. For this construction we need some preparations.

Finite Abelian groups and group characters. We have mentioned groups occasionally without making serious use of their properties in our deductions. We now find them essential.

The definition of a group is well known, but we repeat it here.

DEFINITION: *A group is a system Γ of elements $A, B, \cdots, M$ with a binary operation, which we shall write as multiplication.*
It has the properties:

I. *Closure!* The operation on two elements A, B of Γ produces an element of Γ: $A \cdot B = C$.

II. *Associativity:* The operation is associative:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C) .$$

III. *Unit element:* There exists an element, the unit element I, which leaves the other factor unaltered in operation:

$$A \cdot I = I \cdot A = A .$$

IV. *Inverse element:* For every element $A \in \Gamma$ there exists another element A', the inverse of A, such that

$$A \cdot A' = A' \cdot A = I .$$

If the group operation is commutative for all pairs of elements of Γ,

$$A \cdot B = B \cdot A ,$$

the group is called a *commutative* or *Abelian group*. A group may be *finite* or *infinite*, i.e., may have finitely or infinitely many elements. We shall deal here only with finite groups. The number of elements h of a finite group is called its *order*.

A subset $\Gamma^* \subset \Gamma$, which is itself a group, is called a *subgroup* of Γ. Every group has the trivial subgroup consisting of the unit element alone.

We need a few statements about groups which are easy to prove, and which we state as lemmas.

LEMMA 1: *If A is a fixed element of a group Γ and if X runs through all elements of the group, so does AX.*

This is a direct consequence of the existence of an inverse element A' of A.

LEMMA 2: *The order h^* of a subgroup Γ^* of Γ divides the order h of Γ.*

LEMMA 3: *If h is the order of Γ, then $A^h = I$ for any element A of Γ.*

This lemma follows from the previous one, since the set of all powers of A forms a subgroup. (These lemmas and their proofs are analogous to the facts about residues studied in Chapters 3 and 4.)

We now consider the system of residue classes modulo k which are relatively prime to k. These classes are $h = \varphi(k)$ in number and form a group. Postulates I and II are evidently fulfilled. The unit element in this case is the class which contains 1, i.e., the class of all integers $n \equiv 1 \pmod{k}$. Postulate IV is fulfilled, since for any a which is prime to k there exists an a' such that

$$aa' \equiv 1 \pmod{k} .$$

The class A of all numbers $n \equiv a \pmod{k}$ and the class A' of all numbers $n' \equiv a' \pmod{k}$ are then a pair of inverse elements in our group. The group is obviously Abelian.

A *character* of an Abelian group is a complex-valued function $\chi(A)$ of the element A of Γ, so that $\chi(A)$ is homomorphic to A. This means that if $AB = C$, then $\chi(A) \cdot \chi(B) = \chi(C)$. We assume that $\chi(A)$ is not zero. We express it a little more technically: A character of the Abelian group Γ is a homomorphic mapping of Γ into the group of nonzero complex numbers.

The values of χ can only be certain roots of unity. Indeed for the unit element I we have

$$\chi(I) \cdot \chi(I) = \chi(I^2) = \chi(I)$$

and thus, since χ is not zero, $\chi(I) = 1$. For any element A of Γ we have $A^h = I$, and therefore

$$(\chi(A))^h = \chi(A^h) = \chi(I) = 1 ,$$

so that $\chi(A)$ can only be an hth root of unity. One trivial character always exists, the principal character $\chi_0(A)$, which is 1 for any A of Γ.

In the case which interests us here, in which the group Γ is the congruence residue group modulo k of order $\varphi(k)$, it is convenient to write not the classes but the numbers of the classes, as arguments of χ, so that

$$\chi(n) = \chi(A)$$

if n is an element of the class A. This has as consequence

$$\chi(n_1) = \chi(n_2) \qquad \text{for } n_1 \equiv n_2 \pmod{k} , \tag{14.1}$$

and also

$$\chi(m) \cdot \chi(n) = \chi(mn) . \tag{14.2}$$

We complete the definition of χ as an arithmetical function by setting $\chi(n) = 0$ if n does not belong to any class, i.e., if n is *not* relatively prime to

k. Then (14.1) and (14.2) remain true under this extension of the definition of $\chi(n)$.

The principal character $\chi_0(n)$ is now defined as

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

The crux of the matter is to show the existence of nonprincipal characters, if $h > 1$. We shall not prove this in general for any Abelian group, since that would involve us too deeply in the theory of Abelian groups, but only for our case of the group of congruence residue classes.

Let us first consider some examples of characters. For $k = 2$ there is only one class, $h = 1$, and thus only the principal character $\chi_0(n) = 1$ for n odd, $\chi_0(n) = 0$ for n even.

For $k = 3$ there are 2 characters:

$$\chi_0(n) = \begin{cases} 1 & n \equiv 1, 2 \pmod 3 \\ 0 & n \equiv 0 \pmod 3 , \end{cases} \qquad \chi_1(n) = \begin{cases} 1 & n \equiv 1 \pmod 3 \\ -1 & n \equiv -1 \pmod 3 \\ 0 & n \equiv 0 \pmod 3 . \end{cases}$$

For $k = 4$ there are again 2 characters:

$$\chi_0(n) = \begin{cases} 1 & n \text{ odd} \\ 0 & n \text{ even} , \end{cases} \qquad \chi_1(n) = \begin{cases} 1 & \text{for } n \equiv 1 \pmod 4 \\ -1 & \text{for } n \equiv -1 \pmod 4 \\ 0 & \text{for } n \text{ even} . \end{cases}$$

For $k = 5$ there are 4 characters given in the following table:

$n \equiv$	1	2	3	4	0	(mod 5)
$\chi_0(n)$	1	1	1	1	0	
$\chi_1(n)$	1	i	$-i$	-1	0	
$\chi_2(n)$	1	-1	-1	1	0	
$\chi_3(n)$	1	$-i$	i	-1	0	

We shall see presently that in all these cases we have given complete lists of the possible characters.

Theorems about group characters. Let us take the existence of nonprincipal characters for granted for the moment; it will be proved in Theorem 53.

THEOREM 51: *If χ is not the principal character, then*

$$\sum_{n=1}^{k} \chi(n) = 0 . \tag{14.3}$$

For the principal character χ_0 we obviously have

$$\sum_{n=1}^{k} \chi_0(n) = \varphi(k) . \tag{14.4}$$

Proof: Since χ is not the principal character, there exists an integer a prime to k with $\chi(a) \neq 1$. Then

$$\sum_{n=1}^{k} \chi(n) = \sum_{n=1}^{k} \chi(an) ,$$

since an runs through all residue classes modulo k if n does. But we have $\chi(an) = \chi(a)\chi(n)$, and thus

$$\sum_{n=1}^{k} \chi(n) = \chi(a) \sum_{n=1}^{k} \chi(n) , \qquad (1 - \chi(a)) \sum_{n=1}^{k} \chi(n) = 0 ,$$

which proves (14.3), since $1 - \chi(a) \neq 0$.

