

# AN INTRODUCTION TO THE THEORY OF NUMBERS



*Leo Moser*  
University of Alberta

University of Alberta

An Introduction to the Theory of Numbers

Leo Moser

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (<https://LibreTexts.org>) and like the hundreds of other texts available within this powerful platform, it is freely available for reading, printing and "consuming." Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects.

Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



The LibreTexts mission is to unite students, faculty and scholars in a cooperative effort to develop an easy-to-use online platform for the construction, customization, and dissemination of OER content to reduce the burdens of unreasonable textbook costs to our students and society. The LibreTexts project is a multi-institutional collaborative venture to develop the next generation of open-access texts to improve postsecondary education at all levels of higher learning by developing an Open Access Resource environment. The project currently consists of 14 independently operating and interconnected libraries that are constantly being optimized by students, faculty, and outside experts to supplant conventional paper-based books. These free textbook alternatives are organized within a central environment that is both vertically (from advance to basic level) and horizontally (across different fields) integrated.

The LibreTexts libraries are Powered by [NICE CXOne](#) and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptations contact [info@LibreTexts.org](mailto:info@LibreTexts.org). More information on our activities can be found via Facebook (<https://facebook.com/Libretexts>), Twitter (<https://twitter.com/libretexts>), or our blog (<http://Blog.Libretexts.org>).

This text was compiled on 06/01/2023

## TABLE OF CONTENTS

### Chapters

- [Licensing](#)
- [1.1: Compositions and Partitions](#)
- [1.2: Arithmetic Functions](#)
- [1.3: Distribution of Primes](#)
- [1.4: Irrational Numbers](#)
- [1.5: Congruences](#)
- [1.6: Diophantine Equations](#)
- [1.7: Combinatorial Number Theory](#)
- [1.8: Geometry of Numbers](#)
- [Index](#)
- [Detailed Licensing](#)

### Index

## Licensing

---

*A detailed breakdown of this resource's licensing can be found in [Back Matter/Detailed Licensing](#).*

## CHAPTER OVERVIEW

### Front Matter

[TitlePage](#)

[InfoPage](#)

[Table of Contents](#)

[Licensing](#)

University of Alberta

An Introduction to the Theory of Numbers

Leo Moser

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (<https://LibreTexts.org>) and like the hundreds of other texts available within this powerful platform, it is freely available for reading, printing and "consuming." Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects.

Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



The LibreTexts mission is to unite students, faculty and scholars in a cooperative effort to develop an easy-to-use online platform for the construction, customization, and dissemination of OER content to reduce the burdens of unreasonable textbook costs to our students and society. The LibreTexts project is a multi-institutional collaborative venture to develop the next generation of open-access texts to improve postsecondary education at all levels of higher learning by developing an Open Access Resource environment. The project currently consists of 14 independently operating and interconnected libraries that are constantly being optimized by students, faculty, and outside experts to supplant conventional paper-based books. These free textbook alternatives are organized within a central environment that is both vertically (from advance to basic level) and horizontally (across different fields) integrated.

The LibreTexts libraries are Powered by [NICE CXOne](#) and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptations contact [info@LibreTexts.org](mailto:info@LibreTexts.org). More information on our activities can be found via Facebook (<https://facebook.com/Libretexts>), Twitter (<https://twitter.com/libretexts>), or our blog (<http://Blog.Libretexts.org>).

This text was compiled on 06/01/2023



## TABLE OF CONTENTS

### Chapters

- [Licensing](#)
- [1.1: Compositions and Partitions](#)
- [1.2: Arithmetic Functions](#)
- [1.3: Distribution of Primes](#)
- [1.4: Irrational Numbers](#)
- [1.5: Congruences](#)
- [1.6: Diophantine Equations](#)
- [1.7: Combinatorial Number Theory](#)
- [1.8: Geometry of Numbers](#)
- [Index](#)
- [Detailed Licensing](#)

### Index

## Licensing

---

*A detailed breakdown of this resource's licensing can be found in [Back Matter/Detailed Licensing](#).*

## 1.2: Arithmetic Functions

The next topic we shall consider is that of arithmetic functions. These form the main objects of concern in number theory. We have already mentioned two such functions of two variables, the g.c.d. and l.c.m. of  $m$  and  $n$ , denoted by  $(m, n)$  and  $[m, n]$  respectively, as well as the functions  $c(n)$  and  $p(n)$ . Of more direct concern at this stage are the functions

$$\begin{aligned}\pi(n) &= \sum_{p \leq n} 1 && \text{the number of primes } n \text{ not exceeding } n; \\ w(n) &= \sum_{p|n} 1 && \text{the number of distinct primes factors of } n; \\ \omega(n) &= \sum_{p \leq n} 1 && \text{the number of primes } n \text{ not exceeding } n; \\ \Omega(n) &= \sum_{p^i \leq n} 1 && \text{the number of prime factors of } n; \\ \tau(n) &= \sum_{d|n} 1 && \text{the number of divisors } n; \\ \sigma(n) &= \sum_{d|n} d && \text{the sum of the divisors of } n \\ \varphi(n) &= \sum_{\substack{(a,n)=1 \\ 1 \leq a \leq n}} 1 && \text{the Euler totient function;} \end{aligned}$$

the Euler totient function counts the number of integers  $\leq n$  and relatively prime to  $n$ .

In the section we shall be particularly concerned with the functions  $\tau(n)$ ,  $\sigma(n)$ , and  $\varphi(n)$ . These have the important property that if

$$n = ab \text{ and } (a, b) = 1$$

then

$$f(ab) = f(a)f(b).$$

Any function satisfying this condition is called *weakly multiplicative*, or simply *multiplicative*.

A generalization of  $\tau(n)$  and  $\sigma(n)$  is afforded by

$$\sigma_k(n) = \sum_{d|n} d^k \text{ . then the sum of the } k^{\text{th}} \text{ powers of the divisors of } n,$$

since  $\sigma_0(n) = \tau(n)$  and  $\sigma_1(n) = \sigma(n)$ .

The  $\varphi$  function can also be generalized in many ways. We shall consider later the generalization due to Jordan,  $\varphi_k(n)$  = number of  $k$ -tuples  $\leq n$  whose g.c.d. is relatively prime to  $n$ . We shall derive some elementary properties of these and closely related functions and state some special solved and unsolved problems concerning them. We shall then discuss a theory which gives a unified approach to these functions and reveals unexpected interconnections between them. Later we shall discuss the magnitude of these functions. The functions  $\omega(n)$ ,  $\Omega(n)$ , and, particularly,  $\pi(n)$  are of a different nature and special attention will be given to them.

Suppose in what follows that the prime power factorization of  $n$  is given by

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s} \text{ or briefly } n = \prod p^{\alpha}.$$

We note that 1 is not a prime and take for granted the provable result that, apart from order, the factorization is unique.

In terms of this factorization the functions  $\sigma_k(n)$  and  $\varphi(n)$  are easily determined. It is not difficult to see that the terms in the expansion of the product

$$\prod_{p|n} (1 + p^k + p^{2k} + \cdots + p^{\alpha k})$$

are precisely the divisors of  $n$  raised to the  $k^{\text{th}}$  power. Hence we have the desired expansion for  $\sigma_k(n)$ . In particular

$$\tau(n) = \sigma_0(n) = \prod (\alpha + 1),$$

and

$$\sigma(n) = \sigma_1(n) = \prod_{p|n} (1 + p + p^2 + \cdots + p^{\alpha}) = \prod_{p|n} \frac{p^{\alpha+1} - 1}{p - 1},$$

$$\text{e.g., } 60 = 2^2 \cdot 3^1 \cdot 5^1,$$

$$\tau(60) = (2+1)(1+1)(1+1) = 3 \cdot 2 \cdot 2 = 12,$$

$$\sigma(60) = (1 + 2 + 2^2)(1 + 3)(1 + 5) = 7 \cdot 4 \cdot 6 = 168.$$

These formulas reveal the multiplicative nature of  $\sigma_k(n)$ .

To obtain an explicit formula for  $\varphi(n)$  we make use of the following well-known combinatorial principle.

#### The principle of inclusion and exclusion

Given  $N$  objects each of which may or may not possess any of the characteristics

$$A_1, A_2, \dots$$

Let  $N(A_i, A_j, \dots)$  be the number of objects having the characteristics  $A_i, A_j, \dots$  and possibly others. Then the number of objects which have none of these properties is

$$N - \sum N(A_i) + \sum_{i < j} N(A_i, A_j) - \sum_{i < j < k} N(A_i, A_j, A_k) + \dots,$$

where the summation is extended over all combinations of the subscripts 1, 2, ...,  $n$  in groups of one, two, three and so on, and the signs of the terms alternate.

An integer will be relatively prime to  $n$  only if it is not divisible by any of the prime factors of  $n$ . Let  $A_1, A_2, \dots, A_s$  denote divisibility by  $p_1, p_2, \dots, p_s$  respectively. Then, according to the combinatorial principle stated above

$$\varphi(n) = n - \sum_i \frac{n}{p_i} + \sum_{i < j} \frac{n}{p_i p_j} - \sum_{i < j < k} \frac{n}{p_i p_j p_k} + \dots$$

This expression can be factored into the form

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

e.g.

$$\varphi(60) = 60 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \left(1 - \frac{1}{5}\right) = 60 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} = 16.$$

A similar argument shows that

$$\varphi_k(n) = n^k \prod_{p|n} \left(1 - \frac{1}{p^k}\right).$$

The formula for  $\varphi(n)$  can also be written in the form

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

where  $\mu(d)$  takes on the values 0, 1, -1. Indeed  $\mu(d) = 0$  if  $d$  has a square factor,  $\mu(1) = 1$ , and  $\mu(p_1 p_2 \dots p_s) = (-1)^s$ . This gives some motivation for defining a function  $\mu(n)$  in this way. This function plays an unexpectedly important role in number theory.

Our definition of  $\mu(n)$  reveals its multiplicative nature, but it still seems rather artificial. It has however a number of very important properties which can be used as alternative definitions. We prove the most important of these, namely

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

Since  $\mu(d) = 0$  if  $d$  contains a squared factor, it suffices to suppose that  $n$  has no such factor, i.e.,  $n = p_1 p_2 \dots p_s$ . For such an  $n > 1$


$$\sum_{d|n} \mu(d) = 1 - \binom{n}{1} + \binom{n}{2} - \dots = (1 - 1)^n = 0.$$

By definition  $\mu(1) = 1$  so the theorem is proved.

If we sum this result over  $n = 1, 2, \dots, x$ , we obtain

$$\sum_{d=1}^x \left\lfloor \frac{x}{d} \right\rfloor \mu(d) = 1,$$

which is another defining relation.

Another very interesting defining property, the proof of which I shall leave  as an exercise, is that if

$$M(x) = \sum_{d=1}^x x \mu(d)$$

then

$$\sum_{d=1}^x M\left(\left[\frac{x}{d}\right]\right) = 1.$$

This is perhaps the most elegant definition of  $\mu$ . Still another very important property is that

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

We now turn our attention to Dirichlet multiplication and series.

Consider the set of arithmetic functions. These can be combined in various ways to give new functions. For example, we could define  $f + g$  by

$$(f + g)(n) = f(n) + g(n)$$

and

$$(f \cdot g)(n) = f(n) \cdot g(n)$$

A less obvious mode of combination is given by  $f \times g$ , defined by

$$(f \times g)(n) = \sum_{d|n} f(d)g\left(\frac{n}{d}\right) = \sum_{dd'=n} f(d)g(d').$$

This may be called the divisor product or Dirichlet product.

This motivation for this definition is as follows. If

$$F(s) = \sum_{n=1}^{\infty} f(n)n^{-s}, \quad G(s) = \sum_{n=1}^{\infty} g(n)n^{-s}, \quad \text{and} \quad F(s) \cdot G(s) = \sum_{n=1}^{\infty} h(n)n^{-s},$$

then it is readily checked that  $h = f \times g$ . Thus Dirichlet multiplication of arithmetic functions corresponds to the ordinary multiplication of the corresponding Dirichlet series:

$$f \times g = g \times f, \quad (f \times g) \times h = f \times (g \times h),$$

i.e., our multiplication is commutative and associative. A purely arithmetic proof of these results is easy to supply.

Let us now define the function

$$\ell = \ell(n) : 1, 0, 0, \dots$$

It is easily seen that  $f \times \ell = f$ . Thus the function  $\ell$  is the unity of our multiplication.

It can be proved without difficulty that if  $f(1) \neq 0$ , then  $f$  has an inverse with respect to  $\ell$ . Such functions are called regular. Thus the regular functions form a group with respect to the operation  $\times$ .

Another theorem, whose proof we shall omit, is that the Dirichlet product of multiplicative functions is again multiplicative.

We now introduce the functions

$$I_k : 1^k, 2^k, 3^k, \dots$$

It is interesting that, starting only with the functions  $\ell$  and  $I_k$ , we can build up many of the arithmetic functions and their important properties.

To begin with we may define  $\mu(n)$  by  $\mu = I_0^{-1}$ . This means, of course, that  $\mu \times I_0 = \ell$  or

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

and we have already seen that this is a defining property of the  $\mu$  function. We can define  $\sigma_k$  by

$$\sigma_k = I_0 \times I_k.$$

This means that

$$\sigma_k(n) = \sum_{d|n} (d^k \cdot 1),$$

which corresponds to our earlier definition. Special cases are

$$\tau = I_0 \times I_0 = I_0^2 \quad \text{and} \quad \sigma = I_0 \times I_1$$

Further, we can define

$$\varphi_k = \mu \times I_k = I_0^{-1} \times I_k.$$

This means that

$$\varphi_k(n) = \sum_{d|n} \mu(d) \left(\frac{n}{d}\right)^k,$$

which again can be seen to correspond to our earlier definition.

The special case of interest here is

$$\varphi = \varphi_1 = \mu \times I_1.$$

Now, to obtain some important relations between our functions, we note the so-called Möbius inversion formula. From our point of view this says that

$$g = f \times I_0 \iff f = u \times g$$

This is, of course, quite transparent. Written out in full it states that

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

In this form it is considerably less obvious.

Consider now the following applications. First

$$\sigma_k = I_0 \times I_k \iff I_k = \mu \times \sigma_k.$$

This means that

$$\sum_{d|n} \mu(d) \sigma_k\left(\frac{n}{d}\right) = n^k.$$

Important special cases are

$$\sum_{d|n} \mu(d) \tau\left(\frac{n}{d}\right) = 1,$$

and

$$\sum_{d|n} \mu(d) \sigma\left(\frac{n}{d}\right) = 1.$$

Again

$$\varphi_k = I_0^{-1} \times I_k \iff I_k = I_0 \times \varphi_k,$$

so that

$$\sum_{d|n} \varphi_k(d) = n^k,$$

We can obtain identities of a somewhat different kind. Thus

$$\sigma_k \times \varphi_k = I_0 \times I_k \times I_0^{-1} \times I_k = I_k \times I_k,$$

and hence

$$\sum_{d|n} \sigma_k(d) \varphi_k\left(\frac{n}{d}\right) = \sum_{d|n} d^k \left(\frac{n}{d}\right)^k = \sum_{d|n} n^k = \tau(n) n^k.$$

A special case of interest here is

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

In order to make our calculus applicable to problems concerning distribution of primes, we introduce a unary operation on our functions, called differentiation:

$$f'(n) = -f(n) \log n$$

The motivation for this definition can be seen from

$$\frac{d}{ds} \left( \sum \frac{f(n)}{n^s} \right) = - \sum \frac{(\log n) f(n)}{n^s}.$$

Now let us define

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^\alpha, \\ 0 & \text{if } n \neq p^\alpha. \end{cases}$$

It is easily seen that

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

In our Dirichlet multiplication notation we have

$$\Lambda \times I_0 = -I'_0,$$

so that

$$\Lambda = I_0^{-1} \times (-I'_0) = \mu \times (-I'_0)$$

or

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

Let us now interpret some of our results in terms of Dirichlet series. We have the correspondence

$$F(s) \leftrightarrow f(n) \text{ if } F(s) = \sum \frac{f(n)}{n^s},$$

and we know that Dirichlet multiplication of arithmetic functions corresponds to ordinary multiplication for Dirichlet series. We start with

$$f \leftrightarrow F, 1 \leftrightarrow 1, \text{ and } I_0 \leftrightarrow \zeta(s).$$

Furthermore

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

Also

$$\mu \leftrightarrow \frac{1}{\zeta(s)} \text{ and } I'_0 \leftrightarrow \sum \frac{-\log n}{n^s} = \zeta'(s).$$

This yields

$$\sum \frac{\sigma_k(n)}{n^s} = \zeta(s) \zeta(s-k).$$

Special cases are

$$\sum \frac{\tau(n)}{n^s} = \zeta^2(s)$$

and

$$\sum \frac{\sigma(n)}{n^s} = \zeta(s) \zeta(s-1).$$

Again

$$\sum \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$$

and

$$\sum \frac{\varphi_k(n)}{n^s} = \frac{\zeta(s-k)}{\zeta(s)},$$

with the special case

$$\sum \frac{\varphi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}.$$

To bring a few of these down to quite numerical results we have

$$\begin{aligned}\sum \frac{\tau(n)}{n^2} &= \zeta^2(2) = \frac{\pi^4}{36}, \\ \sum \frac{\sigma_4(n)}{n^2} &= \zeta(2) \cdot \zeta(4) = \frac{\pi^2}{6} \cdot \frac{\pi^4}{90} = \frac{\pi^6}{540}, \\ \sum \frac{\mu(n)}{n^2} &= \frac{6}{\pi^2}\end{aligned}$$

As for our  $\Lambda$  function, we had

$$\Lambda = I_0^{-1} \times I'_0;$$

this means that

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = \frac{-\zeta'(s)}{\zeta(s)}. \quad (1.2.1)$$

The prime number theorem depends on going from this to a reasonable estimate for

$$\Psi(x) = \sum_{n=1}^x \Lambda(n).$$

Indeed we wish to show that  $\Psi(x) \sim x$ .

Any contour integration with the right side of (1) involves of course the need for knowing where  $\zeta(s)$  vanishes. This is one of the central problems of number theory.

