DISTRIBUTION OF ADDITIVE FUNCTIONS IN ALGEBRAIC NUMBER FIELDS

by

Garry Hughes, B.Sc. (Hons.) (Adelaide).

Thesis submitted for the degree of
Master of Science in the Faculty of
Mathematical Sciences at the
University of Adelaide.

July 1987

Department of Pure Mathematics.

CONTENTS

SUMMARY

1. Background

This thesis is based on ideas drawn from classical probabilistic number theory, from the work of Novoselov [2], and from the relevant work on algebraic number fields.

Classical probabilistic number theory (as described in Elliott [1], for example) is concerned with the distribution of arithmetic functions on the ring of (rational) integers, $\mathbb{Z}$. Two well known results in this area are the Hardy-Ramanujan and the Erdös-Wintner theorems. The Hardy-Ramanujan theorem states that, in some sense, every integer n has about $\log\log n$ prime divisors, and the Erdös-Wintner theorem gives conditions under which additive functions have limiting distributions. The original proofs of both results were subsequently considerably simplified by using a result known as the Turán-Kubilius inequality. Although results in this field have a definite probabilistic flavour, it has not proved easy to establish them by a direct appeal to the theory of probability.

Novoselov [2] developed a probability space which provides a natural framework for developing results of probabilistic number theory from results of probability. For example, using standard results from probability theory and some arithmetic estimates (which amount to the Turán-Kubilius inequality) he obtained the Hardy-Ramanujan and Erdös-Wintner theorems.

Many of the results of probabilistic number theory have been generalized to results concerning the distribution of additive functions on the ideals of the ring, $\mathcal{D}$, of integers of an algebraic number field (see Prachar [3], for example). However, work in this area has not used a probabilistic framework as fully as in the classical case of $\mathbb{Z}$.

2. Aims

The aim of this thesis is to set up a space for probabilistic number theory in algebraic number fields analogous to that of Novoselov [2] for $\mathbb{Z}$ and to apply his approach to develop analogues in $\mathcal{D}$ of the Hardy-Ramanujan and Erdös-Wintner theorems. We endeavour to produce as much as possible without the use of sieve results.

3. Contents

Chapter 1 of this thesis is an introduction to the background outlined above, and Chapter 2 gathers together some preliminary material. In Chapter 3 we obtain an analogue of the Turán-Kubilius inequality in $\mathcal{D}$. For this purpose we estimate the number of elements of an ideal which lie in a multiple of the fundamental domain of $\mathcal{D}$ (viewed as a lattice).

In Chapter 4 we construct a probability space Ω, containing $\mathcal{D}$, using two different approaches. One approach is analogous to that in Novoselov [2]. The other views Ω as the product of the completions of $\mathcal{D}$ with respect to its non-Archimedean valuations and this enables us to simplify some proofs.

In Chapter 5 we prove versions of the Hardy-Ramanujan and Erdös-Wintner theorems for additive functions on the principal ideals of $\mathcal{D}$. Some examples are discussed.

In Chapter 6 we consider additive functions on all the ideals of $\mathcal{D}$ (not just the principal ideals). We prove Prachar's version of the Hardy-Ramanujan theorem (see Prachar [3]) by using the results of Chapter 5 and the correspondence between the ideals of $\mathcal{D}$ in a given class and certain elements of $\mathcal{D}$.

References

[1] Elliott,P.D.T.A., *Probabilistic number theory, Volumes I and II.* Springer -Verlag, New York (1979).

[2] Novoselov,E.V., *A new method in probabilistic number theory.* Amer. Math. Soc. Translations (2) 52 (1966), pp. 217-275.

[3] Prachar,K., *Verallgemeinerung eines Satzes von Hardy und Ramanujan auf algebraische Zahlkörper.* Monatsh. Math. 56 (1952), pp. 229-232.

STATEMENT

a) This thesis contains no material which has been accepted for the award of any other degree or diploma in any University and, to the best of my knowledge, contains no material previously published or written by another person, except where due reference is made in this thesis.

b) I consent to this thesis being made available for photocopying and loan if applicable, if accepted for the award of the degree.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Dr. Jane Pitman for her invaluable advice, guidance and constructive criticisms.

I would also like to thank Dr. Keith Matthews for several interesting conversations about some of the topics in this thesis.

Finally, thanks to Carolyn Schultz for her excellent typing (and patience), to Vicki Schofield, and to the hackers in room 7.