THEOREM 52: *The characters modulo k form a finite Abelian group.*

Proof: They form a finite system because they can take only $\varphi(k)$th roots of unity as values. Thus there cannot be more than h^h characters, $h = \varphi(k)$. The product of two characters is a character, since it fulfills (14.1) and (14.2). That accounts for closure of the system. Associativity is inherent in multiplication. The identity of the group is the principal character χ_0:

$$\chi_0(n) \cdot \chi(n) = \chi(n) ,$$

and the inverse of a given character χ is its complex conjugate $\bar{\chi}$:

$$\chi(n)\bar{\chi}(n) = \chi_0(n) .$$

The proofs of these two theorems remain valid for any finite Abelian group. For the proof of existence of nonprincipal characters we shall appeal, however, to the specific nature of the group of residue classes modulo k. We need for this the following remark.

Remark: If d divides k and if χ is known as a character modulo d, then we may construct a character χ^* modulo k by setting

$$\chi^*(n) = \chi(n) \qquad \text{if } (n, k) = 1$$
$$\chi^*(n) = 0 \qquad \text{if } (n, k) \neq 1 .$$

It is simple to test that the χ^* so defined satisfies the postulates (14.1) and (14.2). We shall refer to this process as the *extension* of the character χ modulo d to the modulus k.

We now come to the existence of nonprincipal characters.

THEOREM 53: *If* $(a, k) = 1$ *and* $a \not\equiv 1 \pmod{k}$, *then there exists a character* χ *modulo* k *such that* $\chi(a) \neq 1$.

Proof: Let $k = 2^\alpha p_1^{\beta_1} \cdots p_r^{\beta_r}$ be the decomposition of k into prime factors where the p_j are distinct odd primes. In view of the hypothesis about a, not all the congruences

$$a \equiv 1 \pmod{2^\alpha}$$

$$a \equiv 1 \pmod{p_j^{\beta_j}}, \qquad j = 1, 2, \cdots, r$$

can be fulfilled. First suppose that

$$a \not\equiv 1 \pmod{p^\beta}$$

where p is odd. Let g be a primitive congruence root modulo p^β (see Theorem 28). The powers g^λ, $\lambda = 1, 2, \cdots, \varphi(p^\beta)$, represent all residue classes modulo p^β which are prime to p. We define $\chi(g) = e^{2\pi i/\varphi(p^\beta)}$. Since then

$$\chi(g^\lambda) = (\chi(g))^\lambda = e^{2\pi i\lambda/\varphi(p^\beta)},$$

we have defined a character modulo p^β. With a certain $\mu \not\equiv 0 \pmod{\varphi(p^\beta)}$, we have

$$g^\mu \equiv a \pmod{p^\beta}, \qquad 0 < \mu < \varphi(p^\beta),$$

and thus

$$\chi(a) = e^{2\pi i\mu/\varphi(p^\beta)} \neq 1.$$

Since p^β divides k, this character χ can be extended to a character χ^* modulo k. In this process $\chi(a) = \chi^*(a)$, since $(a, k) = 1$. In this case we have the desired character modulo k.

Second, suppose that $a \not\equiv 1 \pmod{2^\alpha}$. The case $\alpha = 1$ does not occur, for in this case a and k are even, and therefore $(a, k) \neq 1$. Hence $(a, k) = 1$ implies that a is odd, and this would mean $a \equiv 1 \pmod 2$. For $\alpha = 2$, $a \not\equiv 1 \pmod 4$, we must have $a \equiv -1 \pmod 4$. In our table we have given a nonprincipal character χ_1 modulo 4 for which $\chi_1(a) = -1$. We can again extend this character to a character χ^* modulo k.

Now suppose $\alpha \geqq 3$. There is no primitive congruence root modulo 2^α. But 5^λ, $\lambda = 1, \cdots, 2^{\alpha-2}$ represents all numbers of the form $4n + 1$ modulo 2^α, since $5 \equiv 1 \pmod 4$. Thus

$$5^\lambda \equiv 1 \pmod 4$$

on the one hand, and since on the other hand the smallest positive μ with

$$5^\mu \equiv 1 \pmod{2^\alpha}$$

is $\mu = 2^{\alpha-2}$, $\alpha \geqq 3$, as can be seen by induction on α. The numbers $4n - 1$ modulo 2^α can then be represented by -5^λ, so that all odd residue classes modulo 2^α are represented by $\pm 5^\lambda$, $\lambda = 1, 2, \cdots, 2^{\alpha-2}$.

If $a \equiv -1 \pmod{2^\alpha}$, then $a \equiv -1 \pmod 4$ and we proceed as in the previous case with $\chi_1(a) = -1$, where χ_1 can now also be extended to a character modulo k, since $4 \mid k$. Thus we have only to consider the case $a \not\equiv \pm 1 \pmod{2^\alpha}$. Then there exists a $v \not\equiv 0 \pmod{2^{\alpha-2}}$ so that

$$a \equiv \pm 5^v \pmod{2^\alpha}, \qquad 0 < v < 2^{\alpha-2}.$$

We now define a character χ modulo 2^α as follows:

$$\chi(-1) = 1, \qquad \chi(5) = e^{2\pi i/2^{\alpha-2}}.$$

Then

$$\chi(a) = \chi(\pm 5^v) = (\chi(5))^v = e^{2\pi iv/2^{\alpha-2}} \neq 1.$$

If we again extend this character χ modulo 2^α to a character χ^* modulo k, we have also settled this case and the theorem is proved.

THEOREM 54: *If* $a \not\equiv 1 \pmod{k}$, *then* $\sum_\chi \chi(a) = 0$, *where the sum is extended over all characters modulo* k.

Proof: Let χ^* be a character modulo k with $\chi^*(a) \neq 1$; the existence of such a character has been proved in the foregoing theorem. We remember that the characters form a finite group: if χ runs through all characters, so does $\chi^*\chi$. Therefore,

$$\sum_\chi \chi(a) = \sum_\chi \chi^*\chi(a) = \chi^*(a) \sum_\chi \chi(a),$$

and

$$(1 - \chi^*(a)) \sum_\chi \chi(a) = 0.$$

The result follows, since $\chi^*(a) \neq 1$.

THEOREM 55: *There are* $\varphi(k)$ *distinct characters modulo* k.

Proof: By rearranging the summands of the double sums we get the equation

$$\sum_\chi \left\{ \sum_{n=1}^{k} \chi(n) \right\} = \sum_{n=1}^{k} \left\{ \sum_\chi \chi(n) \right\}. \tag{14.5}$$

The inner sum on the left is

$$\sum_{n=1}^{k} \chi(n) = \begin{cases} \varphi(k) & \chi = \chi_0 \\ 0 & \chi \neq \chi_0 \end{cases}.$$

The inner sum on the right side, by Theorem 54, is zero for $n \not\equiv 1 \pmod k$ and is clearly the number c of characters modulo k for $n \equiv 1 \pmod k$, since then each summand contributes 1.

Equation (14.5) therefore reduces to

$$\varphi(k) = c,$$

which had to be proved.

This theorem shows that our tables for characters modulo k up to $k = 5$ are complete, since indeed each table exhibits $\varphi(k)$ characters.

A corollary of Theorem 55 is the equation, for $(a, k) = 1$,

$$\sum_{\chi} \bar{\chi}(a)\chi(n) = \begin{cases} \varphi(k) & \text{if } a \equiv n \pmod{k} \\ 0 & \text{otherwise .} \end{cases} \tag{14.6}$$

Indeed $\bar{\chi}(a)$ is the reciprocal of $\chi(a)$:

$$\bar{\chi}(a) = \chi(a)^{-1} = \chi(a^*) ,$$

where $aa^* \equiv 1 \pmod{k}$. Thus the sum in (14.6) reduces to

$$\sum_{\chi} \chi(a^*n) ,$$

and this is taken care of by Theorems 54 and 55 .