Let us briefly discuss some other Dirichlet series.

If  $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}$  define

$$\lambda(n) = (-1)^{\alpha_1 + \alpha_2 + \cdots + \alpha_s}.$$

The  $\lambda$  function has properties similar to those of the  $\mu$  function. We leave as an exercise to show that

$$\sum_{d|n} \lambda(d) = \begin{cases} 1 & \text{if } n = r^2. \\ 0 & \text{if } n \neq r^2. \end{cases}$$

Now

$$\zeta(2s) = \sum \frac{s(n)}{n^s} \text{ where } s(n) = \begin{cases} 1 & \text{if } n = r^2. \\ 0 & \text{if } n \neq r^2. \end{cases}$$

Hence  $\lambda \times I_0 = s$ , i.e.,

$$\sum \frac{\lambda(n)}{n^s} \cdot \zeta(s) = \zeta(2s)$$

or

$$\sum \frac{\lambda(n)}{n^s} = \frac{\zeta(2s)}{\zeta(s)}.$$

For example

$$\sum \frac{\lambda(n)}{n^2} = \frac{\pi^4}{90} / \frac{\pi^2}{6} = \frac{\pi^2}{15}.$$

We shall conclude with a brief look at another type of generating series, namely Lambert series. These are series of the type

$$\sum \frac{f(n)x^n}{1-x^n}.$$

It is easily shown that if  $F = f \times I_0$  then

$$\sum \frac{f(n)x^n}{1-x^n} = \sum F(n)x^n.$$

Interesting special cases are



$$f = I_0, \sum \frac{x^n}{1-x^n} = \sum \tau(n)x^n;$$

$$f = \mu, \sum \mu(n) \frac{x^n}{1-x^n} = x;$$

$$f = \varphi, \sum \varphi(n) \frac{x^n}{1-x^n} = \sum nx^n = \frac{x}{(1-x)^2}.$$

For example, taking  $x = \frac{1}{10}$  in the last equality, we obtain

$$\frac{\varphi(1)}{9} + \frac{\varphi(2)}{99} + \frac{\varphi(3)}{999} + \cdots = \frac{10}{81}.$$

### Exercise 1.2.1

Prove that  $\sum_{n=1}^{\infty} \frac{\mu(n)x^n}{1+x^n} = x - 2x^2$ .

Prove that  $\sum_{n=1}^{\infty} \frac{\lambda(n)x^n}{1-x^n} = \sum_{n=1}^{\infty} x^{n^2}$ .

### Answer

Add texts here. Do not delete this text first.

This page titled [1.2: Arithmetic Functions](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Leo Moser \(The Trilla Group\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

## 1.3: Distribution of Primes

Perhaps the best known proof in all of “real” mathematics is Euclid’s proof of the existence of infinitely many primes.

### Theorem 1.3.1: existence of infinitely many primes

If  $p$  were the largest prime then  $(2 \cdot 3 \cdot 5 \cdots p) + 1$  would not be divisible by any of the primes up to  $p$  and hence would be the product of primes exceeding  $p$

In spite of its extreme simplicity this proof already raises many exceedingly difficult questions, e.g., are the numbers  $(2 \cdot 3 \cdots p) + 1$  usually prime or composite? No general results are known. In fact, we do not know whether an infinity of these numbers are prime, or whether an infinity are composite.

The proof can be varied in many ways. Thus, we might consider  $(2 \cdot 3 \cdot 5 \cdots p) - 1$  or  $p! + 1$  or  $p! - 1$ . Again almost nothing is known about how such numbers factor. The last two sets of numbers bring to mind a problem that reveals how, in number theory, the trivial may be very close to the most abstruse. It is rather trivial that for  $n > 2$ ,  $n! - 1$  is not a perfect square. What can be said about  $n! + 1$ ? Well,  $4! + 1 = 5^2$ ,  $5! + 1 = 11^2$  and  $7! + 1 = 71^2$ . However, no other cases are known; nor is it known whether any other numbers  $n! + 1$  are perfect squares. We will return to this problem in the lectures on diophantine equations.

After Euclid, the next substantial progress in the theory of distribution of primes was made by Euler. He proved that  $\sum \frac{1}{p}$  diverges, and described this result by saying that the primes are more numerous than the squares. I would like to present now a new proof of this fact -- a proof that is somewhat related to Euclid’s proof of the existence of infinitely many primes.

We need first a (well known) lemma concerning subseries of the harmonic series. Let  $p_1 < p_2 < \dots$  be a sequence of positive integers and let its counting function be

$$\pi(x) = \sum_{p \leq x} 1. \quad (1.3.1)$$

Let

$$R(x) = \sum_{p \leq x} \frac{1}{p}. \quad (1.3.2)$$

### Lemma

If  $R(\infty)$  exists then

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x} = 0. \quad (1.3.3)$$

### Proof

$$\pi(x) = 1(R(1) - R(0)) + 2((R(2) - R(1)) + \dots + x(R(x) - R(x-1))) \quad (1.3.4)$$

or

$$\frac{\pi(x)}{x} = R(x) - \left[ \frac{R(0) + R(1) + \dots + R(x-1)}{x} \right]. \quad (1.3.5)$$

Since  $R(x)$  approaches a limit, the expression within the square brackets approaches this limit and the lemma is proved.

In what follows we assume that the  $p$ ’s are the primes.

To prove that  $\sum \frac{1}{p}$  diverges we will assume the opposite, i.e.,  $\sum \frac{1}{p}$  converges (and hence also that  $\frac{\pi(x)}{x} \rightarrow 0$ ) and derive a contradiction.

$$\sum_{p > n} \frac{1}{p} < \frac{1}{2}. \quad (1.3.6)$$

But now this  $n$  is fixed so there will also be an  $m$  such that

$$\frac{\pi(n!m)}{n!m} < \frac{1}{2n!}. \quad (1.3.7)$$

With such an  $n$  and  $m$  we form the  $m$  numbers

$$T_1 = n! - 1, T_2 = 2n! - 1, \dots, T_m = mn! - 1.$$

Note that none of the  $T$ 's have prime factors  $\leq n$  or  $\geq mn!$ . Furthermore if  $p|T_i$  and  $p|T_j$  then  $p|(T_i - T_j)$  so that  $p|(i - j)$ . In other words, the multiples of  $p$  are  $p$  apart in our set of numbers. Hence not more than  $\frac{m}{p} + 1$  of the numbers are divisible by  $p$ .

Since every number has at least one prime factor we have

$$\sum_{n < p < n!m} \left( \frac{m}{p} + 1 \right) \geq m$$

or

$$\sum_{n < p} \frac{1}{p} + \frac{\pi(n!m)}{m} \geq 1.$$

But now by (1) and (2) the right hand side should be  $< 1$  and we have a contradiction, which proves our theorem.

Euler's proof, which is more significant, depends on his very important identity

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

This identity is essentially an analytic statement of the unique factorization theorem. Formally, its validity can easily be seen. We have

$$\begin{aligned} \prod_p \frac{1}{1 - \frac{1}{p^s}} &= \prod_p \left( 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots \right) \\ &= \left( 1 + \frac{1}{2^s} + \dots \right) \left( 1 + \frac{1}{3^s} + \dots \right) \left( 1 + \frac{1}{5^s} + \dots \right) \\ &= \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \dots. \end{aligned}$$

Euler now argued that for  $s = 1$ ,

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

so that

$$\prod_p \frac{1}{1 - \frac{1}{p}}$$

must be infinite, which in turn implies that  $\sum \frac{1}{p}$  must be infinite.

This argument, although not quite valid, can certainly be made valid. In fact, it can be shown without much difficulty that

$$\sum_{n \leq x} \frac{1}{n} - \prod_{p \leq x} \left( 1 - \frac{1}{p} \right)^{-1}$$

is bounded. Since  $\sum_{n \leq x} \frac{1}{n} - \log x$  is bounded, we can, on taking logs, obtain

$$\log \log x = \sum_{p \leq x} -\log \left( 1 - \frac{1}{p} \right) + O(1)$$

so that

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

We shall use this result later.

Gauss and Legendre were the first to make reasonable estimates for  $\pi(x)$ . Essentially, they conjectured that

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

the famous Prime Number Theorem. Although this was proved in 1896 by J. Hadamard and de la Vallee Poussin, the first substantial progress towards this result is due to Chebycheff. He obtained the following three main results:

1. There is a prime between  $n$  and  $2n$  ( $n > 1$ );
2. There exist positive constants  $c_1$  and  $c_2$  such that

$$\frac{c_2 x}{\log x} < \pi(x) < \frac{c_1 x}{\log x}; \quad (1.3.8)$$

3. If  $\frac{\pi(x)}{\log x}$  approaches a limit, then that limit is 1.

We shall prove the three main results of Chebycheff using his methods as modified by Landau, Erdős and, to a minor extent, L. Moser.

We require a number of lemmas. The first of these relate to the magnitude of

$$n! \text{ and } \binom{2n}{n}.$$

As far as  $n!$  is concerned, we might use **Stirling's approximation**

$$n! \sim n^n e^{-n} \sqrt{2\pi n}.$$

However, for our purposes, a simpler estimate will suffice. Since

$$\frac{n^n}{n!}$$

and we have the following lemma.

#### Lemma 1

$$n^n e^{-n} < n! < n^n.$$

#### Proof

Since

$$(1+1)^{2n} = 1 + \binom{2n}{1} + \cdots + \binom{2n}{n} + \cdots + 1,$$

it follows that

$$\binom{2n}{n} < 2^{2n} = 4^n.$$

This estimate for  $\binom{2n}{n}$  is not as crude as it looks, for it can be easily seen from Stirling's formula that

$$\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi n}}.$$

Using induction we can show for  $n > 1$  that

$$\binom{2n}{n} > \frac{4^n}{2n},$$

and thus we have

### lemma 2

$$\frac{4^n}{2^n} < \binom{2n}{n} < 4^n.$$

#### Proof

Note that  $\binom{2n+1}{n}$  is one of two equal terms in the expansion of  $(1+1)^{2n+1}$ . Hence we also have

### Lemma 3

$$\binom{2n+1}{n} < 4^n.$$

#### Proof

As an exercise you might use these to prove that if

$$n = a + b + c + \dots$$

then

$$\frac{n!}{a!b!c!\dots} < \frac{n^n}{a^a b^b c^c \dots}.$$

Now we deduce information on how  $n!$  and  $\binom{2n}{n}$  factor as the product of primes. Suppose  $e_p(n)$  is the exponent of the prime  $p$  in the prime power factorization of  $n!$ , i.e.

$$n! = \prod p^{e_p(n)}.$$

We easily prove

### lemma 4

$$(Legendre). \quad e_p(n) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \dots$$

#### Proof

In fact  $\left\lfloor \frac{n}{p} \right\rfloor$  is the number of multiples of  $p$  in  $n!$ , the term  $\left\lfloor \frac{n}{p^2} \right\rfloor$  adds the additional contribution of the multiples of  $p^2$ , and so on, e.g.,

$$e_3(30) = \left\lfloor \frac{30}{3} \right\rfloor + \left\lfloor \frac{30}{9} \right\rfloor + \left\lfloor \frac{30}{27} \right\rfloor + \dots = 10 + 3 + 1 = 14.$$

An interesting and sometimes useful alternative expression for  $e_p(n)$  is given by

$$e_p(n) = \frac{n - s_p(n)}{p-1},$$

where  $s_p(n)$  represents the sum of the digits of  $n$  when  $n$  is expressed in base  $p$ . Thus in base 3, 30 can be written 1010 so that  $e_3(30) = \frac{30-2}{2} = 14$  as before. We leave the proof of the general case as an exercise.

We next consider the composition of  $\binom{2n}{n}$  as a product of primes. Let  $E_p(n)$  denote the exponent of  $p$  in  $\binom{2n}{n}$ , i.e.,

$$\binom{2n}{n} = \prod_p p^{E_p(n)}.$$

Clearly

$$E_p(n) = e_p(2n) - 2e_p(n) = \sum_i \left\lfloor \frac{2n}{p^i} \right\rfloor - 2 \left\lfloor \frac{n}{p^i} \right\rfloor.$$

Our alternative expression for  $e_p(n)$  yields

$$E_p(n) = \frac{2s_p(n) - s_p(2n)}{p-1}.$$

In the first expression each term in the sum can easily be seen to be 0 or 1 and the number of terms does not exceed the exponent of the highest power of  $p$  that does not exceed  $2n$ . Hence

#### lemma 5

$$E_p(n) \leq \log_p(2n).$$

#### lemma 6

The contribution of  $p$  to  $\binom{2n}{n}$  does not exceed  $2n$

The following three lemmas are immediate.

#### lemma 7

Every prime in the range  $n < p < 2n$  appears exactly once in  $\binom{2n}{n}$

#### lemma 8

No prime in the range  $p > \sqrt{2n}$  appears more than once in  $\binom{2n}{n}$

Although it is not needed in what follows, it is amusing to note that since  $E_2(n) = 2s_2(n) - s_2(2n)$  and  $s_2(n) = s_2(2n)$ , we have  $E_2(n) = s_2(n)$ .

As a first application of the lemmas we prove the elegant result

#### Theorem 1

$$\prod_{p \leq n} p < 4^n.$$

##### Proof

The proof is by induction on  $n$ . We assume the theorem true for integers  $< n$  and consider the cases  $n = 2m$  and  $n = 2m + 1$ . If  $n = 2m$  then

$$\prod_{p \leq 2m} p = \prod_{p \leq 2m-1} p < 4^{2m-1}$$

by the induction hypothesis. If  $n = 2m + 1$  then

$$\begin{aligned} \prod_{p \leq 2m+1} p &= (\prod_{p \leq m+1} p)(\prod_{m+1 < p < 2m+1} p) \\ &< 4^{m+1} \binom{2m+1}{m} \leq 4^{m+1} 4^m = 4^{2m+1} \end{aligned}$$

and the induction is complete.

It can be shown by much deeper methods (Rosser) that

$$\prod_{p \leq n} p < (2.83)^n.$$

Actually the prime number theorem is equivalent to

$$\sum_{p \leq n} \log p \sim n.$$

From Theorem 1 we can deduce

#### Theorem 2

$$\pi(n) < \frac{cn}{\log n}.$$

##### Proof

Clearly

$$4^n > \prod_{p \leq n} p > \prod_{\sqrt{n} \leq p \leq n} p > \sqrt{n}^{\pi(n) - \pi(\sqrt{n})}$$

Taking logarithms we obtain

$$n \log 4 > (\pi(n) - \pi(\sqrt{n})) \frac{1}{2} \log n$$

or

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

or

$$\pi(n) < (4 \log 2) \frac{n}{\log n} + \sqrt{n} < \frac{cn}{\log n}.$$

Next we prove

### Theorem 3

$$\pi(n) > \frac{cn}{\log n}.$$

#### Proof

For this we use Lemmas 6 and 2. From these we obtain

$$(2n)^{\pi(2n)} > \binom{2n}{n} > \frac{4^n}{2n}.$$

Taking logarithms, we find that

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

Thus, for even  $m$

$$\pi(m) + 1 > \frac{m}{\log m} \log 2$$

and the result follows.

We next obtain an estimate for  $S(x) = \sum_{p \leq x} \frac{\log p}{p}$ . Taking the logarithm of  $n! = \prod_p p^{e_p}$  we find that

$$n \log n > \log n! = \sum e_p(n) \log p > n(\log n - 1).$$

The reader may justify that the error introduced in replacing  $e_p(n)$  by  $\frac{n}{p}$  (of course  $e_p(n) = \sum [\frac{n}{p^i}]$ ) is small enough that

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

or

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

We can now prove

### Theorem 4

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

#### Proof

In fact

$$\begin{aligned}
 R(x) &= \sum_{n=2}^x \frac{S(n) - S(n-1)}{\log n} \\
 &= \sum_{n=2}^x S(n) \left( \frac{1}{\log n} - \frac{1}{\log(n+1)} \right) + O(1) \\
 &= \sum_{n=2}^x (\log n + O(1)) \frac{\log \left(1 + \frac{1}{n}\right)}{(\log n) \log(n+1)} + O(1) \\
 &= \sum_{n=2}^x \frac{1}{n \log n} + O(1) \\
 &= \log \log x + O(1)
 \end{aligned}$$

as desired.

We now outline the proof of Chebycheff's

### Theorem 5

If  $\pi(x) \sim \frac{cx}{\log x}$ , then  $c = 1$ .

#### Proof

Since

$$\begin{aligned}
 R(x) &= \sum_{n=1}^x \frac{\pi(n) - \pi(n-1)}{n} \\
 &= \sum_{n=1}^x \frac{\pi(n)}{n^2} + O(1)
 \end{aligned}$$

$\pi(x) \sim \frac{cx}{\log x}$  would imply

$$\sum_{n=1}^x \frac{\pi(n)}{n^2} \sim c \log \log x.$$

But we already know that  $\pi(x) \sim \log \log x$  so it follows that  $c = 1$ .

We next give a proof of Bertrand's Postulate developed about ten years ago (L. Moser). To make the proof go more smoothly we only prove the somewhat weaker

### Theorem 6: Bertrand's Postulate

For every integer  $r$  there exists a prime  $p$  with

$$3 \cdot 2^{2r-1} < p < 3 \cdot 2^{2r}.$$

We restate several of our lemmas in the form in which they will be used.

1. If  $n < p < 2n$  then  $p$  occurs exactly once in  $\binom{2n}{n}$ .
2. If  $2 \cdot 2^{2r-1} < p < 3 \cdot 2^{2r-1}$  then  $p$  does not occur in  $\binom{3 \cdot 2^{2r}}{3 \cdot 2^{2r-1}}$ .
3. If  $p > 2^{r+1}$  then  $p$  occurs at most once in  $\binom{3 \cdot 2^{2n}}{3 \cdot 2^{2n-1}}$ .
4. No prime occurs more than  $2r + 1$  times in  $\binom{3 \cdot 2^{2r}}{3 \cdot 2^{2r-1}}$ .