CHAPTER 1

INTRODUCTION

This thesis combines ideas from three areas : firstly, classical probabilistic number theory, secondly, the probabilistic framework for probabilistic number theory developed by Novoselov [1], and thirdly, the work of various authors on probabilistic number theory in algebraic number fields.

In this introductory chapter we will discuss, briefly, some typical results in these areas and their history. We then discuss this thesis itself. We will rely heavily, in Section 1 below, on the excellent historical introduction to classical probabilistic number theory given in Elliott [1].

1. Classical Probabilistic Number Theory.

Let us begin with some definitions (which will apply throughout the introduction only). For a set $\mathcal{A}$ of real numbers and an integer $n \geq 1$ we let

$$\nu_n\{m : m \in \mathcal{A}\} = \frac{1}{n}\#\{m : 1 \leq m \leq n \text{ and } m \in \mathcal{A}\} \tag{1}$$

be the frequency of integers from 1 to n which are in $\mathcal{A}$ (as usual # denotes the number of elements in a finite set).

A function f from the positive integers $\mathbb{Z}^+$, to the real numbers $\mathbb{R}$, is called **additive** if, for any relatively prime positive integers m and n, we have

$$f(mn) = f(m) + f(n). \tag{2}$$

Furthermore, f is called **strongly additive** if, in addition to property (2) above, we have for all primes p and positive integers r,

$$f(p^r) = f(p). \tag{3}$$

Such functions are completely determined by their values on prime powers. If f is additive we have

$$f(n) = \sum_{p^r \| n} f(p^r), \tag{4}$$

where the sum is over the prime powers which exactly divide n; and if f is strongly additive,

$$f(n) = \sum_{p|n} f(p). \tag{5}$$

Many of the usual functions of number theory are additive (for example $\omega(n)$, the number of prime divisors of n) or are closely related to additive functions (positive multiplicative functions are exponentials of additive functions). A problem of long standing interest is to determine the behaviour of these additive functions for large values of n in some average sense (this behaviour, hopefully, being smoother than that of the function). One of the earliest, non-trivial, results in this direction was proved by Hardy and Ramanujan [1], in 1917. Among other things they established the following result.

Classical Hardy-Ramanujan Theorem

Let $\theta(n)$ be a function of n such that $\theta(n) \to \infty$ as $n \to \infty$. Then

$$\nu_n\{m : |\omega(m) - \log\log n| \geq \theta(n)\sqrt{\log\log n}\} \tag{6}$$

tends to zero as $n \to \infty$.

This result may be interpreted as saying that almost all integers n have $\log\log n$ prime divisors. It was proved by an arithmetic method, by establishing a precise upper bound for the number of integers from 1 to n, with a given number of prime divisors. The result is of an essentially probabilistic nature, resembling the Law of Large Numbers in probability theory.

In 1934, Turán gave a new proof of the Hardy-Ramanujan Theorem by way of the estimate

$$\frac{1}{n}\sum_{m=1}^{n} (\omega(m) - \log\log n)^2 \leq c_1 \log\log n, \tag{7}$$

(c_1 independent of n). The argument Turán used was similar to that used to obtain the Tchebycheff inequality in probability theory. The result (7) was later extended by Kubilius and there are now a wide class of results, like (7), called Turán-Kubilius inequalities. For example, we have the following (see Elliott [1] Chapter 4 for a proof),

A Classical Turán-Kubilius Inequality

Let f be a strongly additive real function and, for $n \geq 1$ put

$$\begin{aligned} A(n) &= \sum_{p \leq n} \frac{f(p)}{p} \\ B(n) &= \sum_{p \leq n} \frac{(f(p))^2}{p} \end{aligned} \tag{8}$$

where the sums are over primes $\leq n$. There is a constant c_2 (independent of n and f) such that

$$\frac{1}{n} \sum_{m=1}^{n} (f(m) - A(n))^2 \leq c_2 B(n).$$

The next major result was proved in 1938 by Erdös [1] (and the converse by Erdös and Wintner[1] in 1939).

Classical Erdös-Wintner Theorem

Let f be an additive function and suppose the following two series converge:

$$\sum_{p} \frac{f'(p)}{p}, \quad \sum_{p} \frac{(f'(p))^2}{p} \tag{9}$$

where

$$f'(p) = \begin{cases} f(p) & \text{if } |f(p)| < 1 \\ 1 & \text{if } |f(p)| \geq 1. \end{cases}$$

Then, there is a left-continuous function F, such that

$$F(\lambda) = \lim_{n \to \infty} \nu_n \{m : f(m) < \lambda\} \tag{10}$$

for every λ at which F is continuous. (The converse also holds).