THEOREM 56: *For $(a, k) = 1$ let f be the smallest positive exponent such that $a^f \equiv 1 \pmod{k}$. Then $\chi(a)$ is an fth root of unity, and all fth roots of unity appear equally often as $\chi(a)$ if χ runs over the characters modulo k.*

Proof: Since $\chi(a)^f = \chi(a^f) = 1$, the first assertion about $\chi(a)$ is clear. If, moreover, $a \equiv 1 \pmod{k}$, we have $f = 1$, $\chi(a) = 1$ for all χ, and there is nothing to prove. Therefore, let us take $a \not\equiv 1 \pmod{k}$ so that $f > 1$. Let ε be a fixed fth root of unity. We want to find out how often $\chi(a) = \varepsilon$ among the χ, and to show that this frequency does not depend on ε. We consider the following sum:

$$S = \sum_{\chi} \{\varepsilon^{-1}\chi(a) + \varepsilon^{-2}\chi(a^2) + \cdots + \varepsilon^{-f}\chi(a^f)\} ,$$

which we write in two ways:

$$\sum_{\chi} \sum_{l=1}^{f} (\varepsilon^{-1}\chi(a))^l = \sum_{l=1}^{f} \varepsilon^{-l} \sum_{\chi} \chi(a^l) .$$

The inner sum on the right-hand side is 0 for $a^l \not\equiv 1 \pmod{k}$, that is, for $l = 1, \cdots, f-1$, and is $\varphi(k)$ for $l = f$. Thus we have $S = \varphi(k)$. In the inner sum on the left we set

$$\eta = \varepsilon^{-1}\chi(a) ,$$

which is a certain fth root of unity. Hence the sum becomes

$$\sum_{l=1}^{f} \eta^l = \begin{cases} f & \text{for } \eta = 1 \\ 0 & \text{for } \eta \neq 1 . \end{cases}$$

This means $S = e \cdot f$, if e gives the number of times that $\eta = 1$ or that $\chi(a) = \varepsilon$. We have, consequently,

$$e = \frac{\varphi(k)}{f} , \tag{14.7}$$

independent indeed of the particular ε.

The Dirichlet series. We now return to Dirichlet's problem and the construction of a series indicated in the beginning of this chapter. These series are

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

for $s > 1$. The convergence for $s > 1$ follows at once, since $L(s, \chi)$ is majorized termwise by $\zeta(s) = \sum 1/n^s$. The Euler product now appears here immediately as a consequence of the character property of χ and the uniqueness of prime factorization:

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \left(1 + \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{p^{2s}} + \cdots\right) = \prod_p \frac{1}{1 - (\chi(p)/p^s)} . \tag{14.8}$$

For the principal character χ_0, we have in particular

$$L(s, \chi_0) = \sum_{\substack{n=1 \\ (n,k)=1}}^{\infty} \frac{1}{n^s} = \prod_{p \nmid k} \frac{1}{1 - (1/p^s)} = \prod_p \frac{1}{1 - (1/p^s)} \cdot \prod_{p \mid k} \left(1 - \frac{1}{p^s}\right) ,$$

$$L(s, \chi_0) = \zeta(s) \cdot \prod_{p \mid k} \left(1 - \frac{1}{p^s}\right) . \tag{14.9}$$

If we take the logarithms in (14.8), we have

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

$$= \sum_p \sum_{m=1}^{\infty} \frac{1}{m} \frac{\chi(p)^m}{p^{ms}}$$

$$= \sum_p \frac{\chi(p)}{p^s} + \sum_{m=2}^{\infty} \frac{1}{m} \sum_p \frac{\chi(p^m)}{p^{ms}} .$$

If now a is a number relatively prime to k, (14.6) furnishes the fundamental formula

$$\frac{1}{\varphi(k)} \sum_{\chi} \bar{\chi}(a) \log L(s, \chi) = \sum_{p \equiv a \pmod{k}} \frac{1}{p^s} + H_a(s) , \tag{14.10}$$

where

$$H_a(s) = \frac{1}{\varphi(k)} \sum_{\chi} \bar{\chi}(a) \sum_{m=2}^{\infty} \frac{1}{m} \sum_p \frac{\chi(p^m)}{p^{ms}} ,$$

so that

$$|H(s)| \leqq \sum_{m=2}^{\infty} \frac{1}{m} \sum_p \frac{1}{p^{ms}} < \sum_{m=2}^{\infty} \sum_{n=2}^{\infty} \frac{1}{n^{ms}}$$

$$= \sum_{n=2}^{\infty} \sum_{m=2}^{\infty} \frac{1}{n^{ms}} = \sum_{n=2}^{\infty} \frac{1}{n^{2s}} \cdot \frac{1}{1 - (1/n^s)}$$

$$< 2 \sum_{n=2}^{\infty} \frac{1}{n^{2s}} < 2 \sum_{n=1}^{\infty} \frac{1}{n^2} = C < \infty$$

for all $s > 1$. (The constant C happens to be $\pi^2/3$, which is irrelevant here.) In order to prove Dirichlet's theorem, we want to show that

$$\sum_{p \equiv a \pmod k} \frac{1}{p^s} \to \infty$$

as $s \to 1$. Since $H(s)$ remains bounded, this will be accomplished if it is proved that

$$\sum_{\chi} \bar{\chi}(a) \log L(s, \chi) \to \infty \tag{14.11}$$

if $s \to 1$. This is now the remaining problem. One of the summands, namely that for χ_0, does go to infinity as $s \to 1$. In fact, we have

$$\sum_{n=2}^{\infty} \frac{1}{n^s} < \int_1^{\infty} \frac{dx}{x^s} < \sum_{n=1}^{\infty} \frac{1}{n^s}$$

or

$$\zeta(s) - 1 < \frac{1}{s-1} < \zeta(s) .$$

Hence, for $s > 1$,

$$1 < (s-1)\zeta(s) < s,$$

and consequently

$$(s-1)\zeta(s) \to 1 \quad \text{as } s \to 1 .$$

Since $(s-1)$ tends to zero, the factor $\zeta(s)$ must tend to infinity. In (14.9) we see

$$\lim_{s \to 1} \prod_{p \mid k} \left(1 - \frac{1}{p^s}\right) = \prod_{p \mid k} \left(1 - \frac{1}{p}\right) = \frac{\varphi(k)}{k},$$

and thus we find

$$(s-1)L(s, \chi_0) = (s-1)\zeta(s) \prod_{p \mid k} \left(1 - \frac{1}{p^s}\right) \to \frac{\varphi(k)}{k} \quad \text{if } s \to 1 . \tag{14.12}$$

This shows that $\chi_0(a) \log L(s, \chi_0) \to \infty$ if $s \to 1$. In order to prove, however, that the whole sum (14.11) goes to infinity, it is then necessary that no mutual compensation of the terms in that sum might render the limit finite. It will suffice that no other term of the sum goes to ∞. But in order to see that $|\log L(s, \chi)|$ does not go to ∞, it is necessary to show that $L(s, \chi)$ goes *neither to* ∞ *nor to* 0.