We now compare  $\binom{3 \cdot 2^{2r}}{3 \cdot 2^{2r-1}}$  and

$$\binom{2^{2r}}{2^{2r-1}} \binom{2^{2r-1}}{2^{2r-2}} \cdots \binom{2}{1} \binom{2^{r+1}}{2^r} \binom{2^r}{2^{r-1}} \cdots \binom{2}{1}^{2r}.$$

Assume that there is no prime in the range  $3 \cdot 2^{2r-1} < p < 3 \cdot 2^{2r}$ . Then, by our lemmas, we find that every prime that occurs in the first expression also occurs in the second with at least as high a multiplicity; that is, the second expression is not smaller than the first. On the other hand, observing that for  $r \geq 6$



$$3 \cdot 2^{2r} > (2^{2r} + 2^{2r-1} + \cdots + 2) + 2r(2^{r+1} + 2r + \cdots + 2),$$

and interpreting  $\binom{2n}{n}$  as the number of ways of choosing  $n$  objects from  $2n$ , we conclude that the second expression is indeed smaller than the first. This contradiction proves the theorem when  $r > 6$ . The primes 7, 29, 97, 389, and 1543 show that the theorem is also true for  $r \leq 6$ .

The proof of Bertrand's Postulate by this method is left as an exercise.

Bertrand's Postulate may be used to prove the following results.

(1)  $\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$  is never an integer.

(2) Every integer  $> 6$  can be written as the sum of distinct primes.

(3) Every prime  $p_n$  can be expressed in the form

$$p_n = \pm 2 \pm 3 \pm 5 \pm \cdots \pm p_{n-1}$$

with an error of at most 1 (Scherk).

(4) The equation  $\pi(n) = \varphi(n)$  has exactly 8 solutions.

About 1949 a sensation was created by the discovery by Erdős and Selberg of an elementary proof of the Prime Number Theorem. The main new result was the following estimate, proved in an elementary manner:

$$\sum_{p \leq x} \log^2 p + \sum_{pq \leq x} \log p \log q = 2x \log x + O(x).$$

Although Selberg's inequality is still far from the Prime Number Theorem, the latter can be deduced from it in various ways without recourse to any further number theoretical results. Several proofs of this lemma have been given, perhaps the simplest being due to Tatzawa and Iseki. Selberg's original proof depends on consideration of the functions

$$\lambda_n = \lambda_{n,x} = \sum_{d|n} \mu(d) \log^2 \frac{x}{d}$$

and

$$T(x) = \sum_{n=1}^x \lambda_n x^n.$$

Some five years ago J. Lambek and L. Moser showed that one could prove Selberg's lemma in a completely elementary way, i.e., using properties of integers only. One of the main tools for doing this is the following rational analogue of the logarithm function. Let

$$h(n) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \quad \text{and} \quad \ell_k(n) = h(kn) - h(k).$$

We prove in an quite elementary way that

$$|\ell_k(ab) - \ell_k(a) - \ell_k(b)| < \frac{1}{k}.$$

The results we have established are useful in the investigation of the magnitude of the arithmetic functions  $\sigma_k(n)$ ,  $\varphi_k(n)$  and  $\omega_k(n)$ . Since these depend not only on the magnitude of  $n$  but also strongly on the arithmetic structure of  $n$  we cannot expect to approximate them by the elementary functions of analysis. Nevertheless we shall see that "on the average" these functions have a rather simple behavior.

If  $f$  and  $g$  are functions for which

$$f(1) + f(2) + \cdots + f(n) \sim g(1) + g(2) + \cdots + g(n)$$

we say that  $f$  and  $g$  have the same average order. We will see, for example, that the average order of  $\tau(n)$  is  $\log n$ , that of  $\sigma(n)$  is  $\frac{\pi^2}{6}n$  and that of  $\varphi(n)$  is  $\frac{6}{\pi^2}n$ .

Let us consider first a purely heuristic argument for obtaining the average value of  $\sigma_k(n)$ . The probability that  $r | n$  is  $\frac{1}{r}$  and if  $r | n$  then  $\frac{n}{r}$  contributes  $(\frac{n}{r})^k$  to  $\sigma_k(n)$ . Thus the expected value of  $\sigma_k(n)$  is

$$\begin{aligned} \frac{1}{1} \left(\frac{n}{1}\right)^k + \frac{1}{2} \left(\frac{n}{2}\right)^k + \dots + \frac{1}{n} \left(\frac{n}{n}\right)^k \\ = n^k \left( \frac{1}{1^{k+1}} + \frac{1}{2^{k+1}} + \dots + \frac{1}{n^{k+1}} \right) \end{aligned}$$

For  $k = 0$  this will be about  $n \log n$ . For  $n \geq 1$  it will be about  $n^k \zeta(k+1)$ , e.g., for  $n = 1$  it will be about  $n \zeta(2) = n \frac{\pi^2}{6}$ .

Before proceeding to the proof and refinement of some of these results we consider some applications of the inversion of order of summation in certain double sums.

Let  $f$  be an arithmetic function and associate with it the function

$$F(n) = \sum_{d=1}^n f(d)$$

and  $g = f \times I$ , i.e.,

$$g(n) = \sum_{d|n} f(d).$$

We will obtain two expressions for

$$\mathcal{F}(x) = \sum_{n=1}^x g(n).$$

$\mathcal{F}(x)$  is the sum

$$\begin{array}{ccccccccccc} & & f(1) & & & & & & & & \\ + & f(1) & + & f(2) & & & & & & & \\ + & f(1) & & & + & f(3) & & & & & \\ + & f(1) & + & f(2) & & & + & f(4) & & & \\ + & f(1) & & & & & & & + & f(5) & \\ + & f(1) & + & f(2) & + & f(3) & & & & + & f(6) \end{array}$$

Adding along vertical lines we have

$$\sum_{d=1}^x \left[ \frac{x}{d} \right] f(d);$$

if we add along the successive diagonal lines each beginning with  $f(1)$  and with “slopes”  $-1, -2, -3, \dots$ , we obtain

$$\sum_{n=1}^x F\left(\left[ \frac{x}{n} \right] \right).$$

Thus

$$\sum_{n=1}^x \sum_{d|n} f(d) = \sum_{d=1}^x \left[ \frac{x}{d} \right] f(d) = \sum_{n=1}^x F\left(\left[ \frac{x}{n} \right] \right).$$

The special case  $f = \mu$  yields

$$1 = \sum_{d=1}^x \mu(d) \left[ \frac{x}{d} \right] = \sum_{n=1}^x M\left(\left[ \frac{x}{n} \right] \right),$$

which we previously considered.

From

$$\sum_{d=1}^x \mu(d) \left[ \frac{x}{d} \right] = 1,$$

we have, on removing brackets, allowing for error, and dividing by  $x$ ,

$$\left| \sum_{d=1}^x \frac{\mu(d)}{d} \right| < 1.$$

Actually, it is known that

$$\sum_{d=1}^{\infty} \frac{\mu(d)}{d} = 0,$$

but a proof of this is as deep as that of the prime number theorem.

Next we consider the case  $f = 1$ . Here we obtain

$$\sum_{n=1}^x \tau(n) = \sum_{n=1}^x \left\lfloor \frac{x}{n} \right\rfloor = x \log x + O(x).$$

In the case  $f = I_k$  we find that

$$\sum_{n=1}^x \sigma_k(n) = \sum_{d=1}^x d^k \left\lfloor \frac{x}{d} \right\rfloor = \sum_{n=1}^x (1^k + 2^k + \cdots + \left\lfloor \frac{x}{n} \right\rfloor).$$

In the case  $f = \varphi$ , recalling that  $\sum_{d|n} \varphi(d) = n$ , we obtain

$$\frac{x(x+1)}{2} = \sum_{n=1}^x \sum_{d|n} \varphi(d) = \sum_{d=1}^x \left\lfloor \frac{x}{d} \right\rfloor \varphi(d) = \sum_{n=1}^x \Phi\left(\frac{x}{n}\right),$$

where  $\Phi(n) = \sum_{d=1}^n \varphi(d)$ . From this we easily obtain

$$\sum_{d=1}^x \frac{\varphi(d)}{d} \geq \frac{x+1}{2},$$

which reveals that, on the average,  $\varphi(d) > \frac{d}{2}$ .

One can also use a similar inversion of order of summation to obtain the following important second Möbius inversion formula:

#### Theorem

If  $G(x) = \sum_{n=1}^x F\left(\left\lfloor \frac{x}{n} \right\rfloor\right)$  then  $F(x) = \sum_{n=1}^x \mu(n) G\left(\left\lfloor \frac{x}{n} \right\rfloor\right)$ .

#### Proof

$$\begin{aligned} \sum_{n=1}^x \mu(n) G\left(\left\lfloor \frac{x}{n} \right\rfloor\right) &= \sum_{n=1}^x \mu(n) \sum_{m=1}^{\left\lfloor \frac{x}{n} \right\rfloor} F\left(\left\lfloor \frac{x}{mn} \right\rfloor\right) \\ &= \sum_{k=1}^x F\left(\left\lfloor \frac{x}{k} \right\rfloor\right) \sum_{n|k} \mu(n) = F(x). \end{aligned}$$

Consider again our estimate

$$\tau(1) + \tau(2) + \cdots + \tau(n) = n \log n + O(n).$$

It is useful to obtain a geometric insight into this result. Clearly  $\tau(r)$  is the number of lattice points on the hyperbola  $xy = r$ ,  $x > 0$ ,  $y > 0$ . Also, every lattice point  $(x, y)$ ,  $x > 0$ ,  $y > 0$ ,  $xy \leq n$ , lies on some hyperbola  $xy = r$ ,  $r \leq n$ . Hence

$$\sum_{n=1}^n \tau(r)$$

is the number of lattice points in the region  $xy \leq n$ ,  $x > 0$ ,  $y > 0$ . If we sum along vertical lines  $x = 1, 2, \dots, n$  we obtain again

$$\tau(1) + \tau(2) + \cdots + \tau(n) = \left\lfloor \frac{n}{1} \right\rfloor + \left\lfloor \frac{n}{2} \right\rfloor + \cdots + \left\lfloor \frac{n}{n} \right\rfloor.$$

In this approach, the symmetry of  $xy = n$  about  $x = y$  suggests how to improve this estimate and obtain a smaller error term.

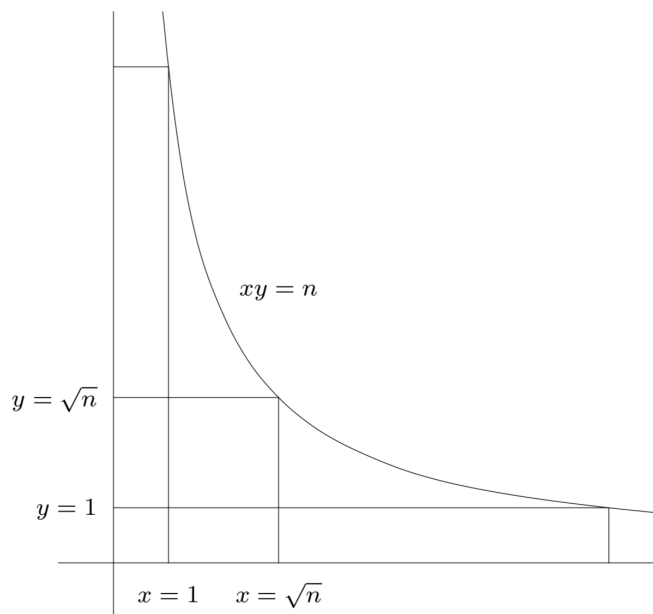


FIGURE 1

Using the symmetry of the above figure, we have, with  $u = \lfloor \sqrt{n} \rfloor$  and  $h(n) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$ ,

$$\begin{aligned} \sum_{r=1}^n \tau(r) &= 2\left(\left[\frac{n}{1}\right] + \cdots + \left[\frac{n}{u}\right]\right) - u^2 \\ &= 2nh(u) - n + O(\sqrt{n}) \\ &= 2n \log \sqrt{u} + (2\gamma - 1)n + O(\sqrt{n}) \\ &= n \log n + (2\gamma - 1)n + O(\sqrt{n}). \end{aligned}$$

Proceeding now to  $\sum \sigma(r)$  we have

$$\sigma(1) + \sigma(2) + \cdots + \sigma(n) = 1 \left[\frac{n}{1}\right] + 2 \left[\frac{n}{2}\right] + \cdots + n \left[\frac{n}{n}\right].$$

In order to obtain an estimate of  $\sum_{n=1}^x \sigma(n)$  set  $k = 1$  in the identity (obtained earlier)

$$\sigma_k(1) + \sigma_k(2) + \cdots + \sigma_k(x) = \sum_{n=1}^x (1^k + 2^k + \cdots + \left[\frac{x}{n}\right]^k).$$

We have immediately

$$\begin{aligned} \sigma(1) + \sigma(2) + \cdots + \sigma(x) &= \frac{1}{2} \sum_{n=1}^x \left[\frac{x}{n}\right] \left[\frac{x}{n} + 1\right] \\ &= \frac{1}{2} \sum_{n=1}^x \frac{x^2}{n^2} + O(x \log x) = \frac{x^2 \zeta(2)}{2} + O(x \log x) \\ &= \frac{\pi^2 x^2}{12} + O(x \log x). \end{aligned}$$

To obtain similar estimates for  $\varphi(n)$  we note that  $\varphi(r)$  is the number of lattice points that lie on the line segment  $x = r$ ,  $0 < y < r$ , and can be seen from the origin. (A point  $(x, y)$  can be seen if  $(x, y) = 1$ .) Thus  $\varphi(1) + \varphi(2) + \cdots + \varphi(n)$  is the number of visible lattice points in the region with  $n > x > y > 0$ .

Let us consider a much more general problem, namely to estimate the number of visible lattice points in a large class of regions.

Heuristically we may argue as follows. A point  $(x, y)$  is invisible by virtue of the prime  $p$  if  $p \mid x$  and  $p \mid y$ . The probability that this occurs is  $\frac{1}{p^2}$ . Hence the probability that the point is invisible is

$$\prod_p \left(1 - \frac{1}{p^2}\right) = \prod_p \left(1 + \frac{1}{p^2} + \frac{1}{p^4} + \cdots\right)^{-1} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}.$$

Thus the number of visible lattice point should be  $\frac{6}{\pi^2}$  times the area of the region. In particular the average order of  $\varphi(n)$  should be about  $\frac{6}{\pi^2}n$ .

We now outline a proof of the fact that in certain large regions the fraction of visible lattice points contained in the region is approximately  $\frac{6}{\pi^2}$ .

Let  $R$  be a region in the plane having finite Jordan measure and finite perimeter. Let  $tR$  denote the region obtained by magnifying  $R$  radially by  $t$ . Let  $M(tR)$  be the area of  $tR$ ,  $L(tR)$  the number of lattice points in  $tR$ , and  $V(tR)$  the number of visible lattice points in  $tR$ .

It is intuitively clear that

$$L(tR) = M(tR) + O(t) \quad \text{and} \quad M(tR) = t^2 M(R).$$

Applying the inversion formula to

$$L(tR) = V(tR) + V\left(\frac{t}{2}R\right) + V\left(\frac{t}{3}R\right) + \cdots$$

we find that

$$\begin{aligned} &= \sum_{d=1}^{\infty} L\left(\frac{t}{d}R\right)\mu(d) = \sum_{d=1}^{\infty} M\left(\frac{t}{d}R\right)\mu(d) \\ &\approx M(tR) \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} \approx M(tR) \frac{6}{\pi^2} = t^2 M(R) \frac{6}{\pi^2}. \end{aligned}$$

With  $t = 1$  and  $R$  the region  $n > x > y > 0$ , we have

$$\varphi(1) + \varphi(2) + \cdots + \varphi(n) \approx \frac{n^2}{2} \cdot \frac{6}{\pi^2} = \frac{3}{\pi^2} n^2$$

It has been shown (Chowla) that the error term here cannot be reduced to  $O(n \log \log \log n)$ . Walfitz has shown that it can be replaced by  $O(n \log^{\frac{3}{4}} n)$ .

Erdddots and Shapiro have shown that

$$\varphi(1) + \varphi(2) + \cdots + \varphi(n) - \frac{3}{\pi^2} n^2$$

changes sign infinitely often.

We will later make an application of our estimate of  $\varphi(1) + \varphi(2) + \cdots + \varphi(n)$  to the theory of distributions of quadratic residues.

Our result can also be interpreted as saying that if a pair of integers  $(a, b)$

are chosen at random the probability that they will be relatively prime is  $\frac{6}{\pi^2}$ .

At this stage we state without proof a number of related results.

At this stage we state without proof a number of related results.

If  $(a, b)$  are chosen at random the expected value of  $(a, b)$  is  $\frac{\pi^2}{6}$ .

If  $f(x)$  is one of a certain class of arithmetic functions that includes  $x^\alpha$ ,  $0 < \alpha < 1$ , then the probability that  $(x, f(x)) = 1$  is  $\frac{6}{\pi^2}$ , and its expected value is  $\frac{\pi^2}{6}$ .

This and related results were proved by Lambek and Moser.

The probability that  $n$  numbers chosen at random are relatively prime is  $\frac{1}{\zeta(n)}$ .

The number  $Q(n)$  of quadratfrei numbers under  $n$  is  $\sim \frac{6}{\pi^2} n$  and the number  $O_k(n)$  of  $k$ th power-free numbers under  $n$  is  $\frac{n}{\zeta(k)}$ .  
The first result follows almost immediately from

$$\sum Q\left(\frac{n^2}{r^2}\right) = n^2,$$

so that by the inversion formula

$$Q(n^2) = \sum \mu(r) \left[\frac{n}{r}\right]^2 \sim n^2 \zeta(2).$$

A more detailed argument yields

$$Q(x) = \frac{6x}{\pi^2} + O(\sqrt{x}).$$

Another rather amusing related result, the proof of which is left as an exercise, is that

$$\sum_{(a,b)=1} \frac{1}{a^2 b^2} = \frac{5}{2}.$$

The result on  $Q(x)$  can be written in the form

$$\sum_{n=1}^x |\mu(n)| \sim \frac{6}{\pi^2} x$$

One might ask for estimates for

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

Indeed, it is known (but difficult to prove) that  $M(x) = o(x)$ .