Furthermore, F is continuous for all λ if and only if the following series diverges,

$$\sum_{f(p) \neq 0} \frac{1}{p}.$$

This theorem was, also, proved using no probability theory as such, but the series that are required to converge in (9) are similar to those that are required to converge in the Three-Series Theorem of Kolmogorov in probability theory (see Chapter 2 Section 3 below).

The next fruitful observation was provided by Kac. It is this: "Whether an integer n is divisible by a prime p is independent of whether it is divisible by a different prime q". Using this notion, in 1940 Erdös and Kac [1] proved the following.

Classical Erdös-Kac Theorem

Let f be a strongly additive real function such that $|f(p)| \leq 1$ for all primes p. Let $A(n)$ and $B(n)$ be as defined in (8) above. If $B(n) \to \infty$ as $n \to \infty$ then for each real λ,

$$\nu_n\{m : f(m) - A(n) \leq \lambda\sqrt{B(n)}\}$$

converges to

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-w^2/2}\, dw \tag{11}$$

as $n \to \infty$.

The functions $A(n)$ and $B(n)$ can be interpreted as, respectively, the expectation and variance of f, and the similarity of the above result to the Central Limit Theorem of probability is clear. Erdös and Kac proved their result by using the Central Limit Theorem and a sieve inequality of the type developed by Brun in the 1920's (for studying the distribution of primes).

Due to the similarity of all the above results with results in probability theory, many authors have attempted proofs using the tools of that theory as much as possible. To put the matter into its historical perspective, we should note that the first axiomatic foundation for probability theory to be widely accepted, had only been presented by Kolmogorov in 1933.

An obvious choice for a probability measure on subsets $\mathcal{A}$, of $\mathbb{Z}$, is the **density of $\mathcal{A}$,**

$$\pi(\mathcal{A}) = \lim_{n\to\infty} \nu_n\{m : m \in \mathcal{A}\}, \tag{12}$$

(when this exists). This choice embodies the idea of Kac, for if $p \neq q$ are primes,

$$\pi(m : p|m \text{ and } q|m) = \frac{1}{pq} = \pi(m : p|m)\,\pi(m : q|m).$$

Unfortunately, π is not a probability measure and we cannot use the theory of probability directly. (Among its many short comings, π is not countably additive. For example $\pi(\cup_n\{n\}) = \pi(\mathbb{Z}^+) = 1$ but $\sum_n \pi(\{n\}) = 0$).

In work of the 1950's, Kubilius dealt with this difficulty by constructing an appropriate finite probability space, in which truncated additive functions,

$$\sum_{p\leq r,\; p|n} f(p),$$

could be studied by using independent functions in that space. Kubilius improved and extended all of the classical results of probabilistic number theory described above, as well as proving many new results. The monograph, Kubilius [1], in which these researches are presented, is still a standard work and, in many areas, is not superseded by Elliott [1].

An essential result needed by Kubilius in the construction of the finite probability space has become known as the "Fundamental Lemma of Kubilius". It is essentially a sieve inequality and was proved by using the sieve method developed in the 1940's by Selberg. The general form of this inequality is a little cumbersome to state here (see Kubilius [1], Lemma 1.6), but the following simple version (taken from Philipp [1], Lemma 5.1.1) shows the nature of the result.

Fundamental Lemma of Kubilius

Let $r = r(N)$ be any integer valued function of the integer N, with $\log r/\log N \to 0$ as $N \to \infty$, and let $2 = p_1 < \cdots < p_t \leq r$ be the primes not exceeding r. If $\alpha_1, \cdots, \alpha_t$ are non-negative integers such that

$$p_1^{\alpha_1} \cdots p_t^{\alpha_t} < \sqrt{N},$$

then

$$\#\{1 \leq m \leq N : p_i^{\alpha_i} \| m,\;\; i = 1, \cdots, t\}$$

$$= \frac{N}{p_1^{\alpha_1} \cdots p_t^{\alpha_t}} \prod_{p \le r} \left(1 - \frac{1}{p}\right) \left\{1 + O\left(\exp\left(\frac{-\log N}{36 \log r}\right)\right)\right\},$$

with an absolute O-constant.

A different approach to the problem of interpreting (12) as a probability measure and using probability theory was developed by Novoselov in the early 1960's. We will discuss this approach in the next section.

For a full discussion of the extensive further developments of probabilistic number theory in the 1960's and 1970's see Elliott [1].