The continuity of $L(s, \chi)$ at $s = 1$ for $\chi \neq \chi_0$. The first part of this task is fairly simple. First, $L(s, \chi) \neq 0$ for $s > 1$, because (14.8) shows $L(s, \chi)$ as an absolutely convergent product of which no factor vanishes. We shall now prove that $L(s, \chi)$, $\chi \neq \chi_0$, is a continuous and differentiable function for $s > 0$. For this purpose we again use Abel's partial summation.

We introduce the auxiliary sum

$$s(m) = \sum_{n=1}^{m} \chi(n) .$$

We have

$$s(m) = \sum_{n=1}^{k[m/k]} \chi(n) + \sum_{n=k[m/k]+1}^{m} \chi(n) ,$$

and since

$$\sum_{n=1}^{k} \chi(n) = 0 ,$$

only the second sum counts. But that sum contains less than k terms. Thus we infer

$$|s(m)| < k .$$

Then for $s > 1$, $M < N$, we have

$$\begin{aligned} S(M, N) &= \sum_{n=M+1}^{N} \frac{\chi(n)}{n^s} = \sum_{n=M+1}^{N} \frac{s(n) - s(n-1)}{n^s} \\ &= \sum_{n=M+1}^{N-1} s(n)\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) + \frac{s(N)}{(N+1)^s} - \frac{s(M)}{(M+1)^s} , \end{aligned}$$

so that

$$\begin{aligned} |S(M, N)| &\leq k \sum_{n=M+1}^{N} \left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) + \frac{k}{(N+1)^s} + \frac{k}{(M+1)^s} \\ &= \frac{2k}{(M+1)^s} . \end{aligned} \tag{14.13}$$

Cauchy's convergence criterion ensures therefore the convergence of the series for $L(s, \chi)$ for $s > 0$, and *uniform* convergence for all $s \geqq \delta > 0$ for a fixed δ. The formal derivative of $L(s, \chi)$ will be

$$-\sum_{n=1}^{\infty} \frac{\chi(n) \log n}{n^s} , \tag{14.14}$$

and the same argument will show uniform convergence of (14.14) for $s \geqq \delta > 0$. Therefore $L(s, \chi)$ is differentiable for $s > 0$ and has (14.14) as its derivative.

In particular, $L(s, \chi)$ is continuous for $s > 0$ and $L(1, \chi)$ is thus *finite*. If $L(s, \chi)$ should *vanish* at $s = 1$, we would have

$$\frac{L(s, \chi) - L(1, \chi)}{s-1} = \frac{L(s, \chi)}{s-1} \to L'(1, \chi)$$

with $s \to 1$ or

$$L(s, \chi) = (s-1)(L'(1, \chi) + \eta(s)) \tag{14.15}$$

with

$$\lim_{s \to 1} \eta(s) = 0 .$$

The nonvanishing of $L(1, \chi)$, $\chi \neq \chi_0$, first step. We now show that $L(1, \chi) \neq 0$ for every nonprincipal character. This we do in two steps.

We form the product of all $L(s, \chi)$:

$$F(s) = \prod_{\chi} L(s, \chi) = \prod_{p \nmid k} \prod_{\chi} \frac{1}{1 - (\chi(p)/p^s)}, \qquad s > 1.$$

We now apply Theorem 56 to $\chi(p)$. If f is the smallest positive integer so that $p^f \equiv 1 \pmod{k}$, then $\chi(p)$ is an fth root of unity, let us say ε, and all such ε occur with the multiplicity $e = (\varphi(k)/f)$ if χ runs through all the characters modulo k. That means

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

where ε runs over all fth roots of unity.

Now since

$$\prod_{\varepsilon} (x - \varepsilon) = x^f - 1,$$

$$\prod_{\varepsilon}\left(1 - \frac{\varepsilon}{x}\right) = 1 - \frac{1}{x^f},$$

we have

$$\prod_{\varepsilon}\left(1 - \frac{\varepsilon}{p^s}\right) = 1 - \frac{1}{p^{fs}},$$

and so

$$\prod_{\chi}\left(1 - \frac{\chi(p)}{p^s}\right) = \left(1 - \frac{1}{p^{fs}}\right)^e \leqq 1 - \frac{1}{p^{efs}}.$$

If we set $h = \varphi(k) = ef$, we have thus

$$F(s) = \prod_{\chi} L(s, \chi) \geqq \prod_{p \nmid k} \frac{1}{1 - (1/p^{hs})} = \zeta(hs) \cdot \prod_{p \mid k}\left(1 - \frac{1}{p^{hs}}\right),$$

and for $s > 1$,

$$F(s) = \prod_{\chi} L(s, \chi) > \zeta(hs) \prod_{p \mid k}\left(1 - \frac{1}{p}\right) > \frac{\varphi(k)}{k}. \tag{14.16}$$

This fact already precludes the vanishing of more than one of the $L(1, \chi)$. Indeed, assume that $L(1, \chi_1) = L(1, \chi_2) = 0$. (We always reserve the notation χ_0 for the principal character.) Then $F(s)$ would contain, besides other factors that are continuous at $s = 1$,

$$\begin{aligned} &L(s, \chi_0) \cdot L(s, \chi_1) \cdot L(s, \chi_2) \\ &\quad = L(s, \chi_0) \cdot (s - 1)(L'(1, \chi_1) + \eta_1(s)) \\ &\qquad \cdot (s - 1)(L'(1, \chi_2) + \eta_2(s)) \quad \text{by (14.15).} \end{aligned}$$

But since $(s - 1)L(s, \chi_0) \to (\varphi(k)/k)$ by (14.12), the second factor $(s - 1)$ would let this product go to 0 in contradiction to (14.16).

If now for a complex character χ (i.e., a character which assumes complex values) we would have

$$L(1, \chi) = 0,$$

then $L(s, \chi)$ and $L(s, \bar{\chi})$ would on the one hand be two different functions, but, since for real $s > 1$

$$L(s, \bar{\chi}) = \sum_n \frac{\bar{\chi}(n)}{n^s} = \overline{L(s, \chi)},$$

we would also have $L(1, \bar{\chi}) = 0$, which we just excluded. Therefore, the problem is further reduced; it remains to be shown that $L(1, \chi) \neq 0$ for all *real* nonprincipal characters.

Now if there were just *one* $L(1, \chi) = 0$, we would have in $F(s)$, besides continuous factors, the product

$$\begin{aligned} &L(s, \chi_0)L(s, \chi) \\ &\quad = L(s, \chi_0)(s - 1)(L'(1, \chi) + \eta(s)) \to \frac{\varphi(k)}{k} L'(1, \chi) \quad \text{for } s \to 1. \end{aligned}$$

This would imply that $F(s)$, for $s \to 1$ would have a finite limit. If we could show that this is not so, but actually $F(s) \to \infty$ for $s \to 1$, then the possibility of a single vanishing $L(1, \chi)$ would also be ruled out. But this apparently simple plan leads to complications. We take another road.

The nonvanishing of $L(1, \chi)$, $\chi \neq \chi_0$, second step. There are a number of proofs for the rather famous problem concerning $L(1, \chi) \neq 0$. Only real χ, that is $\chi(n) = \pm 1$ for $(n, k) = 1$, have to be considered. Dirichlet solved the problem by reducing it to another one. He had shown that $L(1, \chi)$ for a real character χ has a meaning in the theory of quadratic forms and represents a number which by its definition must be positive. We proceed here directly, following a proof of Mertens (1897).