Let us turn our attention to  $\omega(n)$ . We have

$$\omega(1) + \omega(2) + \cdots + \omega(n) = \sum_{p \leq n} \left[\frac{n}{p}\right] \sim n \log \log n.$$

Thus the average value of  $\omega(n)$  is  $\log \log n$ .

The same follows in a similar manner for  $\Omega(n)$

A relatively recent development along these lines, due to Erdős, Kac, Leveque, Tatum and others is a number of theorems of which the following is typical.

Let  $A(x; \alpha, \beta)$  be the number of integers  $n \leq x$  for which

$$\alpha \sqrt{\log \log n} + \log \log n < \omega(n) < \beta \sqrt{\log \log n} + \log \log n.$$

Then

$$\lim_{x \rightarrow \infty} \frac{A(x; \alpha, \beta)}{x} = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{u^2}{2}} du.$$

Thus we have for example that  $\omega(n) < \log \log n$  about half the time.

One can also Prove (Tatum) that a similar result holds for  $B(x; \alpha, \beta)$ , the number of integers in the set  $f(1), f(2), \dots, f(x)$  ( $f(x)$  is an irreducible polynomial with integral coefficients) for which  $\omega(f(n))$  lies in a range similar to those prescribed for  $\omega(n)$ .

Another type of result that has considerable applicability is the following.

The number  $C(x, \alpha)$  of integers  $\leq x$  having a prime divisor  $> x\alpha$ ,  $1 > \alpha > \frac{1}{2}$ , is  $\sim -x \log \alpha$ . In fact, we have

$$\begin{aligned} C(x, \alpha) &= \sum_{x^{\alpha \log p} < p < x} \frac{x}{p} \sim x \sum_{x^{\alpha \log p} < p < x} \frac{1}{p} \\ &= x(\log \log x - \log \log \alpha) \\ &= x(\log \log x - \log \log x - \log \alpha) = -x \log \alpha. \end{aligned}$$

For example the density of numbers having a prime factor exceeding their square root is  $\log 2 \approx .7$ .

Thus far we have considered mainly average behavior of arithmetic functions. We could also inquire about absolute bounds. One can prove without difficulty that

$$1 > \frac{\varphi(n)\sigma(n)}{n^2} > \epsilon > 0 \text{ for all } n.$$

Also, it is known that

$$n > \varphi(n) > \frac{cn}{\log \log n}$$

and

$$n < \sigma(n) < cn \log \log n.$$

As for  $\tau(n)$ , it is not difficult to show that

$$\tau(n) > (\log n)^k$$

infinitely often for every  $k$  while  $\tau(n) < n^\epsilon$  for every  $\epsilon$  and  $n$  sufficient large.

We state but do not prove the main theorem along these lines.

If  $\epsilon > 0$  then

$$\tau(n) < 2^{(1+\epsilon) \log n / \log \log n} \text{ for all } n > n_0(\epsilon)$$

while

$$\tau(n) > 2^{(1-\epsilon) \log n / \log \log n} \text{ infinitely often.}$$

A somewhat different type of problem concerning average value of arithmetic functions was the subject of a University of Alberta master's thesis of Mr. R. Trollope a couple of years ago.

Let  $s_r(n)$  be the sum of the digits of  $n$  when written in base  $r$ . Mirsky has proved that

$$s_r(1) + s_r(2) + \cdots + s_r(n) = \frac{r-1}{2} n \log_r n + O(n).$$

Mr. Trollope considered similar sums where the elements on the left run over certain sequences such as primes, squares, etc.

Still another quite amusing result he obtained states that

$$\frac{s_1(n) + s_2(n) + \cdots + s_n(n)}{n^2} \sim 1 - \frac{\pi^2}{12}.$$

---

This page titled [1.3: Distribution of Primes](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Leo Moser \(The Trilla Group\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

## 1.4: Irrational Numbers

The best known of all irrational numbers is  $\sqrt{2}$ . We establish  $\sqrt{2} \neq \frac{a}{b}$  with a novel proof which does not make use of divisibility arguments.

Suppose  $\sqrt{2} = \frac{a}{b}$  ( $a, b$  integers), with  $b$  as small as possible. Then  $b < a < 2b$  so that

$$\frac{2ab}{ab} = 2, \frac{a^2}{b^2} = 2, \text{ and } \frac{2ab - a^2}{ab - b^2} = 2 = \frac{a(2b - a)}{b(a - b)}.$$

Thus

$$\sqrt{2} = \frac{2b - a}{a - b}.$$

But  $a < 2b$  and  $a - b < b$ ; hence we have a rational representation of  $\sqrt{2}$  with denominator smaller than the smallest possible!

To convince students of the existence of irrationals one might begin with a proof of the irrationality of  $\log_{10} 2$ . If  $\log_{10} 2 = \frac{a}{b}$  then  $10^{a/b} = 2$  or  $10^a = 2^b$ . But now the left hand side is divisible by 5 while the right hand side is not.

Also not as familiar as it should be is the fact that  $\cos 1^\circ$  (and  $\sin 1^\circ$ ) is irrational. From

$$\cos 45^\circ + i \sin 45^\circ = (\cos 1^\circ + i \sin 1^\circ)^{45}$$

we deduce that  $45^\circ$  can be expressed as a polynomial in integer coefficients in  $\cos 1^\circ$ . Hence if  $\cos 1^\circ$  were rational so would be  $\cos 45^\circ = \frac{1}{\sqrt{2}}$ .

The fact that

$$\cos 1 = 1 - \frac{1}{2!} + \frac{1}{4!} - \dots$$

is irrational can be proved in the same way as the irrationality of  $e$ . In the latter case, assuming  $e$  rational,

$$\frac{b}{a} = e = 1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{(a+1)!} + \frac{1}{(a+2)!} + \dots,$$

which, after multiplication by  $a!$ , would imply that  $\frac{1}{a+1} + \frac{1}{(a+1)(a+2)} + \dots$  is a positive integer less than 1.

A slightly more complicated argument can be used to show that  $e$  is not of quadratic irrationality, i.e., that if  $a, b, c$  are integers then  $ae^2 + be + c \neq 0$ . However, a proof of the transcendentality of  $e$  is still not easy. The earlier editions of Hardy and Wright claimed that there was no easy proof that  $\pi$  is transcendental but this situation was rectified in 1947 by I. Niven whose proof of the irrationality of  $\pi$  we now present.

Let

$$\pi = \frac{a}{b}, f(x) = \frac{x^n(a-by)^n}{n!}, \text{ and } F(x) = f(x) - f^{(2)}(x) + f^{(4)}(x) - \dots,$$

the positive integer  $n$  being specified later. Since  $n!f(x)$  has integral coefficients and terms in  $x$  of degree  $\leq 2n$ ,  $f(x)$  and all its derivatives will have integral values at  $x = 0$ . Also for  $x = \pi = \frac{a}{b}$ , since  $f(x) = f(\frac{a}{b} - x)$ . By elementary calculus we have

$$\frac{d}{dx}[F'(x) \sin x - F(x) \cos x] = F''(x) \sin x + F(x) \sin x = f(x) \sin x.$$

Hence

$$\int_0^\pi f(x) \sin x dx = [F'(x) - F(x) \cos x]_0^\pi = \text{an integer}.$$

However, for  $0 < x < \pi$ ,

$$0 < f(x) \sin x < \frac{\pi^n a^n}{n!} \rightarrow 0$$



for large  $n$ . Hence the definite integral is positive but arbitrarily small for large  $n$ ; this contradiction shows that the assumption  $\pi = \frac{a}{b}$  is untenable.

This proof has been extended in various ways. For example, Niven also proved that the cosine of a rational number is irrational. If now  $\pi$  were rational,  $\cos \pi = -1$  would be irrational. Further, the method can also be used to prove the irrationality of certain numbers defined as the roots of the solutions of second order differential equations satisfying special boundary conditions. Recently, a variation of Niven's proof has been given which, although more complicated, avoids the use of integrals or infinite series. A really simple proof that  $\pi$  is transcendental, i.e., does not satisfy any polynomial equation with integer coefficients is still lacking.

With regard to transcendental numbers there are essentially three types of problems: to prove the existence of such numbers, to construct such numbers, and finally (and this is much more difficult than the first two) to prove that certain numbers which arise naturally in analysis are transcendental. Examples of numbers which have been proved transcendental are  $\pi$ ,  $e$ ,  $e^{-\pi}$ , and  $\frac{\log 3}{\log 2}$ . It is interesting to remark here that Euler's constant  $\gamma$  and

$$\sum_{n=1}^{\infty} \frac{1}{n^{2s+1}} \quad (s \text{ is an integer})$$

have not even been proved irrational.

Cantor's proof of the existence of transcendental numbers proceeds by showing that the algebraic numbers are countable while the real numbers are not. Thus every uncountable set of numbers contains transcendental numbers. For example there is a transcendental number of the form  $e^{i\theta}$ ,  $0 < \theta < \frac{\pi}{2}$ , say.

Although it is not entirely relevant here we will perform now a little disappearing stunt using such a transcendental number  $e^{i\theta}$  and a construction due to Kuratowski.

Consider the following set of points in the complex plane. Start with the point  $O$  and let  $\tilde{S}$  be the set of all points obtainable from it by a succession of the operations of translating the points 1 unit to the right and rotating them through an angle  $\theta$  about  $O$ . If we denote such translations and rotations by  $T$  and  $R$  respectively then a typical point of our set  $\tilde{S}$  may be denoted by  $T^a R^b T^c R^d \dots$ . We next observe that every point of  $\tilde{S}$  must have a unique representation in this form. Indeed,  $T$  means adding 1 to the complex number corresponding to the point and  $R$  means multiplication by  $e^{i\theta}$ . Hence all our points are polynomials in  $e^{i\theta}$  with positive coefficients, say  $z = P(e^{i\theta})$ . But now if a point has a double representation, then  $P(e^{i\theta}) = R(e^{i\theta})$  and we would obtain a polynomial in  $e^{i\theta}$  which would negate the transcendental character of  $e^{i\theta}$ .

Let  $\tilde{T}$  denote the subset of  $\tilde{S}$  which consists of those points of  $\tilde{S}$  for which the last operation needed to reach them is a  $T$ , and let  $\tilde{R}$  denote the subset which consist of those points of  $\tilde{S}$  for which the last operation needed to reach them is an  $R$ . Clearly  $\tilde{S} = \tilde{T} \cup \tilde{R}$  and  $\tilde{T} \cap \tilde{R} = \emptyset$ . A translation of  $\tilde{S}$  of one unit to the right sends  $\tilde{S}$  into  $\tilde{T}$ , i.e., it makes  $\tilde{R}$  vanish! On the other hand, a rotation of the plane through  $\theta$  sends  $\tilde{S}$  into  $\tilde{R}$  making  $\tilde{T}$  vanish!

So far we have discussed only the existence of transcendental numbers. The easiest approach to the actual construction of such numbers is via a theorem due to Liouville.

We say that an algebraic number is of degree  $n$  if it satisfies a polynomial equation of degree  $n$ . We say that a real number  $\lambda$  is approximable to order  $n$  provided the inequality

$$\left| \lambda - \frac{a}{b} \right| < \frac{c}{b^n}$$

has an infinity of solutions for some constant  $c$ . Liouville's theorem states that a real algebraic number of degree  $n$  is not approximable to any order greater than  $n$ .

Suppose  $\lambda$  is of degree  $n$ . Then it satisfies an equation

$$f(\lambda) = a_0 \lambda^n + a_1 \lambda^{n-1} + \dots + a_n = 0.$$

There is a number  $M = M(\lambda)$  such that  $|f'(x)| < M$  where  $\lambda - 1 < x < \lambda + 1$ . Suppose now that  $\frac{p}{q} \neq \lambda$  is an approximation to  $\lambda$ . We may assume the approximation good enough to ensure that  $\frac{p}{q}$  lies in the interval  $\lambda - 1, \lambda + 1$ ,

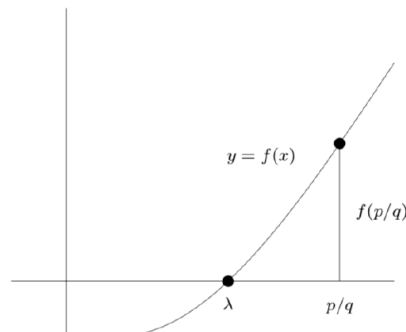


FIGURE 2

and is nearer to  $\lambda$  than any other root of  $f(x) = 0$ , so that  $f(p/q) \neq 0$ .

Clearly (see Figure 2),

$$\left|f\left(\frac{p}{q}\right)\right| = \frac{1}{q^n} |a_0 p^n + a_1 p^{n-1} q + \cdots + a_n q^n| \geq \frac{1}{q^n}$$

and

$$\left|\frac{f(p/q)}{\lambda - p/q}\right| < M$$

so that

$$\left|\lambda - \frac{p}{q}\right| > \frac{c}{q^n}$$

and the theorem is proved.

Although Liouville's theorem suffices for the construction of many transcendental numbers, much interest centers on certain refinements. In particular, it is desirable to have a theorem of the following type. If  $\lambda$  is of degree  $n$  then

$$\left|\lambda - \frac{p}{q}\right| < \frac{M}{q^{f(n)}}$$

has at most a finite number of solutions. Here  $f(n)$  may be taken as  $n$  by Liouville's theorem. Can it be decreased? Thue, about 1909, first showed that one could take  $f(n) = \frac{n}{2}$  and Siegel (1921) showed that we can take  $f(n) = 2\sqrt{n}$ . This was slightly improved by Dyson and Schneider to  $\sqrt{2n}$ . Very recently (1955), F. K. Roth created a sensation by proving that we can take  $f(n) = 2 + \epsilon$ . His proof is long and complicated. That we cannot take  $f(n) = 2$  (hence Roth's result is in a way best possible) can be seen from the following result due to Dirichlet.

For irrational  $\lambda$  there exist infinitely many solutions of

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

The proof is not difficult. Let  $\lambda$  be irrational and consider, for fixed  $n$ , the numbers  $(\lambda), (2\lambda), \dots, (n\lambda)$ , where  $(x)$  means "fractional part of  $x$ ". These  $n$  points are distinct points on  $[0, 1)$ ; hence there exist two of them say  $i\lambda$  and  $j\lambda$  whose distance apart is  $\leq \frac{1}{n}$ . Thus we have

$$(i\lambda) - (j\lambda) < \frac{1}{n}$$

or

$$k\lambda - m \leq \frac{1}{n} \quad (k \text{ and } m \text{ integers } \leq n)$$

and

$$\left|\lambda - \frac{m}{k}\right| \leq \frac{1}{nk} \leq \frac{1}{n^2},$$

as required.

We now return to the application of Liouville's theorem to the construction of transcendental numbers.

Consider

$$\frac{1}{10^{1!}} + \frac{1}{10^{2!}} + \cdots + \frac{1}{10^{p!}} = \lambda_p$$

as well as the real number  $\lambda = \lambda_\infty$ . It is easily checked that  $|\lambda_\infty - \lambda_p| < \frac{1}{10^{p+1}}$  for every  $p$ . Hence  $\lambda$  is approximable to order  $n$  for any  $n$  and hence is not algebraic.

---

This page titled [1.4: Irrational Numbers](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Leo Moser \(The Trilla Group\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

## 1.5: Congruences

In this section we shall develop some aspects of the theory of divisibility and congruences.

If

$$a = \prod p^\alpha \text{ and } b = \prod p^\beta$$

then it is easily seen that

$$(a, b) = \prod p^{\min(\alpha, \beta)} \text{ while } [a, b] = \prod p^{\max(\alpha, \beta)}.$$

From these it follows easily that  $(a, b) \cdot [a, b] = a \cdot b$ . We leave it as an exercise to show that

$$(a, b) = \frac{1}{a} \sum_{a=1}^{a-1} \sum_{\beta=1}^{a-1} e^{2\pi i \frac{b}{a} \alpha \beta}.$$

The notation  $a \equiv b \pmod{m}$  for  $m \mid (a - b)$  is due to Gauss. Rather obvious properties of this congruence are  $a \equiv a$ ,  $a \equiv b \Rightarrow b \equiv a$ , and  $a \equiv b$  and  $b \equiv c \Rightarrow a \equiv c$ , i.e.,  $\equiv$  is an equivalence relation. It is also easily proved that  $a \equiv b$  and  $c \equiv d$  together imply  $ac \equiv bd$ ; in particular  $a \equiv b \Rightarrow ka \equiv kb$ . However the converse is not true in general. Thus  $2 \times 3 = 4 \times 3 \pmod{6}$  does not imply  $2 \equiv 4 \pmod{6}$ . However, if  $(k, m) = 1$  then  $ka \equiv kb$  does imply  $a \equiv b$ .

Another important result is the following.

### Theorem

If  $a_1, a_2, \dots, a_{\varphi(m)}$  form a complete residue system mod  $m$  then so does  $aa_1, aa_2, \dots, aa_{\varphi(m)}$  provided  $(a, m) = 1$ .

### Proof

We have  $\varphi(m)$  residues. If two of them are congruent  $aa_i \equiv aa_j$ ,  $a(a_i - a_j) \equiv 0$ . But  $(a, m) = 1$  so that  $a_i \equiv a_j$ .

An application of these ideas is to the important Euler's theorem:

### Theorem

If  $(a, m) = 1$  then  $a^{\varphi(m)} \equiv 1 \pmod{m}$ .

### Proof

Since  $a_1, a_2, \dots, a_{\varphi(m)}$  are congruent to  $aa_1, aa_2, \dots, aa_{\varphi(m)}$  in some order, their products are congruent. Hence

$$a^{\varphi(m)} a_1 a_2 \cdots a_{\varphi(m)} \equiv a_1 a_2 \cdots a_{\varphi(m)}$$

and the result follows.

A special case of primary importance is the case  $m = p$  where we have

$$(a, p) = 1 \Rightarrow a^{p-1} \equiv 1 \pmod{p}$$

Multiplying by  $a$  we have for all cases  $a^p \equiv a \pmod{p}$ . Another proof of this result goes by induction on  $a$ ; we have

$$(a+1)^p = a^p + \binom{p}{1} a^{p-1} + \cdots + 1 \equiv a^p \equiv a \pmod{p}$$

One could also use the multinomial theorem and consider  $(1+1+\cdots+1)^p$ .