2. Novoselov's Space.

In a series of papers, Novoselov constructed a probability space Ω, which seems to be a natural one for proving results of probabilistic number theory (like the theorems of Hardy-Ramanujan, Erdös-Wintner and Erdös-Kac above). See, especially, Novoselov [1], and the references contained there. Another exposition is given in Babu [1].

Novoselov's space is the completion of $\mathbb{Z}$ with respect to a metric topology in which the basic open sets are the arithmetic progressions. This space is equivalent to the following space:

$$\Omega = \prod_p \mathbb{Z}_p,$$

which is the Cartesian product of the completions, $\mathbb{Z}_p$, of $\mathbb{Z}$ with respect to the p-adic valuations. (This equivalence is a special case of the results of Chapter 4, Section 4 below). Ω has a normalized Haar measure P, and we can therefore use the results of probability theory directly.

In the space Ω, it is easy to extend the usual notion, in $\mathbb{Z}$, of divisibility by a prime p. If $\underset{\sim}{x} \in \Omega$ we say $p|\underset{\sim}{x}$ if the p-th component of $\underset{\sim}{x}$ is non-zero and has p-adic valuation $\le 1/p$. In this case the set $\{\underset{\sim}{x} \in \Omega : p|\underset{\sim}{x}\}$ has probability $1/p$ and the idea of Kac is easy to formalize. For, if $p \ne q$ are primes,

$$P(\underset{\sim}{x} \in \Omega : p|\underset{\sim}{x} \text{ and } q|\underset{\sim}{x}) = P(\underset{\sim}{x} \in \Omega : p|\underset{\sim}{x})P(\underset{\sim}{x} \in \Omega : q|\underset{\sim}{x}).$$

In this way the probability measure P mimics the density function π. It is now possible to establish some of the results quoted in the last section by a direct appeal to the theory of probability. As an example, we outline, for strongly additive functions, Novoselov's proof of the Erdös-Wintner Theorem (see Novoselov [1], Proposition 46).

A strongly additive function f may be extended to a function on Ω by defining a new function,

$$\overline{f}(\underset{\sim}{x}) = \sum_{p} \overline{f}_p(\underset{\sim}{x}) \tag{13}$$

where

$$\overline{f}_p(\underset{\sim}{x}) = \begin{cases} f(p) & \text{if } p|\underset{\sim}{x} \\ 0 & \text{if } p \not| \underset{\sim}{x}. \end{cases}$$

There is no *a priori* guarantee that $\overline{f}(\underset{\sim}{x})$ will even converge, because it is possible that $p|\underset{\sim}{x}$ for infinitely many primes p. However, the functions $\overline{f}_p(\underset{\sim}{x})$ are independent functions on Ω (in the probabilistic sense) and the convergence of the series in (9) is exactly what is needed to apply Kolmogorov's Three Series Theorem to the functions $\overline{f}_p(\underset{\sim}{x})$. In this way Novoselov deduced that $\overline{f}(\underset{\sim}{x})$ converges almost everywhere on Ω. He also showed that the distribution function of $\overline{f}(\underset{\sim}{x})$,

$$P(\underset{\sim}{x} \in \Omega : \overline{f}(\underset{\sim}{x}) < \lambda),$$

coincides precisely with the density,

$$\pi(m : f(m) < \lambda),$$

and the Erdös-Wintner Theorem (sufficiency part) was proved. Novoselov made use of a few arithmetic estimates (which amount to the Turán-Kubilius inequality) but no sieve results were used. Because of this, he obtained no more than a special case of the Erdös-Kac Theorem (see Example 2 of Section 6 in Novoselov [1]). Some sort of sieve result seems essential in obtaining the Erdös-Kac Theorem (see Elliott [1], Chapters 3 and 12). In this context, it is interesting to quote Mackey [1] (see page 40): " It is almost certainly true that the results of Kac and his collaborators can be deduced from this observation (*that the functions $\overline{f}_p$ are independent)* and the known properties of independent functions.", (my parenthetical comment). It appears that Mackey underestimated the difficulty of the transition between density results and the probability P. Perhaps, "almost certainly" should be interpreted in the technical sense.

Several authors have made use of the ideas of Novoselov in further studies. Notable among these is Babu (see [1], [2] and the references listed there).