Let us put

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

It is immediately seen that $f(n_1 n_2) = f(n_1)f(n_2)$ whenever $(n_1, n_2) = 1$. Since

$$\begin{aligned} f(p^l) &= \chi(1) + \chi(p) + \cdots + \chi(p^l) \\ &= 1 + \chi(p) + \chi(p)^2 + \cdots + \chi(p)^l, \end{aligned}$$

and $\chi(p) = \pm 1$, it follows that $f(p^l) \geqq 0$, and hence

$$f(n) \geqq 0$$

for all natural numbers n. If l is *even*, then the above equation shows $f(p^l)$ either equal to $l + 1$ or to 1, and in any case $\geqq 1$. From this and the multiplicativity, it follows that

$$f(m^2) \geqq 1.$$

Since now

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^{\frac{1}{2}}} \geqq \sum_{m=1}^{\infty} \frac{f(m^2)}{m} \geqq \sum_{m=1}^{\infty} \frac{1}{m},$$

it follows that

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^{\frac{1}{2}}}$$

diverges. Let us investigate this divergence a little more closely. We set

$$G(x) = \sum_{n=1}^{x} \frac{f(n)}{n^{\frac{1}{2}}} = \sum_{n=1}^{x} \frac{1}{n^{\frac{1}{2}}} \sum_{d|n} \chi(d) = \sum_{t \cdot d \leq x} \frac{\chi(d)}{(td)^{\frac{1}{2}}}.$$

In this last sum we are summing over all lattice points under the hyperbola

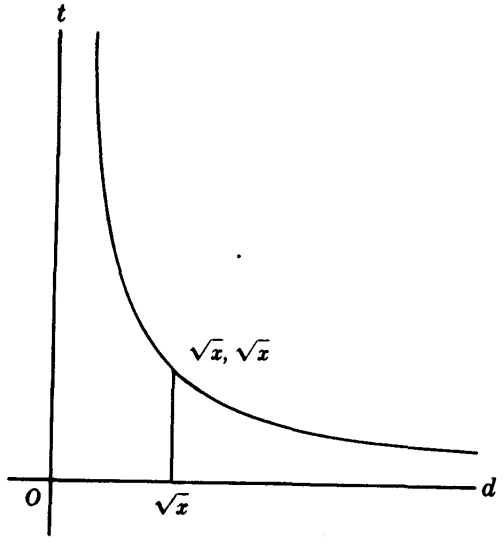

Figure 7

$t \cdot d = x$ in a (d, t) plane. We break the area under the hyperbola into two pieces by the ordinate erected at $\sqrt{x}$. See Fig. 7. Then we have

$$G(x) = \sum_{d=1}^{\sqrt{x}} \frac{\chi(d)}{d^{\frac{1}{2}}} \sum_{t=1}^{x/d} \frac{1}{t^{\frac{1}{2}}} + \sum_{t=1}^{\sqrt{x}} \frac{1}{t^{\frac{1}{2}}} \sum_{d=\sqrt{x}+1}^{x/t} \frac{\chi(d)}{d^{\frac{1}{2}}}.$$

Now (12.1) of the Lemma in Chapter 12 shows that

$$\sum_{t=1}^{z} \frac{1}{t^{\frac{1}{2}}} = \int_1^z \frac{du}{u^{\frac{1}{2}}} + C' + O(x^{-\frac{1}{2}})$$

$$= 2\sqrt{x} + C + O(x^{-\frac{1}{2}}), \qquad (14.17)$$

so that

$$G(x) = \sum_{d=1}^{\sqrt{x}} \frac{\chi(d)}{d^{\frac{1}{2}}} \left(2\sqrt{\frac{x}{d}} + C\right) + O\left(x^{-\frac{1}{2}} \sum_{d=1}^{\sqrt{x}} \frac{|\chi(d)|}{d^{\frac{1}{2}}}\right) + O\left(\sum_{t=1}^{\sqrt{x}} \frac{1}{t^{\frac{1}{2}}} \cdot \frac{1}{x^{\frac{1}{2}}}\right),$$

where the last term comes from an application of (14.13) to

$$\sum_{d=\sqrt{x}+1}^{x/t} \frac{\chi(d)}{d^{\frac{1}{2}}}.$$

Continuing, we have

$$G(x) = 2\sqrt{x} \sum_{d=1}^{\sqrt{x}} \frac{\chi(d)}{d} + C \cdot O(1) + O(x^{-\frac{1}{2}}) + O(1),$$

where we have twice employed (14.17) and the fact that

$$\sum_{d} \frac{\chi(d)}{d^{\frac{1}{2}}}$$

remains bounded because of convergence. Thus, with application of (14.13) again,

$$\begin{aligned} G(x) &= 2\sqrt{x}\left\{\sum_{d=1}^{\infty} \frac{\chi(d)}{d} - \sum_{d=\sqrt{x}+1}^{\infty} \frac{\chi(d)}{d}\right\} + O(1) \\ &= 2\sqrt{x}\, L(1, \chi) + 2\sqrt{x}\, O\left(\frac{1}{\sqrt{x}}\right) + O(1) \\ &= 2\sqrt{x}\, L(1, \chi) + O(1). \end{aligned}$$

We know that $G(x) \to \infty$ if $x \to \infty$. But this evidently is possible only if

$$L(1, x) \neq 0,$$

which we had to prove.

Altogether, we have now proved (14.11) for $(a, k) = 1$.

Since this implies, as we have seen,

$$\sum_{p \equiv a \pmod{k}} \frac{1}{p^s} \to \infty \quad \text{if } s \to 1, \qquad (14.18)$$

we have proved Dirichlet's famous theorem.

THEOREM 57: *The arithmetic progression $a + kn$, $n = 1, 2, 3, \cdots$, contains infinitely many prime numbers if a and k are coprime.*

Remarks. Two further remarks may be in place.

I. All that was needed, as we have observed above, was to show that in (14.16) $|F(s)| \to \infty$ as $s \to 1$. Now (14.16) shows that $F(s)$ cannot tend to 0 if $s \to 1$, and thus it will suffice to show $\log F(s) \to \infty$ if $s \to 1$. But

$$\log F(s) = \sum_{\chi} \log L(s, \chi),$$

which in comparison with (14.10), putting $a = 1$, means

$$\log F(s) = \varphi(k)\left(\sum_{p \equiv a \pmod k} \frac{1}{p^s} + H_1(s)\right).$$

Therefore, if one could prove that

$$\sum_{p \equiv 1 \pmod k} \frac{1}{p^s} \to \infty,$$

then (14.18) would follow for any a which is prime to k, and thus Dirichlet's theorem would follow immediately.

II. Our reasoning contains one difficulty which we have glossed over, the meaning of the expression $\log L(s, \chi)$, where χ and therefore $L(s, \chi)$ are complex numbers.