Still another proof goes by considering the number of regular convex  $p$ -gons where each edge can be colored in one of  $a$  colors. The number of such polygons is  $a^p$ , of which  $a$  are monochromatic. Hence  $a^p - a$  are not monochromatic and these come in sets of  $p$  each by rotation through  $\frac{2\pi n}{p}$ ,  $n = 1, 2, \dots, p-1$ . The idea behind this proof has considerable applicability and we shall return to at least one other application a little later. We also leave as an exercise the task of finding a similar proof of Euler's theorem.

The theorems of Fermat and Euler may also be conveniently viewed from a group-theoretic viewpoint. The integers relatively prime to  $m$  and  $< m$  form a group under multiplication mod  $m$ . The main thing to check here is that every element has an inverse

in the system. If we seek an inverse for  $a$  we form  $aa_1, aa_2, \dots, aa_{\varphi(m)}$ . We have already seen that these are  $\varphi(m)$  numbers incongruent mod  $m$  and relatively prime to  $m$ . Thus one of them must be the unit and the result follows.

We now regain Euler's proof from that of Lagrange's which states that if  $a$  is an element of a group  $G$  of order  $m$ , then  $a^m = 1$ . In our case this means  $a^{\varphi(m)} \equiv 1$  or  $a^{p-1} \equiv 1$  if  $p$  is a prime.

The integers under  $p$  form a field with respect to  $+$  and  $\times$ . Many of the important results of number theory are based on the fact that the multiplicative part of this group (containing  $p-1$  elements) is cyclic, i.e., there exists a number  $g$  (called primitive root of  $p$ ) such that  $1 = g^0, g^1, g^2, \dots, g^{p-1}$  are incongruent mod  $p$ . This fact is not trivial but we omit the proof. A more general group-theoretic result in which it is contained states that every finite field is automatically Abelian and its multiplicative group is cyclic.

In the ring of polynomials with coefficients in a field many of the theorems of elementary theory of equations holds. For example if  $f(x)$  is a polynomial whose elements are residue classes mod  $p$  then  $f(x) \equiv 0 \pmod{p}$  has at most  $p$  solutions. Further if  $r$  is a root then  $x - r$  is a factor. On the other hand, it is not true that  $f(x) \equiv 0 \pmod{p}$  has at least one root.

Since  $x^p - x \equiv 0$  has at most  $p$  roots we have the factorization

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

Comparing coefficients of  $x$  we have  $(p-1)! \equiv -1 \pmod{p}$  which is Wilson's theorem.

Cayley gave a geometrical proof of Wilson's theorem as follows. Consider the number of directed  $p$ -gons with vertices of a regular  $p$ -gon. (See Figure 3.)

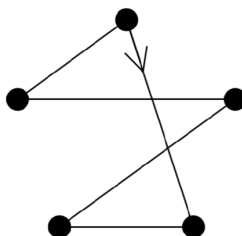


FIGURE 3

There are  $(p-1)!$  in number of which  $p-1$  are regular. Hence the nonregular ones are  $(p-1)! - (p-1)$  in number and these come in sets of  $p$  by rotation. Hence

$$(p-1)! - (p-1) \equiv 0 \pmod{p}$$

and

$$(p-1)! + 1 \equiv 0 \pmod{p}$$

follows.

One can also give a geometrical proof which simultaneously yields both

Fermat's and Wilson's theorems and we suggest as a problem finding such a proof.

Wilson's theorem yields a necessary and sufficient condition for primality:  $p$  is prime if and only if  $(p-1)! \equiv -1$ , but this is hardly a practical criterion.

## Congruences with Given Roots

Let  $a_1, a_2, \dots, a_k$  be a set of distinct residue classes mod  $n$ . If there exists a polynomial with integer coefficients such that  $f(x) \equiv 0 \pmod{n}$  has roots  $a_1, a_2, \dots, a_k$  and no others, we call this set compatible mod  $n$ . Let the number of compatible sets mod  $n$  be denoted by  $C(n)$ . Since the number of subsets of the set consisting of  $0, 1, 2, \dots, n-1$  is  $2^n$ , we call  $c(n) = \frac{C(n)}{2^n}$  the coefficient of compatibility of  $n$ .

If  $n = p$  is a prime then the congruence

$$(x-a_1)(x-a_2) \cdots (x-a_k) \equiv 0 \pmod{p}$$

has precisely the roots  $a_1, a_2, \dots, a_n$ . Hence  $c(p) = 1$ . In a recent paper M. M. Chokjnsacka Pniewska has shown that  $c(4) = 1$  while  $c(n) < 1$  for  $n = 6, 8, 9, 10$ . We shall prove that  $c(n) < 1$  for every composite  $n \neq 4$ . We can also prove that the average value of  $c(n)$  is zero, i.e.,

$$\lim_{n \rightarrow \infty} \frac{1}{n} (c(1) + c(2) + \dots + c(n)) = 0$$

Since  $c(n) = 1$  for  $n = 1$  and  $n = p$  we consider only the case where  $n$  is composite. Suppose then that the unique prime factorization of  $n$  is given by  $p_1^{\alpha_1} p_2^{\alpha_2} \dots$  with

$$p_1^{\alpha_1} > p_2^{\alpha_2} > \dots$$

Consider separately the cases

(1)  $n$  has more than one prime divisor, and

(2)  $n = p^\alpha, \alpha > 1$ .

In case (1) we can write

$$n = a \cdot b, (a, b) = 1, a > b > 1.$$

We now show that if  $f(a) \equiv f(b) \equiv 0$  then  $f(0) \equiv 0$ .

**Proof.** Let  $f(x) = c_0 + c_1x + \dots + c_mx^m$ . Then

$$0 \equiv af(b) \equiv ac_0 \pmod{n} \text{ and}$$

$$0 \equiv bf(a) \equiv bc_0 \pmod{n}.$$

Now since  $(a, b) = 1$  there exist  $r, s$ , such that  $ar + bs = 1$  so that  $c_0(ar + bs) \equiv 0$  and  $c_0 \equiv 0$ .

In case (2) we can write

$$n = p^{\alpha-1}p.$$

We show that  $f(p^{\alpha-1}) \equiv 0$  and  $f(0) \equiv 0$  imply  $f(kp^{\alpha-1}) \equiv 0, k = 2, 3, \dots$

**Proof.** Since  $f(0) \equiv 0$ , we have  $c_0 \equiv 0$ ,

$$f(x) \equiv c_1x + c_2x^2 + \dots + c_mx^m, \text{ and}$$

$$f(p^{\alpha-1}) \equiv c_1p^{\alpha-1} \equiv 0 \pmod{p^\alpha},$$

so that  $c_1 \equiv 0 \pmod{p}$ . But now  $f(kp^{\alpha-1}) \equiv 0$ , as required.

## On Relatively Prime Sequences Formed by Iterating Polynomials (Lambek and Moser)

Bellman has recently posed the following problem. If  $p(x)$  is an irreducible polynomial with integer coefficients and  $p(x) > x$  for  $x > c$ , prove that  $\{p^n(c)\}$  cannot be prime for all large  $n$ . We do not propose to solve this problem but wish to make some remarks.

If  $p(x)$  is a polynomial with integer coefficients, then so is the  $k^{\text{th}}$  iterate defined recursively by  $p^0(x) = x, p^{p+1}(x) = p(p^k(x))$ . If  $a$  and  $b$  are integers then

$$p^k(a) \equiv p^k(b) \pmod{(a-b)} \quad (1.5.1)$$

In particular for  $a = p^n(c)$  and  $b = 0$  we have  $p^{k+n}(c) \equiv p^k(0) \pmod{p^n(c)}$ .

Hence

$$(p^{k+n}(c), p^n(c)) = (p^k(0), p^n(c)). \quad (1.5.2)$$

We shall call a sequence  $\{a_n\}, n \geq 0$ , relatively prime if  $(a_m, a_n) = 1$  for all values of  $m, n$ , with  $m \neq n$ . From 5.2 we obtain

### Theorem 1

$\{p^n(c)\}, n \geq 0$ , is relatively prime if and only if  $(p^k(0), p^n(c)) = 1$  for all  $k \geq 0, n \geq 0$

From the follows immediately a result of Bellman: If  $p^k(0) = p(0) \neq 1$  for  $k \geq 1$  and if  $(a, p(0)) = 1$  implies  $(p(a), p(0)) = 1$  then  $\{p^n(a)\}$ ,  $n \geq a$  is relatively prime whenever  $(c, p(0)) = 1$ .

We shall now construct all polynomials  $p(x)$  for which  $\{p^n(c)\}$ ,  $n \geq 0$ , is relatively prime for all  $c$ . According to Theorem 1,  $p^k(0) = \pm 1$  for all  $k \geq 1$ , as is easily seen by taking  $n = k$  and  $c = 0$ . But then  $\{p^k(0)\}$  must be one of the following six sequences:

$$1, 1, 1, \dots$$

$$1, -1, 1, \dots$$

$$1, -1, -1, \dots$$

$$-1, 1, 1, \dots$$

$$-1, 1, -1, \dots$$

$$-1, -1, 1, \dots$$

it is easily seen that the general solution of  $p(x)$  (with integer coefficients) of  $m$  equations

$$p(a_k) = a_{k+1}, \quad k = 0, 1, 2, \dots, m-1.$$

is obtained from a particular solution  $p_1(x)$  as follows.

$$p(x) = p_1(x) + (x - a_1)(x - a_2) \cdots (x - a_{m-1}) \cdot Q(x), \quad (1.5.3)$$

\

where  $Q(x)$  is any polynomial with integer coefficients.

### Theorem 2

$\{p^n(c)\}$ ,  $n \geq 0$ , is relatively prime for all  $c$  if and only if  $p(x)$  belongs to one of the following six classes of polynomials.

$$1 + x(x-1) \cdot Q(x)$$

$$1 - x - x^2 + x(x^2 - 1) \cdot Q(x)$$

$$1 - 2x^2 + x(x^2 - 1) \cdot Q(x)$$

$$2x^2 - 1 + x(x^2 - 1) \cdot Q(x)$$

$$x^2 - x - 1 + x(x^2 - 1) \cdot Q(x)$$

$$-1 + x(x+1) \cdot Q(x)$$

### Proof

In view of (5.3) we need only verify that the particular solutions yield the six sequences given above.

## On the Distribution of Quadratic Residues

A large segment of number theory can be characterized by considering it to be the study of the first digit on the right of integers. Thus, a number is divisible by  $n$  if its first digit is zero when the number is expressed in base  $n$ . Two numbers are congruent (mod  $n$ ) if their first digits are the same in base  $n$ . The theory of quadratic residues is concerned with the first digits of the squares. Of particular interest is the case where the base is a prime, and we shall restrict ourselves to this case.

If one takes for example  $p = 7$ , then with congruences (mod 7) we have  $1^2 \equiv 1 \equiv 6^2$ ,  $2^2 \equiv 4 \equiv 5^2$ , and  $3^2 \equiv 2 \equiv 4^2$ ; obviously,  $0^2 \equiv 0$ . Thus 1, 2, 4 are squares and 3, 5, 6 are nonsquares or nonresidues. If  $a$  is a residue of  $p$  we write

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

while if  $a$  is a nonresidue we write

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

For  $p \mid c$  we write  $\left(\frac{c}{p}\right) = 0$ . This notation is due to Legendre. For  $p = 7$  the sequence  $\left(\frac{a}{p}\right)$  is thus

+ + - - -.

For  $p = 23$  it turns out to be

+ + + + - - + + - - + + - - -.

The situation is clarified if we again adopt the group theoretic point of view. The residue classes (mod  $p$ ) form a field, whose multiplicative group (containing  $p - 1$  elements) is cyclic. If  $g$  is a generator of this group then the elements may be written  $g^1, g^2, \dots, g^{p-1} = 1$ . The even powers of  $g$  are the quadratic residues; they form a subgroup of index two. The odd powers of  $g$  are the quadratic nonresidues. From this point of view it is clear

$$\text{res} \times \text{res} = \text{res}, \text{res} \times \text{nonres} = \text{nonres}, \text{nonres} \times \text{nonres} = \text{res}.$$

Further  $1/a$  represents the unique inverse of  $a$  (mod  $p$ ) and will be a residue or nonresidue according as  $a$  itself is a residue or nonresidue.

The central theorem in the theory of quadratic residues and indeed one of the most central results of number theory is the Law of Quadratic Reciprocity first proved by Gauss about 1800. It states that

It leads to an algorithm for deciding the value of  $\left(\frac{p}{q}\right)$ .

Over 50 proofs of this law have been given including recent proofs by Zassenhaus and by Lehmer. In Gauss' first proof (he gave seven) he made use of the following lemma—which he tells us he was only able to prove with considerable difficulty. For  $p \equiv 1 \pmod{4}$  the least nonresidue of  $p$  does not exceed  $2\sqrt{p} + 1$ . The results we want to discuss today are in part improvements of this result and more generally are concerned with the distribution of the sequence of + and - 's in  $\left(\frac{a}{p}\right)$ ,  $a = 1, 2, \dots, p - 1$ .

In 1839 Dirichlet, as a by-product of his investigation of the class number of quadratic forms, established the following theorem: If  $p \equiv 3 \pmod{4}$  then among the integers  $1, 2, \dots, \frac{p-1}{2}$ , there are more residues than nonresidues. Though this is an elementary statement about integers, all published proofs, including recent ones given by Chung, Chowla, Whitman, Carlitz, and Moser involve Fourier series. Landau was quite anxious to have an elementary proof. Though somewhat related results have been given by Whitman and by Carlitz, Dirichlet's result is quite isolated. Thus, no similar nontrivial result is known for other ranges.

In 1896 Aladow, in 1898 von Sterneck, and in 1906 Jacobsthal, took up the question of how many times the combinations ++, +-, -+, and -- appear. They showed that each of the four possibilities appeared, as one might expect, with frequency 1/4. In 1951 Perron examined the question again and proved that similar results hold if, instead of consecutive integers, one considers integers separated by a distance  $d$ . J. B. Kelly recently proved a result that, roughly speaking, shows that the residues and non residues are characterized by this property. Jacobsthal also obtained partial results for the cases of 3 consecutive residues and non residues. Let  $R_n$  and  $N_n$  be the number of blocks of  $n$  consecutive residues and non residues respectively. One might conjecture that  $R_n \sim N_n \sim \frac{p}{2^n}$ . Among those who contributed to this question are Vandiver, Bennet, Dorge, Hopf, Davenport, and A. Brauer. Perhaps the most interesting result is that of A. Brauer. He showed that for  $p > p_0(n)$ ,  $R_n > 0$  and  $N_n > 0$ . We shall sketch part of his proof. It depends on a very interesting result of Van der Waerden (1927). Given  $k, \ell$ , there exists an integer  $N = N(k, \ell)$  such that if one separates the integers  $1, 2, \dots, N$  into  $k$  classes in any manner whatsoever, at least one of the classes will contain an arithmetic progression of length  $\ell$ . There are a number of unanswered questions about this theorem to which we shall return in a later section.

Returning to Brauer's work, we shall show he proves that all large primes have, say, 7 consecutive residues. One separates the numbers  $1, 2, \dots, p - 1$  into 2 classes, residues and nonresidues. If  $p$  is large enough one of these classes will contain, by Van der Waerden's theorem, 49 terms in arithmetic progression, say

$$a, a + b, a + 2b, \dots, a + 48b.$$

Now if  $\frac{a}{b} = c$  then we have 49 consecutive numbers of the same quadratic character, namely

$$c, c + 1, c + 2, \dots, c + 48.$$

and the result is complete.



The proof of the existence of nonresidues is considerably more complicated. Furthermore it is interesting to note that the existence of blocks like  $+-+-\dots$  is not covered by these methods.

We now return to the question raised by Gauss. What can be said about the least nonresidue  $n_p$  of a prime? Since 1 is a residue, the corresponding question about residues is “what is the smallest prime residue  $r_p$  of  $p$ ?”. These questions were attacked in the 1920s by a number of mathematicians including Nagel, Schur, Polya, Zeitz, Landau, Vandiver, Brauer, and Vinogradov. Nagel, for example, proved that for  $p \neq 7, 23$ ,  $n_p < \sqrt{p}$ . Polya and Schur proved that

$$\sum_{n=a}^b \left(\frac{n}{p}\right) < \sqrt{p} \log p.$$

This implies that there are never more than  $\sqrt{p} \log p$  consecutive residues or nonresidues and that ranges much larger than  $\sqrt{p} \log p$  have about as many residues as nonresidues. Using this result and some theorems on the distribution of primes, Vinogradov proved that for  $p > p_0$ ,

$$n_p < p^{\frac{1}{2\sqrt{e}}} \log^2 p < n^{.303}.$$

Checking through Vinogradov's proof we found that by his method the  $p_0$  is excessively large, say  $p_0 > 10^{10^{10}}$ . Nevertheless, in spite of numerous attempts this result of Vinogradov has not been much improved.

In 1938 Erdős and Ko showed that the existence of small nonresidues was intimately connected with the nonexistence of the Euclidean Algorithm in quadratic fields. This led Brauer, Hua, Min to re-examine the question of explicit bounds for the least nonresidue. Brauer already in 1928 had proved a number of such results, typical of which is that for all  $p \equiv 1 \pmod{8}$ ,

$$n_p < (2p)^{.4} + 3(2p)^{.2} + 1$$

and Hua and Min proved, for example, that for  $p > e^{250}$ ,

$$n_p < (60\sqrt{p})^{.625}.$$

Small primes (under 10,000,000) were considered by Bennet, Chatland, Brauer, Moser, and others.

Quite recently, the unproved extended Riemann hypothesis has been applied to these problems by Linnik, Chowla, Erdős, and Ankeny. Thus, for example, Ankeny used the extended Riemann hypothesis to prove that  $n_p \neq O(\log^2 p)$ . In the opposite direction Pillai (1945) proved that  $p \neq o(\log \log p)$ . Using first the Riemann hypothesis, and later some deep results of Linnik on primes in arithmetic progression, Friedlander and Chowla improved this to  $n_p \neq o(\log p)$ .

Quite recently there have been a number of results in a somewhat different direction by Brauer, Nagel, Skolem, Redei, and Kanold. Redei's method is particularly interesting. He uses a finite projective geometrical analogue of the fundamental theorem of Minkowski on convex bodies to prove that for  $p \equiv 1 \pmod{4}$ , at least  $\frac{1}{7}$  of the numbers up to  $\sqrt{p}$  are residues and at least 1 are non residues. Our own recent contributions to the theory are along these lines. We shall outline some of this work.

Consider first the lattice of points in a square of side  $m$ . We seek an estimate for  $V(m)$ , the number of visible lattice points in the square. As on the previous paragraph we find

$$[m]^2 = V(m) + V\left(\frac{m}{2}\right) + \dots$$

and inverting by the Mobius inversion formula yields

$$V(m) = \sum_{d \geq 1} \mu(d) \left[\frac{m}{d}\right]^2.$$

As before this leads to the asymptotic estimate

$$V(m) \sim \frac{6}{\pi^2} m^2.$$

We can however obtain explicit estimates for  $v(m)$  also. Indeed, from the exact formula for  $V(m)$  above one can show that for all  $m$ ,  $V(m) > .6m^2$ . We now take  $m = [\sqrt{p}]$ . For reasonably large  $p$  we shall have  $V([\sqrt{p}]) > .59m^2$ . Now with each visible lattice point  $(a, b)$  we associate the number  $\frac{a}{b} \pmod{p}$ . We now show that distinct visible points correspond to distinct numbers. Thus if  $\frac{a}{b} = \frac{c}{d}$  then  $ad \equiv bc$ . But  $ad < p$  and  $bc < p$ . Hence  $ad = bc$  and  $\frac{a}{b} = \frac{c}{d}$ . However  $(a, b) = (c, d) = 1$  so that  $a = c$  and  $b = d$ .

Since we have at least .59 distinct numbers represented by fractions  $\frac{a}{b}$ ,  $a < \sqrt{p}$ ,  $b < \sqrt{p}$ , at least .09 of these will correspond to nonresidues. If  $R$  denotes the number of residues  $< \sqrt{p}$  and  $N$  the number of nonresidues  $< \sqrt{p}$ , then  $R + N = \sqrt{p}$  and  $2RN > .09p$ . Solving these inequalities gives  $R, N > .04\sqrt{p}$ . This is weaker than Nagel's result but has the advantage of holding for primes  $p \equiv 3 \pmod{4}$  as well as  $p \equiv 1 \pmod{4}$ . Thus, exceptions turn out to be only the primes 7 and 23. For primes  $p \equiv 1 \pmod{4}$ ,  $-1$  is a nonresidue and this can be used together with above method to get stronger results. One can also use the existence of many nonresidues  $< \sqrt{p}$  to prove the existence of one small nonresidue, but the results obtainable in this way are not as strong as Vinogradov's result.

---

This page titled [1.5: Congruences](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Leo Moser](#) ([The Trilla Group](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

## 1.6: Diophantine Equations

Volume 2 of Dickson's *History of the Theory of Numbers* deals with Diophantine equations. It is as large as the other two volumes combined. It is therefore clear that we shall not cover much of this ground in this section. We shall confine our attention to some problems which are interesting though not of central importance.

One such problem is the Diophantine equation  $n! + 1 = x^2$  mentioned in an earlier section. The problem dates back to 1885 when H. Brocard conjectured that the only solutions are  $4! + 1 = 5^2$ ,  $5! + 1 = 11^2$  and  $7! + 1 = 71^2$ . About 1895 Ramanujan made the same conjecture but no progress towards a solution of the problem. About 1940 the problem appeared as an elementary (!) problem in the *Monthly*. No solutions were offered. However in 1950 an incorrect solution was published and since that time several faulty attempts to prove the result have been made. Again, about 1950 someone took the trouble to check, by brute force, the conjecture up to  $n = 50$ . However, earlier, in his book on the theory of numbers Kraitichik already had proved the result up to 5000. As far as we know that is where the problems stands. We shall now give an indication of Kraitichik's method.

Suppose we want to check  $100! + 1$ . Working (mod 103) we have

$$100!(-2)(-2) \equiv -1, 100! + \frac{1}{2} \equiv 0, 100! + 1 \equiv \frac{1}{2} \equiv 52.$$

If now 52 is a nonresidue of 103 we have achieved our goal. Otherwise we could carry out a similar calculation with another  $p > 100$ , say 107. Note that  $100! + 1 \equiv 0 \pmod{101}$  gives no information. Variations of this method can be used to eliminate many numbers wholesale and this is what Kraitichik did. We now outline a proof that  $n! + 1 = x^8$  has only a finite number of solutions. This proof depends on two facts which we have not proved:

(1) Every odd prime divisor of  $x^2 + 1$  is of the form  $4n + 1$ ;

(2) There are roughly as many primes  $4n + 1$  as  $4n + 3$ .

Now  $n! + 1 = x^8$  gives  $n! = x^8 - 1 = (x^4 + 1)(x^2 + 1)(x^2 - 1)$ ; on the right the contribution of primes  $4k + 1$  and  $4k - 1$  is about the same while on the left all the odd prime factors of  $(x^4 + 1)(x^2 + 1)$  i.e., about  $(n!)^{3/4}$  of the product, are of the form  $4n + 1$ .

We now go on to quite a different problem. Has the equation

$$1^n + 2^n + \cdots + (m-1)^n = m^n$$

any solutions in integers other than  $1 + 2 = 3$ ? Here are some near solutions:

$$3^2 + 4^2 = 5^2,$$

$$3^3 + 4^3 + 5^3 = 6^3,$$

$$1^6 + 2^6 + 4^6 + 7^6 + 9^6 + 12^6 + 13^6 + 15^6 + 16^6 + 18^6 + 20^6 + 22^6 + 23^6 = 28^6.$$

We now outline a proof that if other solutions exist then  $m > 10^{1000000}$ . The rest of this section appeared originally as the paper "On the Diophantine Equation  $1^n + 2^n + \cdots + (m-1)^n = m^n$ ," *Scripta Mathematica*, 19 (1953), pp. 84-88. (Pieter Moree discusses this theorem and proof in "A top hat for Moser's four mathematical rabbits," *The American Mathematical Monthly*, 118 (2011), 364-370.)

A number of isolated equations expressing the sum of the  $n^{\text{th}}$  powers of integers as an  $n^{\text{th}}$  power of an integer have long been known. Some examples are:

$$3^3 + 4^3 + 5^3 = 6^3$$

$$\sum_{i=1}^{100} i^4 - 1^4 - 2^4 - 3^4 - 8^4 - 10^4 - 72^4 = 212^4$$

$$1^6 + 2^6 + 4^6 + 5^6 + 6^6 + 7^6 + 9^6 + 12^6 + 13^6 + 15^6 + 16^6 + 18^6 + 20^6 + 21^6 + 22^6 + 23^6 = 28^6$$

Further examples and references to such results are given in [1, p. 682]. On the other hand the only known solution in integers to the equation in the title is the trivial one  $1 + 2 = 3$ . In a letter to the author, P. Erdős conjectured that this is the only solution. The object of this note is to show that if the equation has a solution with  $n > 1$ , then  $m > 10^{1000000}$ .

Let  $S_n(m)$  denote  $\sum_{i=1}^{m-1} i^n$ . In what follows we assume

$$S_n(m) \equiv m^n, n > 1. \tag{1.6.1}$$

It is possible to examine (6.1) with various moduli and thereby obtain restrictions on  $m$  and  $n$ . This is essentially our method, but the moduli are so chosen that we are able to combine the resulting congruences so as to obtain extremely large bounds for  $m$  without laborious computation.

We use the following lemma.

**Lemma 1.** If  $p$  is a prime and  $\epsilon_n(p)$  is defined by  $\epsilon_n(p) = -1$  when  $(p-1) \mid n$  and  $\epsilon_n(p) = 0$  when  $(p-1)$  does not divide  $n$  then

$$S_n(p) \equiv \epsilon_n(p) \pmod{p}. \quad (1.6.2)$$

A simple proof of (2) is given in [2, p. 90].

Now suppose  $p \mid (m-1)$ , then

$$s_n(m) = \sum_{i=0}^{\frac{m-1}{p}-1} \sum_{j=1}^p (j+ip)^n \equiv \frac{m-1}{p} \cdot \epsilon_n(p) \pmod{p}.$$

On the other hand  $m \equiv 1 \pmod{p}$  so that by (6.1)

$$\frac{m-1}{p} \cdot \epsilon_n(p) \equiv 1 \pmod{p}. \quad (1.6.3)$$

Hence  $\epsilon(p) \not\equiv 0 \pmod{p}$  so that from the definition of  $\epsilon_n(p)$  it follows that  $\epsilon_n(p) = -1$  and

$$p \mid (m-1) \text{ implies } (p-1) \mid n. \quad (1.6.4)$$

Thus (6.3) can be put in the form

$$\frac{m-1}{p} + 1 \equiv 0 \pmod{p} \quad (1.6.5)$$

or

$$m-1+p \equiv 0 \pmod{p^2}. \quad (1.6.6)$$

From (6.6) it follows that  $m-1$  is squarefree. Further, since it is easily checked that  $m-1 \neq 2$  it follows that  $m-1$  must have at least one prime divisor, so by (6.4)  $n$  is even.

We now multiply together all congruences of the type (6.5), that is one for each prime dividing  $m-1$ . Since  $m-1$  is squarefree, the resulting modulus is  $m-1$ . Furthermore, products containing two or more distinct factors of the form  $(m-1)/p$  will be divisible by  $m-1$ . Thus we obtain

$$(m-1) \sum_{p \mid (m-1)} \frac{1}{p} + 1 \equiv 0 \pmod{m-1} \quad (1.6.7)$$

or

$$\sum_{p \mid (m-1)} \frac{1}{p} + \frac{1}{m-1} \equiv 0 \pmod{1}. \quad (1.6.8)$$

The only values of  $m \leq 1000$  which satisfy (6.8) are 3, 7, 43.

We proceed to develop three more congruences, similar to (6.8), which when combined with (6.8) lead to the main result. Equation (6.1) can be written in the form

$$S_n(m+2) = 2m^n + (m+1)^n. \quad (1.6.9)$$

Suppose that  $p \mid (m+1)$ . Using (6.2) and the fact that  $n$  is even, we obtain as before

$$p \mid (m+1) \text{ implies } (p-1) \mid n \quad (1.6.10)$$

and

$$\frac{m+1}{p} + 2 \equiv 0 \pmod{p}. \quad (1.6.11)$$

From (6.11) it follows that no odd prime appears with the exponent greater than one in  $m+1$ . The prime 2 however, requires special attention. If we examine (1) with modulus 4, and we use the fact that  $n$  is even, then we find that  $m+1 \equiv 1$  or  $4 \pmod{8}$ . Thus  $m+1$  is odd or contains 2 exactly to the second power. If we assume the second of these possibilities then (6.11) can be put in the form

$$\frac{m+1}{2p} + 1 \equiv 0 \pmod{p}. \quad (1.6.12)$$

We multiply together all the congruences of the type (12). This modulus then becomes  $\frac{m+1}{2}$ . Further any term involving two or more distinct factors  $\frac{m+1}{2p}$  will be divisible by  $\frac{m+1}{2}$  so that on simplification we obtain

$$\sum_{p \mid (m+1)} \frac{1}{p} + \frac{2}{m+1} \equiv 0 \pmod{1}. \quad (1.6.13)$$

We proceed to find two or more congruences similar to (6.13) without using the assumption that  $m+1$  is even. Suppose that  $p \mid 2m-1$  and let  $t = \frac{1}{2}(\frac{2m-1}{p} - 1)$ . Clearly  $t$  is an integer and  $m-1 = tp + \frac{p-1}{2}$ . Since  $n$  is even  $a^n = (-a)^n$  so that

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

Now

$$S_n(m) = \sum_{i=0}^{t-1} \sum_{j=1}^{p-1} (j+ip)^n + \sum_{i=1}^{(p-1)/2} i^n \equiv \left(t + \frac{1}{2}\right) \epsilon_n(p) \pmod{p}. \quad (1.6.14)$$

On the other hand  $m^n \equiv 0 \pmod{p}$  so that (6.1) and (6.14) imply  $\epsilon_n(p) \not\equiv 0$ . Hence  $p-1/n$  and by Fermat's theorem  $m^n \equiv 1 \pmod{p}$ . Thus (6.1) and (6.14) yield  $-(t + \frac{1}{2}) \equiv 1 \pmod{p}$ . Replacing  $t$  by its value and simplifying we obtain

$$\frac{2m-1}{p} + 2 \equiv 0 \pmod{p}. \quad (1.6.15)$$

Since  $2m-1$  is odd (6.15) implies that  $2m-1$  is squarefree. Multiplying congruences of type (6.15), one for each of the  $r$  prime divisors of  $2m-1$  yields

$$2^{r-1}((2m-1) \sum_{p \mid (2m-1)} \frac{1}{p} + 2) \equiv 0 \pmod{2m-1}.$$

Since the modulus is odd this gives

$$\sum_{p \mid (2m-1)} \frac{1}{p} + \frac{2}{2m-1} \equiv 0 \pmod{1}. \quad (1.6.16)$$

Finally we obtain a corresponding congruence for primes dividing  $2m+1$ . For this purpose we write (6.1) in the form

$$S_n(m+1) = 2m^n. \quad (1.6.17)$$

Suppose  $p \mid 2m+1$ . Set  $v = \frac{1}{2}(\frac{2m+1}{p} - 1)$ . Using again an argument similar to that employed to obtain (6.16) we find that  $(p-1) \mid n$  and  $2m+1$  is squarefree. Finally we obtain

$$\sum_{p \mid (2m+1)} \frac{1}{p} + \frac{4}{2m+1} \equiv 0 \pmod{p}. \quad (1.6.18)$$

We assume again that  $m+1$  is even so that (6.13) holds. If we now add the left hand sides of (6.8), (6.13), (6.16), and (6.18) we get an integer, at least 4. No prime  $p > 3$  can divide more than one of the numbers  $m-1$ ,  $m+1$ ,  $2m-1$ ,  $2m+1$ . Further, 2 and 3 can divide at most two of these numbers. Hence if  $M = (m-1)(m+1)(2m-1)(2m+1)$  then

$$\sum_{p|M} \frac{1}{p} + \frac{1}{m-1} + \frac{2}{m+1} + \frac{2}{2m-1} + \frac{4}{2m+1} \geq 4 - \frac{1}{2} - \frac{1}{3}. \quad (1.6.19)$$

We have already seen that the only possibilities for  $m$  with  $m \leq 1000$  are 3, 7, and 43. These are easily ruled out by (6.16). Thus (6.19) yields

$$\sum_{p|M} \frac{1}{p} > 3.16. \quad (1.6.20)$$

From (6.20) it follows that if  $\sum_{p \leq x} \frac{1}{p} < 3.16$  then  $M > \prod_{p \leq x} p$ . We shall prove the following lemma.

#### lemma 2

$$\sum_{p \leq 10^7} \frac{1}{p} < 3.16.$$

#### Proof

By direct computation

$$\sum_{p \leq 10^8} \frac{1}{p} < 2.18. \quad (1.6.21)$$

From Lehmer's table [3] and explicit estimates for  $\pi(x)$  due to Rosser [4] it can easily be checked that for  $10^3 < x < 10^7$

$$\pi(x) < \frac{1.2x}{\log x}. \quad (1.6.22)$$

Now in [2, p. 339] it is shown that

$$\sum_{p \leq x} \frac{1}{p} = \frac{\pi(x)}{x} + \int_2^x \frac{\pi(t)}{t^2} dt. \quad (1.6.23)$$

combining (6.21), (6.22) and (6.23) gives the required result.

In [4] it is proved that

$$\sum_{p \leq x} \log p > (1 - \frac{1}{\log x})x, x < e^{100}. \quad (1.6.24)$$

Hence

$$\log M > \log \prod_{p \leq 10^7} p = \sum_{p \leq 10^7} \log p > (1 - \frac{1}{7 \log 10})10^7 > (.93)10^7.$$

Now  $M < 4n^2$  so that

$$\log m > (\frac{\log M - \log 4}{2}) > (.231)10^7$$

and  $m > e^{(.231)10^7} > 10^{1000000}.$

Returning to the case  $m-1$  odd, we note that in this we cannot use (6.13). Letting  $N = (m-1)(2m-1)(2m+1)$  we get from (6.8), (6.16) and (6.18)

$$\sum_{p|N} \frac{1}{p} + \frac{1}{m-1} + \frac{2}{2m-1} + \frac{4}{2m+1} > 3 - \frac{1}{3}. \quad (1.6.25)$$

However, since the prime 2 does not appear on the left side (6.25) is actually a stronger condition on  $m$  than is (6.19) so that in any case  $m > 10^{1000000}.$

#### References

- [1] L.E. Dickson, *History of the Theory of Numbers*, vol. 2
  - [2] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*.
  - [3] D.H. Lehmer, “List of Prime Numbers from 1 to 10,006,721.”
  - [4] B. Rosser, “Explicit Bounds for Some Functions of Prime Numbers”, *Amer. Jour. of Math.*, 63 (1941), 211–232.
- 

This page titled [1.6: Diophantine Equations](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Leo Moser \(The Trilla Group\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

## 1.7: Combinatorial Number Theory

There are many interesting questions that lie between number theory and combinatorial analysis. We consider first one that goes back to I. Schur (1917) and is related in a surprising way to Fermat's Last Theorem. Roughly speaking, the theorem of Schur states that if  $n$  is fixed and sufficiently many consecutive integers  $1, 2, 3, \dots$  are separated into  $n$  classes, then at least one class will contain elements  $a, b, c$  with  $a + b = c$ .