The logarithm of a complex number

$$z = |z| \cdot e^{i\varphi} \tag{14.19}$$

is defined as

$$\log z = \log |z| + i\varphi,$$

where $\log |z|$ is the usual natural logarithm of a positive number. But the argument φ in (14.19) is not uniquely determined and leaves additive multiples of 2π free. Now, for real s,

$$L(s, \chi) = 1 + \frac{\chi(2)}{2^s} + \frac{\chi(3)}{3^s} + \cdots \to 1$$

as $s \to \infty$. As we have seen, $L(s, \chi)$ varies continuously for $s \geqq 1$, without becoming zero. Thus $\log L(s, \chi)$ is meant as *that continuous function* which goes to $\log 1 = 0$ if $s \to \infty$. This fixation eliminates doubtful additions of multiples of $2\pi i$ to $\log L(s, \chi)$, and with this definition of the logarithm all previous arguments are valid.

15

The Sieve of Eratosthenes and a Theorem of V. Brun

The sieve of Eratosthenes. A method to detect the prime numbers in the sequence of all natural numbers was found by Eratosthenes (third century B.C.). It utilizes the fact that a proper divisor of a number must precede it but cannot be 1, the unit. Hence 2 is the first prime number. After 2 no even number can be a prime number, so all even numbers can be stricken out ("sieved" out). The first remaining number is 3; and since it is not a multiple of a preceding number, it must be a prime. Again, all multiples of 3 can be dismissed and are stricken out.

1 2 3 ~~4~~ 5 ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11 ~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19 ~~20~~ ~~21~~ ~~22~~ 23 ~~24~~ ~~25~~.

This leaves 5 as the next surviving number, which thus must be a prime. Now all multiples of 5 are sieved out, and so on. This procedure is known as the sieve of Eratosthenes.

Actually this process achieves a good deal more. If a number n is composite, at least one factor must obviously be $\leqq \sqrt{n}$. If in our list above, extending to 25, a number is composite, it must have at least one factor $\leqq 5$, and thus be a multiple of 2 or 3 or 5. When these multiples have been deleted, the remaining numbers up to 25 must all be primes: 7, 11, 13, 17, 19, 23. This goes on, in general; the primes between $\sqrt{n}$ and n are obtained by deleting the multiples of all primes up to $\sqrt{n}$.

A direct application of the sieve of Eratosthenes to the problem of the distribution of primes has not been successful so far. However, since 1919 the Norwegian mathematician Viggo Brun and many followers have used a sieve method to study certain problems connected with primes. We have already mentioned in the Introduction the occurrence of twin primes, i.e., pairs of prime numbers which differ by 2. They are relatively infrequent among all primes. Whereas we know that the sum of the reciprocals of all primes

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \cdots = \sum_p \frac{1}{p}$$

is divergent (see Chapter 13, Theorem 50), we shall prove in this chapter the theorem of Viggo Brun.

THEOREM 58: *The sum*

$$\tfrac{1}{3}+\tfrac{1}{5}+\tfrac{1}{7}+\tfrac{1}{11}+\tfrac{1}{13}+\tfrac{1}{17}+\tfrac{1}{19}+\tfrac{1}{29}+\tfrac{1}{31}+\cdots$$

of the reciprocals of twin primes is convergent (possibly containing only finitely many terms).

Remark: The prime 5 belongs to the two twin pairs 3,5 and 5,7. No other primes can have this property!

First step of V. Brun's method. Viggo Brun applied a *double sieving* to the sequence of natural numbers, so that all those numbers n were stricken out for which n or $n+2$ are composite. Only those n's were retained which are a first member of a pair of twin primes. In fact, as we shall see, this program cannot be carried out in full rigor; some more numbers will escape the sieve, but not enough to disturb the convergence of the series in Theorem 58.

Let $T(x)$ be the number of the first members n of pairs of twin primes for which $n \leqq x$. Moreover, let $U(x; p_1, p_2, \cdots, p_r)$ be the number of odd numbers $n \leqq x$ for which $n(n+2)$ is not divisible by any of the odd primes $p_1, p_2, \cdots, p_r$. If we take as this set all primes $p_j \leqq \sqrt{x+2}$, then $U(x; p_1, \cdots, p_r)$ counts only the first members of twin primes. Since some twin primes may be among the primes $p_1, p_2, \cdots, p_r$, we can state

$$T(x) \leqq r + U(x; p_1, \cdots, p_r)\,. \tag{15.1}$$

If, however, we take, for the sieving process, only odd primes $p_j \leqq y < \sqrt{x+2}$, the previous inequality will remain correct *a fortiori*, since on the right-hand side we may also count some odd numbers which are not first members of twin primes. Let us use the abbreviation

$$U(x; p_1, p_2, \cdots, p_r) = U(x; y)\,,$$

if $p_1, p_2, \cdots, p_r$ are the odd primes $p_j \leqq y$ for any $y \leqq \sqrt{x+2}$:

$$T(x) \leqq r + U(x; y) \leqq \frac{y}{2} + U(x; y)\,. \tag{15.2}$$

Let us further designate by $B(x; p_i \cdot p_j \cdots p_r)$ the number of odd numbers $n \leqq x$ for which $n(n+2)$ is divisible by $p_i p_j \cdots p_r$. Then

$$U(x; y) = \left[\frac{x+1}{2}\right] - \sum_i B(x; p_i) + \sum_{i<j} B(x; p_i \cdot p_j) - \sum_{i<j<k} B(x; p_i p_j p_k) + \cdots + (-1)^r B(x; p_1 p_2 \cdots p_r)\,, \tag{15.3}$$

where all primes are taken from the set of odd primes $\leqq y$. The validity of this formula can be seen by a process of enumeration.† On the left side are

† This formula is actually a special case of a formula of mathematical logic, sometimes called Sylvester's formula (see, e.g., G. Birkhoff and S. MacLane, *A Survey of Modern Algebra* (New York: Macmillan Co., 1953), pp. 347–348).

counted (by definition) those among the $[(x+1)/2]$ odd numbers $n \leqq x$ for which $n(n+2)$ is not divisible by any of the odd primes $3, 5, \cdots, p_r \leqq y$. All these are mentioned in the count $[(x+1)/2]$ on the right-hand side. However, all odd integers for which $n(n+2)$ is divisible by at least one of these primes is counted exactly zero times on the right-hand side. Indeed let $n(n+2)$ be divisible by the f primes $p_\alpha, p_\beta, \cdots, p_\lambda$, and only by these. Then n is counted once in $[(x+1)/2]$, f times in

$$\sum_i B(x; p_i)\,,$$

$\binom{f}{2}$ times in

$$\sum_{i<j} B(x; p_i p_j)\,,$$

$\binom{f}{3}$ times in

$$\sum_{i<j<k} B(x; p_i p_j p_k)\,,$$

and so on, and with the observation of the $\pm$ signs, altogether

$$1 - \binom{f}{1} + \binom{f}{2} - + \cdots + (-1)^f \binom{f}{f} = (1-1)^f = 0 \tag{15.4}$$

times, as had to be shown.

Let us abbreviate a number which is the product of f different prime factors taken from $3, 5, \cdots, p_r$ as $\rho^{(f)}$. Then (15.3) can be abbreviated as

$$U(x; y) = \left[\frac{x+1}{2}\right] + \sum_{f=1}^{r} (-1)^f \sum_{\rho^{(f)}} B(x; \rho^{(f)}) \tag{15.5}$$

where in the last inner sum $\rho^{(f)}$ runs over all products of f different prime factors, each taken from $3, 5, \cdots, p_r$.