Consider the fact that if we separate the positive integers less than  $2^n$  into  $n$  classes by putting 1 in class 1, the next 2 in class 2, the next 4 in class 3, etc., then no class contains the sum of two of its elements. Alternatively, we could write every integer  $m$  in form  $2^k \theta$  where  $\theta$  is odd, and place  $m$  in the  $k$ th class. Again the numbers less than  $2^n$  will lie in  $n$  classes and if  $m_1 = 2^{k_1} \theta_1$  and  $m_2 = 2^{k_2} \theta_2$  are in class  $k$  then  $m_1 + m_2 = 2^k (\theta_1 + \theta_2)$  lies in a higher numbered class. The more complicated manner of distributing integers outlined below enables us to distribute  $1, 2, \dots, \frac{3^n - 1}{2}$  into  $n$  classes in such away that no class has a solution to  $a + b = c$ :

1 2 5  
3 4 6  
10 11 7  
13 12 8  
...  
...  
...

On the other hand, the theorem of Schur states that if one separates the numbers  $1, 2, 3, \dots, [n!e]$  into  $n$  classes in any manner whatsoever then at least one class will contain a solution to  $a + b = c$ . The gap between the last two statements reveals an interesting unsolved problem, namely, can one replace the  $[n!e]$  in Schur's result by a considerably smaller number? The first two examples given show that we certainly cannot go as low as  $2n - 1$ , and the last example shows that we cannot go as low as  $\frac{3^n - 1}{2}$ .

We now give a definition and make several remarks to facilitate the proof of Schur's theorem.

Let  $T_0 = 1, T_n = nT_{n-1} + 1$ . It is easily checked that

$$T_n = n! \left( 1 + \frac{1}{1!} + \frac{2}{2!} \cdots + \frac{1}{n!} \right) = [n!e].$$

Thus Schur's theorem can be restated as follows: If  $1, 2, \dots, T_n$  are separated into  $n$  classes in any manner whatever, at least one class will contain a solution of  $a + b = c$ . We will prove this by assuming that the numbers  $1, 2, \dots, T_n$  have been classified  $n$  ways with no class containing a solution of  $a + b = c$  and from this obtain a contradiction. Note that the condition  $a + b \neq c$  means that no class can contain the difference of two of its elements.

Suppose that some class, say  $A$ , contains elements  $a_1 < a_2 < \dots$ . We form differences of these in the following manner:

$$\begin{aligned} b_1 &= a_2 - a_1, b_2 = a_3 - a_1, b_3 = a_4 - a_1, \dots \\ c_1 &= b_2 - b_1, c_2 = b_3 - b_1, c_3 = b_4 - b_1, \dots \\ d_1 &= c_2 - c_1, d_2 = c_3 - c_1, d_3 = c_4 - c_1, \dots \end{aligned}$$

and so on. We note that all the  $b$ 's,  $c$ 's,  $d$ 's, etc., are differences of  $a$ 's and hence cannot lie in  $A$ .

Now, we start with  $T_n$  elements. At least

$$\left\lfloor \frac{T_n}{n} + 1 \right\rfloor = T_{n-1} + 1$$

of these must lie in a single class, say  $A_1$ . We then form  $T_{n-1}$   $b$ 's. These do not lie in  $A_1$ , and hence lie in the remaining  $n - 1$  classes. At least

$$\left\lfloor \frac{T_{n-1}}{n-1} + 1 \right\rfloor = T_{n-2} + 1$$

of them must lie in a single class, say  $A_2$ . Form their  $T_{n-2}$  differences, the  $c$ 's. These yield  $T_{n-2}$  numbers neither in  $A_1$  nor  $A_2$ . Continuing in this manner yields  $T_{n-3}$  numbers not in  $A_1, A_2, A_3$ . In this manner we eventually obtain  $T_0 = 1$  number not



belonging to  $A_1, A_2, \dots, A_n$ . But all numbers formed are among the numbers  $1, 2, \dots, T_n$  so we have a contradiction, which proves the theorem.

We state, without proof, the connection with Fermat's last theorem. A natural approach to Fermat's theorem would be to try to show that  $x^n + y^n = z^n \pmod{p}$  is insolvable modulo some  $p$ , provided  $p$  does not divide  $x \cdot y \cdot z$ . However, Schur's theorem can be used to show that this method must fail and indeed if  $p > n!e$  then  $x^n + y^n = z^n \pmod{p}$  has a solution with  $p$  not a factor of  $xyz$ .

Somewhat related to Schur's theorem is a famous theorem of Van der Waerden, which we briefly investigate. In the early 1920's the following problem arose in connection with the theory of distribution of quadratic residues. Imagine the set of all integers to be divided in any manner into two classes. Can one assert that arithmetic progressions of arbitrary length can be found in at least one of these classes? The problem remained unsolved for several years in spite of concentrated efforts by many outstanding mathematicians. It was finally solved in 1928 by Van der Waerden. As is not uncommon with such problems, Van der Waerden's first step was to make the problem more general, and hence easier.

Van der Waerden proved the following: Given integers  $k$  and  $\ell$ , there exists an integer  $W = W(k, \ell)$  such that if the numbers  $1, 2, 3, \dots, W$  are separated into  $k$  classes in any manner, then at least one class will contain  $\ell$  terms in arithmetic progression. We will not give Van der Waerden's proof here. It is extremely tricky, difficult to see through, and leads only to fantastically large bound for  $W(k, \ell)$ . For this reason the reader might consider the very worthwhile unsolved problem of finding an alternative simpler proof that  $W(k, \ell)$  exists and finding reasonable bounds for it. We will have a little more to say about the function  $W(k, \ell)$  a little later.

Our next problem of combinatorial number theory deals with "nonaveraging" sequences. We call a sequence  $A: a_1 < a_2 < a_3 < \dots$  non-averaging if it does not contain the average of two of its elements, i.e.,  $a_i + a_j \neq 2a_k$  ( $i \neq j$ ). Let  $A(n)$  denote the number of elements in  $A$  not exceeding  $n$ . The main problem is to estimate how large  $A(n)$  can be if  $A$  is nonaveraging. We can form a nonaveraging sequence by starting with 1, 2, ... and then always taking the smallest number that does not violate the condition for nonaveraging sets. In this way we obtain 1, 2, 4, 5, 10, 11, 13, 14, 28, 29, 31, ... . It is an interesting fact that this sequence is related to the famous Cantor ternary set. Indeed, we leave it as an exercise to prove that this sequence can be obtained by adding 1 to each integer whose representation in base 3 contains only 0's and 1's. This sequence is maximal in the sense that no new number can be inserted into the sequence without destroying its nonaveraging character. This, as well as other facts, led Szekeres (about 1930) to conjecture that this set was as dense as any nonaveraging set. For this set, the counting function can easily be estimated to be  $\sim n^{\log 2 / \log 3}$ . It therefore came as a considerable surprise when Salem and Spencer (1942) proved that one could have a nonaveraging set of integers  $\leq n$  containing at least  $n^{1-c/\sqrt{\log \log n}}$  elements.

Given a number  $x$ , written in base ten, we decide whether  $x$  is in  $R$  on the basis of the following rules.

First we enclose  $x$  in a set of brackets, putting the first digit (counting from right to left) in the first bracket, the next two in the second bracket, the next three in the third bracket, and so on. If the last nonempty bracket (the bracket furthest to the left that does not consist entirely of zeros) does not have a maximal number of digits, we fill it with zeros. For instance, the numbers

$$a = 32653200200 \quad b = 100026000150600 \quad c = 1000866600290500$$

would be bracketed

$$\begin{aligned} a &= (00003)(2653)(200)(20)(0), \\ b &= (10002)(6100)(150)(60)(0), \\ c &= (10008)(6600)(290)(50)(0), \end{aligned}$$

respectively. Now suppose the  $r^{\text{th}}$  bracket in  $x$  contains nonzero digits, but all further brackets to the left are 0. Call the number represented by the digits in the  $i^{\text{th}}$  bracket  $x_i$ ,  $i = 1, 2, \dots, r-2$ . Further, denote by  $\bar{x}$  the number represented by the digit in the last two brackets taken together, but excluding the last digit. For  $x$  to belong to  $R$  we require

1. the last digit of  $x$  must be 1,
2.  $x_i$  must begin with 0 for  $i = 1, 2, \dots, r-2$ ,
3.  $x_1^2 + \dots + x_{r-2}^2 = \bar{x}$ .

In particular, note that  $a$  satisfies (2) but violates (1) and (3) so that  $a$  is not in  $R$ ; but  $b$  and  $c$  satisfy all three conditions and are in  $R$ . To check (3) we note that  $60^2 + 150^2 = 26100$ .

We next prove that no three integers in  $R$  are in arithmetic progression. First note that if two elements of  $R$  have a different number of nonempty brackets their average cannot satisfy (1). Thus we need only consider averages of elements of  $R$  having the

same number of nonempty brackets. From (1) and (3) it follows that the two elements of  $R$  can be averaged bracket by bracket for the first  $r - 2$  brackets and also for the last two brackets taken together. Thus, in our example,

$$\begin{aligned}\frac{1}{2}(60 + 50) &= 55, \quad \frac{1}{2}(150 + 290) = 220, \\ \frac{1}{2}(100026100 + 100086600) &= 100056350, \\ \frac{1}{2}(b + c) &= (10005)(6350)(220)(55)(0)\end{aligned}$$

This violates (3) and so is not in  $R$ . In general we will prove that if  $x$  and  $y$  are in  $R$  then  $\bar{z} = \frac{1}{2}(x + y)$  violates (3) and so is not in  $R$ .

Since  $x$  and  $y$  are in  $R$ ,

$$\bar{z} = \frac{\bar{x} + \bar{y}}{2} = \sum_{i=1}^{r-2} \frac{x_i^2 + y_i^2}{2}.$$

On the other hand,  $z$  in  $R$  implies

$$\bar{z} = \sum_{i=1}^{r-2} z_i^2 = \sum_{i=1}^{r-2} \frac{(x_i + y_i)^2}{2}.$$

Hence, if  $z$  is in  $R$  then

$$\sum_{i=1}^{r-2} \frac{x_i^2 + y_i^2}{2} = \sum_{i=1}^{r-2} \frac{(x_i + y_i)^2}{2}.$$

Thus

$$\sum_{i=1}^{r-2} \frac{(x_i - y_i)^2}{2} = 0,$$

which implies  $x_i = y_i$  for  $i = 1, 2, \dots, r - 2$ . This together with (1) and (2) implies that  $x$  and  $y$  are not distinct.

Szekeres' sequence starts with 1, 2, 4, 5, 10, 11, ... . Our sequence starts with

$$100000, 1000100100, 1000400200, \dots$$

Nevertheless, the terms of this sequence are eventually much smaller than the corresponding terms of Szekeres' sequence. We now estimate how many integers in  $R$  contain exactly  $r$  brackets. Given  $r$  brackets we can make the first digit in each of the  $r - 2$  brackets 0. We can fill up the first  $r - 2$  brackets in as arbitrary manner. This can be done in

$$10^{0+1+2+\dots+(r-2)} = 10^{\frac{1}{2}(r-1)(r-2)}$$

ways. The last two brackets can be filled in such a way as to satisfy (1) and (3).

To see this we need only check that the last two brackets will not be overfilled, and that the last digit, which we shall set equal to 1, will not be interfered with. This follows from the inequality

$$(10^1)^2 + (10^2)^2 + \dots + (10^{r-2})^2 < 10^{2(r-1)}.$$

For a given  $n$  let  $r$  be the integer determined by

$$10^{\frac{1}{2}r(r+1)} \leq n < 10^{\frac{1}{2}(r+1)(r+2)}. \quad (1.7.1)$$

Since all the integers with at most  $r$  brackets will not exceed  $n$ , and since  $r$  brackets can be filled to specification in  $10^{\frac{1}{2}(r-2)(r-1)}$  ways, we have

$$R(n) \geq 10^{\frac{1}{2}(r-2)(r-1)} \quad (1.7.2)$$

From the right hand side of (7.1) we have

$$r + 2 > \sqrt{2 \log n}$$

so that (7.2) implies that

$$R(n) \geq 10^{\frac{1}{2}(r-2)(r-1)} > 10^{\log n - c\sqrt{\log n}} > 10^{(\log n)(1-c/\sqrt{\log n})}$$

where all logs are to base 10.

An old conjecture was that  $\frac{A(n)}{n} \rightarrow 0$  for every nonaveraging sequence. This has only been proved quite recently (1954) by K. F. Roth. His proof is not elementary.

L. Moser has used a similar technique to get lower bounds for the Van der Waerden function  $W(k, \ell)$ . He proved that  $W(k, \ell) > \ell k^{\log k}$ , i.e., he showed how to distribute the numbers, 1, 2, ...,  $[\ell k^{\log k}]$  into  $k$  classes in such a way that no class contains 3 terms in arithmetic progression. Using a quite different method Erdős and Rado have shown that  $W(k, \ell) > \sqrt{2\ell k^\ell}$ .

Erdős has raised the following question: What is the maximum number of integers  $a_1 < a_2 < \dots < a_k \leq n$  such that  $2^k$  sums of distinct  $a$ 's are all distinct? The powers of 2 show that one can give  $k+1$   $a$ 's not exceeding  $2^k$  and one can in fact give  $k+2$   $a$ 's under  $2^k$  satisfying the required condition. On the other hand, all the sums involved are less than  $kn$  so that

$$2^k \leq kn, \quad (1.7.3)$$

which implies

$$k < \frac{\log n}{\log 2} + (1 + o(1)) \frac{\log \log n}{\log 2}. \quad (1.7.4)$$

We now show how Erdős and Moser improved these estimates (Publisher's note: The current best lower bound may be found in I. Aliev, "Siegel's lemma and sum-distinct sets," Discrete Comput. Geom. 39 (2008), 59–66.) to

$$2^k < 4\sqrt{k}n, \quad (1.7.5)$$

which implies

$$k < \frac{\log n}{\log 2} + (1 + o(1)) \frac{\log \log n}{2 \log 2}. \quad (1.7.6)$$

The conjecture of Erdős is that

$$k = \frac{\log n}{\log 2} + o(1). \quad (1.7.7)$$

Denote the sum of distinct  $a$ 's by  $s_1, s_2, \dots, s_{2^k}$  and let  $A = a_1 + a_2 + \dots + a_k$ . Observe that the average sum is  $\frac{A}{2}$  since we can pair each sum with the sum of the complementary set. This suggests that we estimate  $\sum_i (s_i - \frac{A}{2})^2$ .

We have

$$\sum_i (s_i - \frac{A}{2})^2 = \sum \frac{1}{2} (\pm a_1 \pm a_2 \pm \dots \pm a_k)^2,$$

where the last sum runs over the  $2^k$  possible distributions of sign. Upon squaring we find that all the cross terms come in pairs while each  $a_i^2$  will appear  $2^k$  times. Thus

$$\sum_i (s_i - \frac{A}{2})^2 = 2^k \sum a_i^2 < 2^{k-2} n^2 k.$$

Thus the number of sums  $s_i$  for which

$$|s_i - \frac{A}{2}| \geq n\sqrt{k}$$

cannot exceed  $2^{k-1}$ . Since all the sums are different, we have  $2^{k-1}$  distinct numbers in a range of length  $2n\sqrt{k}$ . This yields  $2^{k-1} \leq 2n\sqrt{k}$  as required.

Let  $a_1 < a_2 < \dots$  be an infinite sequence of integers and define  $f(n)$  to be the number of solutions of  $n = a_i + a_j$  where all solutions count once. G. A. Dirac and D. J. Newman gave the following interesting proof that  $f(n)$  cannot be constant from some

stage on. If  $f(\ell+1) = f(\ell+2) = \dots$  we would have

$$\begin{aligned} \frac{1}{2}(\sum a_k)^2 + \sum z^{2a_k} &= \sum f(n)z^n \\ &= P_\ell(z) + a \frac{z^{\ell+1}}{1-z}, \quad (f(\ell+1) = a), \end{aligned}$$

where  $P(z)$  is a polynomial of degree  $\leq \ell$ . If  $z \rightarrow -1$  on the real axis the right side remains bounded, but the left side approaches infinity, since both terms on the left side are positive, and the second tends to infinity. This contradiction proves the theorem.

Turan and Erdős conjectured that if  $f(n) > 0$  for all sufficiently large  $n$  then  $\limsup f(n) = \infty$  but this seems very difficult to prove. A still stronger conjecture would be that if  $a_k > ck^2$  then  $\limsup f(n) = \infty$ . The best known result in this direction is only  $\limsup f(n) \geq 2$ .

Fuchs and Erdős recently proved that

$$\sum_{k=1}^n f(k) = cn + o\left(\frac{n^{\frac{1}{4}}}{\log n}\right)$$

is impossible. If  $a_k = k^2$  one comes to the problem of lattice points in a circle of radius  $n$ . Here Hardy and Landau proved

$$\sum_{k=1}^n f(k) = \pi n + o(n \log n)$$

does not hold. Though not quite as strong as this, the result of Erdős and Fuchs is applicable to a much more general situation and is much easier (but not very easy) to prove.

Let  $a_1 < a_2 < \dots$  be an infinite sequence of integers. Erdős conjectured, and G. G. Lorentz proved, that there exists a sequence  $b_i$  of zero density such that every integer is of the form  $a_i + b_j$ .

An interesting unsolved problem along these lines is to find a sequence  $B: b_1 < b_2 < \dots$  with counting function  $B(n) < \frac{cn}{\log n}$  such that every integer is of the form  $2^k + b_j$ .

Let  $a_1 < a_2 < \dots < a_{2n}$  be  $2n$  integers in the interval  $[1, 4n]$  and  $b_1 < b_2 < \dots < b_{2n}$  the remaining numbers in the interval. Erdős conjectured that there exists an integer  $x$  such that the number of solutions of  $a_i + x = b_j$  is at least  $n$ . It is quite easy to show that there exists an  $x$  so that the number of solutions of  $a_i + x = b_j$  is at least  $\frac{n}{2}$ . We merely observe that the number of solutions of  $a_i + y = b_j$  is  $4n^2$  and that there are  $8n$  possible choices of  $y$ , i.e.,  $-4n \leq y \leq 4n, y \neq 0$ . Thus for some  $y_0$  there are at least  $\frac{n}{2}$   $b$ 's in  $a_i + y_0$  as stated.