It is now important, and this was Brun's decisive observation, not to use the full sum on the right-hand side, but to break it off at a suitably chosen index $f = m < r$. If we choose m *even*, we have

$$U(x; y) < \left[\frac{x+1}{2}\right] + \sum_{f=1}^{m} (-1)^f \sum_{\rho^{(f)}} B(x; \rho^{(f)})\,. \tag{15.6}$$

Instead of considering the full sum (15.4) we need the following lemma for counting.

LEMMA 1

$$\sum_{\lambda=0}^{m} (-1)^\lambda \binom{f}{\lambda} \begin{cases} = 0 \text{ for } m \geqq f > 0 \\ > 0 \text{ for } m < f,\ m \text{ even} \\ < 0 \text{ for } m < f,\ m \text{ odd}\,. \end{cases}$$

We leave the simple proof of this lemma to the reader. (For a proof observe that the binomial coefficients are increasing up to $\lambda = [f/2]$; for $m > [f/2]$ use the symmetry of the binomial coefficients and (15.4).)

The next step is the establishment of a handy expression for $B(x; \rho^{(f)})$. This is done in the following lemma, which we formulate, for the sake of convenience, a little more generally than we actually need it.

LEMMA 2: *Let ρ be an odd number and $\nu(\rho)$ the number of its* different *prime factors. Then the number $B(x; \rho)$ of odd numbers $n \leqq x$ for which $n(n+2)$ is divisible by ρ is*

$$B(x; \rho) = 2^{\nu(\rho)} \left\{ \frac{x}{2\rho} + \theta \right\}, \qquad |\theta| \leqq 1 . \tag{15.7}$$

Proof: Let

$$\rho = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_\nu^{\alpha_\nu}, \qquad \nu = \nu(\rho) .$$

Since the p_j are odd, none of these primes can divide both n and $n+2$. If $n(n+2)$ is divisible by ρ, then some of the prime powers in ρ must divide n, the others $n+2$; for instance,

$$\begin{aligned} n &\equiv 0 \pmod{p_1^{\alpha_1} \cdots p_\lambda^{\alpha_\lambda}} \\ n+2 &\equiv 0 \pmod{p_{\lambda+1}^{\alpha_{\lambda+1}} \cdots p_\nu^{\alpha_\nu}} . \end{aligned} \tag{15.8}$$

These two congruences together are equivalent to *one* congruence modulo ρ, according to the Chinese remainder theorem. In the range $1 \leqq n \leqq x$ for odd n the congruences will therefore have $[x/2\rho]$ or $[x/2\rho]+1$ solutions,† or in any case $x/2\rho + \theta$ solutions with a certain θ, $|\theta| \leqq 1$.

Now for $n(n+2) \equiv 0 \pmod{\rho}$ we can have exactly $2^\nu = 2^{\nu(\rho)}$ distributions of the $\nu(\rho)$ prime powers between n and $n+2$, which means that we have $2^{\nu(\rho)}$ pairs of congruences like (15.8). Consequently we obtain the result (15.7) with a new meaning of θ.

Remark: Because $x/2\rho$ for large ρ is small compared with the upper limit 1 of $|\theta|$, the formula (15.7) is not useful for large ρ. It is this fact which makes it advisable to stop the sum in (15.6) at a certain suitable $m < r$.

Collecting now our results from (15.1), (15.6), and (15.7) we obtain

$$T(x) \leqq r + \frac{x}{2} \sum_{f=0}^{m} (-1)^f \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} + \sum_{f=0}^{m} \sum_{\rho^{(f)}} 2^f , \tag{15.9}$$

where for convenience we have introduced the term $f = 0$ with $\rho^{(0)} = 1$. Evidently $\nu(\rho^{(f)}) = f$. We remember that m is even.

† These solutions are equidistant in the range and have the distance 2ρ.

Second step of V. Brun's method: Estimations. The last term in (15.9) is the easiest to estimate. Since $\rho^{(f)}$ runs through all products of f prime factors, each taken from the set $3, 5, \cdots, p_r$, we have

$$\sum_{f=0}^{m} \sum_{\rho^{(f)}} 2^f = \sum_{f=0}^{m} \binom{r}{f} 2^f < \sum_{f=0}^{m} (2r)^f < \frac{(2r)^{m+1}}{2r-1} \leqq (2r)^{m+1} .$$

On the other hand, we have

$$\sum_{f=0}^{m} (-1)^f \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} = \sum_{f=0}^{r} (-1)^f \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} - \sum_{f=m+1}^{r} (-1)^f \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} .$$

The first sum is obtained as the result of multiplying out

$$\prod_{j=1}^{r} \left(1 - \frac{2}{p_j}\right) .$$

In the second we put

$$s_f = \sum_{\rho^{(f)}} \frac{2^f}{\rho^{(f)}} \tag{15.10}$$

and observe for a later purpose that s_f is the fth elementary symmetric function of the quantities

$$\frac{2}{3}, \frac{2}{5}, \cdots, \frac{2}{p_r} . \tag{15.11}$$

We thus have

$$T(x) \leqq r + \frac{x}{2} \prod_{j=1}^{r} \left(1 - \frac{2}{p_j}\right) + \frac{x}{2} \sum_{f=m+1}^{r} (-1)^{f-1} s_f + (2r)^{m+1} . \tag{15.12}$$

Now, between successive elementary symmetric functions of any r positive quantities, the following inequality holds:

$$s_1 \cdot s_f \geqq (f+1) s_{f+1}, \qquad f = 1, 2, \cdots, \tag{15.13}$$

which remains true for $f > r$ if we put $s_{r+1} = s_{r+2} = \cdots = 0$. The above inequality becomes obvious through multiplication on the left-hand side, where each term of s_{f+1} appears $(f+1)$ times, besides some other (positive) terms. We deduce by iteration:

$$s_2 \leqq \frac{s_1^2}{2!}, \qquad s_3 \leqq \frac{1}{3} s_1 s_2 \leqq \frac{s_1^3}{3!},$$

and, in general,

$$s_f \leqq \frac{s_1^f}{f!} .$$

Inequality (15.13) shows moreover that

$$s_f \geqq s_{f+1}$$

provided

$$f + 1 \geqq s_1 .$$

In our case we have

$$s_1 = \frac{2}{3} + \frac{2}{5} + \cdots + \frac{2}{p_r} = 2 \sum_{3 \leq p \leq y} \frac{1}{p}, \tag{15.14}$$

where s_1 evidently depends on y, which has still to be chosen. If we take m therefore to satisfy

$$m + 1 \geqq s_1 . \tag{15.15}$$

then the sum on the right side of (15.12) is an alternating sum whose terms are decreasing in absolute value. Therefore

$$\sum_{f=m+1}^{r} (-1)^{f-1} s_f \leqq s_{m+1} \leqq \frac{s_1^{m+1}}{(m+1)!} \leqq \left(\frac{e s_1}{m+1}\right)^{m+1}$$

where we have used an estimate of the factorial from Chapter 13.