P. Scherk improved the  $\frac{n}{2}$  to  $n(2 - \sqrt{2}) = .586n$ . By an entirely different method L. Moser improved this further to  $.712n$ . On the other hand Selfridge, Ralston and Motzkin have used S.W.A.C. to disprove the original conjecture and have found examples where no number is representable more than  $.8n$  times as a difference between an  $a$  and a  $b$ .

Still another set of interesting problems of combinatorial number theory revolve about the concept of addition chain introduced by A. Scholz. An addition chain for  $n$  is a set of integers  $1 = a_0 < a_1 < \dots < a_r = n$  such that every element  $a_p$  can be written as a sum  $a_\sigma + a_\tau$  of preceding elements of the chain. For example for  $n = 666$

$$1, 2, 4, 8, 16, 24, 40, 80, 160, 320, 640, 666$$

form a chain with  $r = 12$ ; the same holds for

$$1, 2, 3, 6, 9, 18, 27, 54, 81, 162, 324, 648, 666.$$

In any case we must have  $a_1 = 2$ , and  $a_2 = 3$  or  $4$ . By the length  $\ell = \ell(n)$  Scholtz understands the smallest  $\ell$  for which there exists an addition chain  $a_0, a_1, \dots, a_\ell = n$ .

Scholtz stated the following:

$$\begin{aligned} m+1 &\leq \ell(n) \leq 2m \text{ for } 2^m + 1 \leq n \leq 2^{m+1} \quad (m \geq 1); \\ \ell(ab) &\leq \ell(a) + \ell(b); \\ \ell(2^{m+1} - 1) &\leq m + \ell(m+1). \end{aligned}$$

The first two of these are easy to prove. The third we would conjecture to be false. Scholtz surmised that the first could be improved and raised the question of whether or not

$$1 \leq \limsup_{n \rightarrow \infty} \frac{\ell(n)}{\log_2 n} \leq 2$$

could be improved.

In what follows we prove (1) and outline a proof due to A. Brauer that

$$\ell(n) \sim \log_2 n.$$

Suppose integers are written in base 2 and we seek an addition chain for 10110110 say. We might form the chain

$$1, 10, 100, 101, 1010, 1011, 10110, 101100, 101101, 1011010, \\ 1011011, 10110110, 101101100, 101101101.$$

In this process, each digit “costs” at most two elements in the chain so that  $\ell < 2 \log_2 n$ . Since the left hand side of the inequality of (1) is trivial the method suggested above yields a proof of (1).

Brauer’s idea is to build up a large stock of numbers first and use it when the occasion arises. Suppose  $n$  is about  $2^m$ . We start out with the chain  $1, 2, \dots, 2^r$ , where  $r$  will be determined later. We can now break up the digits of  $n$  into  $m/r$  blocks with  $r$  digits in each block. For example, suppose

$$n = (101)(110)(010)(101)(111)$$

Here  $m = 15, r = 3$ .

Starting with our stock of all 3 digit numbers we can proceed as follows:

$$1, 10, 100, \underline{101}, 1010, 10100, 101000, \underline{101110}, \\ 1011100, 10111000, 101110000, \underline{101110010}, \dots$$

where between the underlined stages we double and at the underlined stages we add the appropriate number from our stock to build up  $n$ . In this case we would need  $2^3 + 2^{15} + 5$  steps. In general, the number of steps for a number under  $2^m$  would be about  $2^r + m + \frac{m}{c}$ . By appropriate choice of  $r$  we could make  $2^r + \frac{m}{c}$  as small as we please in comparison with  $m$ . Indeed, using this idea Brauer proved in general

$$\ell(n) < \log_2 n + \frac{1}{\log \log n} + \frac{2 \log 2}{(\log n)^{1-\log 2}}.$$

---

This page titled [1.7: Combinatorial Number Theory](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Leo Moser \(The Trilla Group\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.

## 1.8: Geometry of Numbers

We have already seen that geometrical concepts are sometimes useful in illuminating number theoretic considerations. With the introduction by Minkowski of geometry of numbers a real welding of important parts of number theory and geometry was achieved. This branch of mathematics has been in considerable vogue in the last 20 years, particularly in England where it was and is being developed vigorously by Mordell, Davenport, Mahler and their students.

We shall consider a very brief introduction to this subject. First, we shall examine a proof of the fundamental theorem of Minkowski due to Hajos (1934), then we shall discuss some generalizations and applications of this theorem, and finally we shall investigate some new results and conjectures that are closely related.

In its simplest form the fundamental theorem of Minkowski is the following.

### Theorem 1.8.1: Fundamental Theorem of Minkowski

Let  $R$  be a region in the  $x - y$  plane of area  $A > 4$ , symmetric about the origin and convex. Then  $R$  contains a lattice point other than the origin.

First, some preliminary remarks. In the condition  $A > 4$ , the 4 cannot be replaced by any smaller number. This may be seen by considering the square of side  $2 - \epsilon$ , centered at the origin. Indeed this example might at first suggest that the theorem is quite intuitive, as it might seem that squeezing this region in any direction and keeping its area fixed would necessarily force the region to cover some lattice point. However the matter is not quite so simple, as other examples reveal that neither central symmetry nor convexity are indispensable. As far as convexity is concerned what is really needed is that with the vectors  $\vec{V}_1$  and  $\vec{V}_2$  the region should also contain  $\frac{1}{2}(\vec{V}_1 + \vec{V}_2)$ . The symmetry means that with  $\vec{V}_1$  the vector  $-\vec{V}_1$  should also be in  $R$ . Thus the symmetry and convexity together imply that, if  $\vec{V}_1$  and  $\vec{V}_2$  are in  $R$ , so is  $\frac{1}{2}(\vec{V}_1 - \vec{V}_2)$ . This last condition is really sufficient for our purpose and may replace the conditions of symmetry and convexity. It is implied by symmetry and convexity but does not imply either of these conditions.

Another example that perhaps illuminates the significance of Minkowski's theorem is the following. Consider a line through  $O$  having irrational slope  $\tan \theta$ ; see Figure 4. This line passes through no lattice point other than the origin. If we take a long segment of this line, say extending length  $R$  on either side of  $O$ , then there will be a lattice point closest to, and a distance  $r$  from,

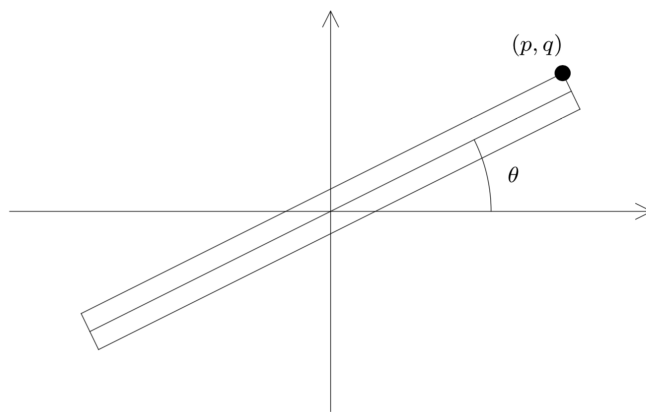


FIGURE 4. The long side of the rectangle is  $2R$ , the short  $2r$ .

this segment. Hence, no matter how large  $R$  is, we can construct a rectangle containing this line segment, which contains no lattice point other than  $O$ . By the fundamental theorem of Minkowski the area  $4rR$  of this rectangle does not exceed 4. Thus  $r \leq \frac{1}{R}$ . Note that if  $(p, q)$  is a lattice point on the border of the rectangle then  $\frac{p}{q} \approx \tan \theta$ , so that the fundamental theorem of Minkowski will give some information about how closely an irrational number can be approximated by rationals.

Let us now return to Hajos proof of the fundamental theorem of Minkowski. Consider the  $x - y$  plane cut up into an infinite chessboard with the basic square of area 4 determined by  $|x| \leq 1, |y| \leq 1$ . We now cut up the chessboard along the edges of the squares and superimpose all the squares that contain parts of the region  $R$ . We have now compressed an area  $> 4$  into a region of area 4. This implies that there will be some overlapping, i.e., one can stick a pin through the square so as to pierce  $R$  into two points say  $V_1$  and  $V_2$ . Now reassemble the region and let the points  $V_1$  and  $V_2$  be the vectors  $\vec{V}_1$  and  $\vec{V}_2$ . Consider the fact that the  $x$  and  $y$  coordinates of  $V_1$  and  $V_2$  differ by a multiple of 2. We write  $V_1 \equiv V_2 \pmod{2}$ , which implies  $\frac{1}{2}(V_1 - V_2) \equiv 0 \pmod{1}$ . Thus  $\frac{1}{2}(V_1 - V_2)$  is a lattice point different from  $O$  (since  $V_1 \neq V_2$ ) in  $R$ .

The fundamental theorem of Minkowski can easily be generalized to  $n$ -dimensional space. Indeed we need only replace 4 in the fundamental theorem of Minkowski by  $2n$  and Hajos' proof goes through. Many extensions and refinements of the fundamental theorem of Minkowski have been given. I shall return to some of them later.

One of Polya's earliest papers has the long and curious title "Zahlentheoretisches und Wahrscheinlichkeitstheoretisches über die Sichtweite in Walde und durch Schneefall". A proof of Polya's main result in this paper can be greatly simplified and somewhat refined using the fundamental theorem of Minkowski. The problem is this.

Suppose every lattice point other than  $O$  is surrounded by a circle of radius  $r \leq \frac{1}{2}$  (a tree in a forest). A man stands at  $O$ . In direction  $\theta$  he can see a distance  $f(r, \theta)$ . What is the furthest he can see in any direction? That is, determine

$$F(r) = \max_{\theta} f(r, \theta)$$

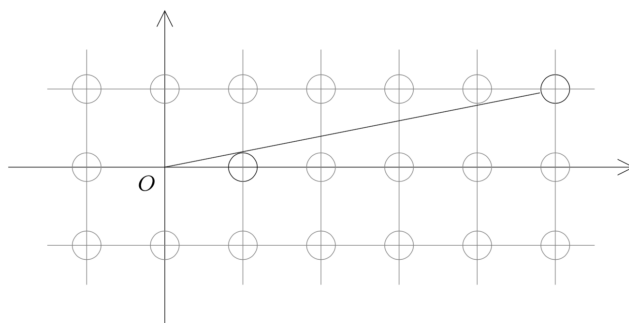


FIGURE 5

By looking past the circle centered at  $(1, 0)$  (Figure 5), we can see almost a distance  $\frac{1}{r}$ . On the other hand we can prove that  $F(r) \leq \frac{1}{r}$ . For suppose that we can see a distance  $F(r)$  in direction  $\theta$ . About this line of vision construct a rectangle with side  $2r$ . This rectangle contains no lattice point, for otherwise the tree centered at such a lattice point would obstruct our line of vision; see Figure 6.

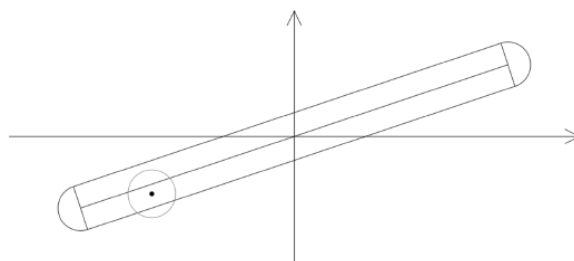


FIGURE 6

Hence, by the fundamental theorem of Minkowski  $4F(r)r \leq 4$  and  $F(r) \leq \frac{1}{r}$  as required. Note that no lattice point can be in either semicircle in the diagram. This enables us to improve slightly on Polya's result. I shall leave the details as an exercise.

A more significant application of the fundamental theorem of Minkowski concerns the possibility of solving in integers a set of linear inequalities.

Consider the inequalities

$$\begin{aligned} |a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n| &\leq \lambda_1, \\ |a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n| &\leq \lambda_2, \\ &\vdots \\ |a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n| &\leq \lambda_n, \end{aligned}$$

where the  $a_{ij}$  are real numbers and the  $\lambda_1, \lambda_2, \dots, \lambda_n$  are positive numbers. The problem is to find sufficient conditions for the existence of integers  $x_1, \dots, x_n$ , not all 0 satisfying the system. The fundamental theorem of Minkowski can be used to prove that a solution will exist provided the determinant  $\det(a_{ij})$  of the coefficients is, in absolute value, less than the product  $\lambda_1 \cdot \lambda_2 \cdots \lambda_n$ . This comes about in the following way. Geometrically, the inequalities determine an  $n$ -dimensional parallelepiped whose volume (or content) is

$$\frac{1}{\det(a_{ij})} \cdot 2^n \cdot \lambda_1 \cdot \lambda_2 \cdots \lambda_n.$$

If  $\lambda_1 \cdot \lambda_2 \cdots \lambda_n > \det(a_{ij})$  then the content exceeds  $2^n$  and so contains a lattice point different from  $O$ .

A very recent analogue of the fundamental theorem of Minkowski is the following. Let  $R$  be a convex region, not necessarily symmetric about  $O$ , but having its centroid at  $O$ . If its area exceeds  $\frac{9}{2}$ , then it contains a lattice point not  $O$ . The constant  $\frac{9}{2}$  is again best possible, but an  $n$ -dimensional analogue of this result is unknown.

The following is a conjectured generalization of the fundamental theorem of Minkowski, which we have unfortunately been unable to prove. Perhaps you will be able to prove or disprove it. Let  $R$  be a convex region containing the origin and defined by  $r = f(\theta)$ ,  $0 \leq \theta < 2\pi$ . If

$$\int_0^\pi f(\theta)f(\theta + \pi)d\theta > 4$$

then  $R$  contains a nontrivial lattice point. For symmetrical regions  $f(\theta) = f(\theta + \pi)$ , and the conjecture reduces to the fundamental theorem of Minkowski.

Here is a somewhat related and only partially solved problem. Let  $M(n)$  be defined as the smallest number such that any convex region of area  $M(n)$  can be placed so as to cover  $n$  lattice points. Clearly  $M(1) = 0$ . It is not difficult to show that  $M(2) = \frac{\pi}{4}$ , i.e., any convex region whose area exceeds that of a circle of diameter 1 can be used to cover 2 lattice points. To determine  $M(3)$  already seems difficult. What one can easily prove is that  $M(n) \leq n - 1$  and we conjecture the existence of a positive constant  $c$  such that  $M(n) < n - c\sqrt{n}$ .

---

This page titled [1.8: Geometry of Numbers](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Leo Moser \(The Trilla Group\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform; a detailed edit history is available upon request.



## CHAPTER OVERVIEW

### Back Matter

[Index](#)

[Detailed Licensing](#)

## Index

---

### C

congruences

[1.5: Congruences](#)

### D

Diophantine Equations

[1.6: Diophantine Equations](#)

### I

Irrational Numbers

[1.4: Irrational Numbers](#)

## Detailed Licensing

---

### Overview

**Title:** [An Introduction to the Theory of Numbers \(Moser\)](#)

**Webpages:** 25

**All licenses found:**

- [Undeclared](#): 64% (16 pages)
- [CC BY 4.0](#): 36% (9 pages)

### By Page

- [An Introduction to the Theory of Numbers \(Moser\) - CC BY 4.0](#)
  - [Front Matter - Undeclared](#)
    - [TitlePage - Undeclared](#)
    - [InfoPage - Undeclared](#)
    - [Table of Contents - Undeclared](#)
  - [Chapters - Undeclared](#)
    - [Front Matter - Undeclared](#)
      - [TitlePage - Undeclared](#)
      - [InfoPage - Undeclared](#)
      - [Table of Contents - Undeclared](#)
      - [Licensing - Undeclared](#)
    - [1.1: Compositions and Partitions - CC BY 4.0](#)
    - [1.2: Arithmetic Functions - CC BY 4.0](#)
    - [1.3: Distribution of Primes - CC BY 4.0](#)
    - [1.4: Irrational Numbers - CC BY 4.0](#)
    - [1.5: Congruences - CC BY 4.0](#)
    - [1.6: Diophantine Equations - CC BY 4.0](#)
    - [1.7: Combinatorial Number Theory - CC BY 4.0](#)
    - [1.8: Geometry of Numbers - CC BY 4.0](#)
    - [Back Matter - Undeclared](#)
      - [Index - Undeclared](#)
      - [Glossary - Undeclared](#)
      - [Detailed Licensing - Undeclared](#)
  - [Back Matter - Undeclared](#)
    - [Index - Undeclared](#)

## Index

---

### C

congruences

[1.5: Congruences](#)

### D

Diophantine Equations

[1.6: Diophantine Equations](#)

### I

Irrational Numbers

[1.4: Irrational Numbers](#)

## Detailed Licensing

---

### Overview

**Title:** [An Introduction to the Theory of Numbers \(Moser\)](#)

**Webpages:** 25

**All licenses found:**

- [Undeclared](#): 64% (16 pages)
- [CC BY 4.0](#): 36% (9 pages)

### By Page

- [An Introduction to the Theory of Numbers \(Moser\) - CC BY 4.0](#)
  - [Front Matter - Undeclared](#)
    - [TitlePage - Undeclared](#)
    - [InfoPage - Undeclared](#)
    - [Table of Contents - Undeclared](#)
  - [Chapters - Undeclared](#)
    - [Front Matter - Undeclared](#)
      - [TitlePage - Undeclared](#)
      - [InfoPage - Undeclared](#)
      - [Table of Contents - Undeclared](#)
      - [Licensing - Undeclared](#)
    - [1.1: Compositions and Partitions - CC BY 4.0](#)
    - [1.2: Arithmetic Functions - CC BY 4.0](#)
    - [1.3: Distribution of Primes - CC BY 4.0](#)
    - [1.4: Irrational Numbers - CC BY 4.0](#)
    - [1.5: Congruences - CC BY 4.0](#)
    - [1.6: Diophantine Equations - CC BY 4.0](#)
    - [1.7: Combinatorial Number Theory - CC BY 4.0](#)
    - [1.8: Geometry of Numbers - CC BY 4.0](#)
    - [Back Matter - Undeclared](#)
      - [Index - Undeclared](#)
      - [Glossary - Undeclared](#)
      - [Detailed Licensing - Undeclared](#)
  - [Back Matter - Undeclared](#)
    - [Index - Undeclared](#)