We know that s_1 defined in (15.14) is increasing without bounds with y:

$$s_1 = 2 \log\log y + O(1) ,$$

after Theorem 50.
If we determine $m + 1$ so that

$$e^2 s_1 < m + 1 < 9 s_1 ,$$

which certainly will satisfy (15.15), we have

$$\sum_{f=m+1}^{r} (-1)^{f-1} s_f \leqq \left(\frac{1}{e}\right)^{m+1} < e^{-e^2 s_1} < e^{-s_1} .$$

Furthermore, since for any real z

$$1 - z \leqq e^{-z} ,$$

we have

$$\prod_{j=1}^{r} \left(1 - \frac{2}{p_j}\right) < e^{-2 \sum_{j=1}^{r} \frac{1}{p_j}} = e^{-s_1} ,$$

so that (15.12) goes over into

$$T(x) \leqq y + x e^{-s_1} + y^{9 s_1} , \tag{15.16}$$

where we have made the trivial observation $2r < y$.

Third step of V. Brun's method: Choice of a parameter. We have, for y large enough, $2 \log\log y - B < s_1 < 3 \log\log y$ for a suitable positive B. Thus (15.16) can be replaced by

$$T(x) \leqq y + e^B \frac{x}{(\log y)^2} + y^{27 \log\log y} .$$

For the choice of $y \leqq \sqrt{x}$ our attention has to be focused on the last two terms of this inequality. Let us put

$$y = x^\gamma , \qquad 0 < \gamma \leqq \tfrac{1}{2} .$$

We can see at once that any choice of a constant positive γ, however small, will make the last term grow faster than x, thus nullifying our efforts. We have

$$T(x) < x^{\frac{1}{2}} + e^B \frac{x}{(\gamma \log x)^2} + x^{27\gamma \log\log x} .$$

Now choose

$$\gamma = \frac{1}{30 \log\log x} , \qquad x \geqq 3 .$$

This leads to

$$T(x) < x^{\frac{1}{2}} + 900\, e^B x \left(\frac{\log\log x}{\log x}\right)^2 + x^{\frac{1}{15}} .$$

For large x the second summand will be the predominant one here.

Let $T^*(x)$ be the number of all twin primes $\leqq x$ (not only the first members of each pair as counted in $T(x)$). Then obviously

$$T^*(x) \leqq 2T(x) .$$

We have thus obtained Viggo Brun's theorem.

THEOREM 59: *There exists a positive constant C so that $T^*(x)$, the number of twin primes not exceeding x, satisfies, for $x > 3$,*

$$T^*(x) < Cx \left(\frac{\log\log x}{\log x}\right)^2 . \tag{15.17}$$

The sum of the reciprocals of the twin primes. The discussion of this sum is now a simple matter of partial summation. We have

$$S(x) = \sum_{\substack{p \text{ twin prime} \\ p \leqq x}} \frac{1}{p} = \sum_{\substack{3 \leq n \leq x \\ n \text{ odd}}} \frac{1}{n} \left(T^*(n) - T^*(n-2)\right) .$$

Partial summation yields

$$S(x) = \sum_{\substack{3 \leq n \leq x \\ n \text{ odd}}} T^*(n) \left(\frac{1}{n} - \frac{1}{n+2}\right) + T^*(\xi) \frac{1}{\xi + 2} ,$$

where $\xi = 2[(x-1)/2] + 1$ is the greatest odd integer not exceeding x. Thus, in view of (15.17),

$$S(x) = 2 \sum_{\substack{3 \leq n \leq x \\ n \text{ odd}}} T^*(n) \frac{1}{n(n+2)} + O\left(\frac{\log\log x}{\log x}\right)^2$$

$$= O\left(\sum_{3 \leq n \leq x} \frac{(\log\log n)^2}{n (\log n)^2}\right) .$$

It is well known (and easily proved) that

$$\sum_{2 \leq n < \infty} \frac{1}{n(\log n)^{1+\varepsilon}}$$

is convergent for $\varepsilon > 0$. *A fortiori* the sum

$$\sum_{3 \leq n < \infty} \frac{(\log \log n)^2}{n(\log n)^2}$$

is convergent. This proves Viggo Brun's Theorem 58.

Additional remarks. What we have done in proving (15.17) is to count not only twin primes but such numbers n, $n+2$ below x which both have relatively large prime factors, since we have sieved out all multiples of $p_j \leqq y = x^{1/80 \log \log x}$.

The sieve method has subsequently been refined so that multiples of primes $p_j \leqq x^c$ with a certain *fixed* c could all be eliminated. In this way

$$T^*(x) < C \frac{x}{(\log x)^2}$$

could be established.

The problem of the twin primes is in some respects akin to Goldbach's problem: Is every even number $\geqq 4$ the sum of 2 primes? Viggo Brun could indeed apply his idea of the double sieving to this problem and obtained a result weaker than Goldbach's conjecture, but of a similar nature: Every large enough even number is the sum of two numbers, each of which is a prime or a product of at most 9 primes.

This has been improved further, in particular by A. Selberg, and the best-known result deals with numbers which are products of at most 3 primes.

If we insist, however, that the summands have to be primes, then one can prove by the sieve method together with the ingenious arguments of Schnirelmann about the "density" of certain sequences of integers among all natural numbers, that every large enough number is the sum of at most 20 primes.

However, since the 1920's a completely different method, far from elementary, has been used in problems of this sort. It was invented by Hardy and Littlewood and utilizes power series and the theory of functions of a complex variable. The first result of Hardy and Littlewood in this direction was still based on a certain unproved hypothesis. The Russian mathematician I. M. Vinogradov later improved the method so that the unproved hypothesis was eliminated. Vinogradov still has not proved Goldbach's theorem in full, but we know through him that every large enough *odd* number is the sum of 3 primes.

Index

It is well known (and easily proved) that

$$\sum_{2 < n < \infty} \frac{1}{n(\log n)^{1+\varepsilon}}$$

is convergent for $\varepsilon > 0$. *A fortiori* the sum

$$\sum_{3 \leq n < \infty} \frac{(\log \log n)^2}{n(\log n)^2}$$

is convergent. This proves Viggo Brun's Theorem 58.

Additional remarks. What we have done in proving (15.17) is to count not only twin primes but such numbers n, $n+2$ below x which both have relatively large prime factors, since we have sieved out all multiples of $p_j < y = x^{1/30 \log \log x}$.

The sieve method has subsequently been refined so that multiples of primes $p_j \leqq x^c$ with a certain *fixed* c could all be eliminated. In this way

$$T^*(x) < C \frac{x}{(\log x)^2}$$

could be established.

The problem of the twin primes is in some respects akin to Goldbach's problem: Is every even number >4 the sum of 2 primes? Viggo Brun could indeed apply his idea of the double sieving to this problem and obtained a result weaker than Goldbach's conjecture, but of a similar nature: Every large enough even number is the sum of two numbers, each of which is a prime or a product of at most 9 primes.

This has been improved further, in particular by A. Selberg, and the best-known result deals with numbers which are products of at most 3 primes.

If we insist, however, that the summands have to be primes, then one can prove by the sieve method together with the ingenious arguments of Schnirelmann about the "density" of certain sequences of integers among all natural numbers, that every large enough number is the sum of at most 20 primes.

However, since the 1920's a completely different method, far from elementary, has been used in problems of this sort. It was invented by Hardy and Littlewood and utilizes power series and the theory of functions of a complex variable. The first result of Hardy and Littlewood in this direction was still based on a certain unproved hypothesis. The Russian mathematician I. M. Vinogradov later improved the method so that the unproved hypothesis was eliminated. Vinogradov still has not proved Goldbach's theorem in full, but we know through him that every large enough *odd* number is the sum of 3 primes.

Index

PRINTED IN THE UNITED STATES OF AMERICA