3. Probabilistic Results in Algebraic Number Fields.

Many of the results of Section 1 have been extended to results about the ring of integers $\mathcal{D}$, of an algebraic number field. A natural frequency to use here is

$$\nu_n\{\mathcal{I} : \mathcal{I} \in \mathcal{A}\} = \frac{1}{n}\#\{\mathcal{I} : N(\mathcal{I}) \leq n \text{ and } \mathcal{I} \in \mathcal{A}\} \tag{14}$$

which counts the number of ideals $\mathcal{I}$, of $\mathcal{D}$, with norm no larger than n, which lie in a set $\mathcal{A}$ of ideals.

In 1952, Prachar [1] proved a version of the Hardy-Ramanujan Theorem for ideals. This was later extended by Fluch [1] as follows.

Ideal Hardy-Ramanujan Theorem

Let $\omega(\mathcal{I})$ be the number of prime ideals dividing the ideal $\mathcal{I}$ and suppose that $\theta(n) \to \infty$ as $n \to \infty$. Then,

$$\nu_n\{\mathcal{I} : |\omega(\mathcal{I}) - \log\log n| \geq \theta(n)\sqrt{\log\log n}\}$$

tends to zero as $n \to \infty$.

de Kroon [1] investigated additive functions restricted to the principal ideals of $\mathcal{D}$ (see Chapter 2 Section 1 below, for a definition of these functions), and, in particular, investigated an analogue of the Erdös-Kac Theorem. However, his paper does not seem to provide a firm probabilistic foundation for his results (see Chapter 5 Section 1 below, for a discussion of this).

An alternative definition of frequency to that in (14) is used by Rieger [2]. For a set $\mathcal{A}$, of algebraic integers, let

$$\nu'_n\{d : d \in \mathcal{A}\} = \frac{1}{n}\#\{d \in \mathcal{D} : |d| < n^{1/s} \text{ and } d \in \mathcal{A}\} \tag{15}$$

where $|d| < n^{1/s}$ means that each conjugate of d is, in absolute value, smaller than $n^{1/s}$ (s being the degree of the number field). Using this definition of frequency, Rieger proved an Erdös-Kac type result:

$$\nu_n' \{d \in \mathcal{D} : \ |d| < n^{1/s}, \omega(d) - \log\log n < \lambda\sqrt{\log\log n}\}$$
$$\to \frac{c_3}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-w^2/2}\, dw$$

as $n \to \infty$, for a constant c_3, where $\omega(d)$ is the number of prime divisors of the principal ideal generated by d. (See Satz 2 of Rieger [2], where an estimate of the rate of convergence is also given).

Both Rieger and de Kroon used analogues of classical sieve methods in $\mathcal{D}$. (Rieger [1], Satz 14 is a Selberg sieve inequality and de Kroon [1], Lemma 1 is a Brun sieve inequality similar to the fundamental Lemma of Kubilius). Several authors have studied sieve results in algebraic number fields for their own sake (for example see Rieger [1], Schaal [1] and Wilson [1]).

As far as I am aware, no author has proved an Erdös-Wintner theorem in $\mathcal{D}$.

In each case, the proofs of the aforementioned extensions of classical results to $\mathcal{D}$ are analogous to the original proofs in $\mathbb{Z}$. In general, work in this area has not used a probabilistic framework as fully as in the case of $\mathbb{Z}$.

4. This Thesis.

The aim of this thesis is to set up a space for probabilistic number theory in algebraic number fields analogous to that of Novoselov [1] in $\mathbb{Z}$, and to apply his approach to develop analogues of the Hardy-Ramanujan and Erdös-Wintner Theorems in $\mathcal{D}$. We endeavour to produce as much as possible without the use of sieve results.

In Chapter 2 we gather together some preliminary material. We firstly review some of the standard results about algebraic number fields and then establish a few arithmetic estimates for the distribution of prime ideals. Finally, we collect the probability theory we need.

In Chapter 3 we obtain an analogue of the classical Turán-Kubilius inequality (as stated in Section 1 of this introduction) for additive functions on the ideals of $\mathcal{D}$. For this purpose we estimate the number of elements of an ideal which lie in a multiple of a fundamental domain of $\mathcal{D}$ (viewed as a lattice in $\mathbb{R}^s$, where s is the degree of $\mathcal{D}$).

In Chapter 4 we construct a probability space Ω, containing $\mathcal{D}$, using two different approaches. These approaches are the analogues of the two ways of constructing Novoselov's space discussed in Section 2 of this introduction. We then establish the equivalence of these two spaces from a topological and measure point of view.

In Chapter 5 we prove versions of the Hardy-Ramanujan and Erdös-Wintner Theorems for additive functions restricted to the principal ideals of $\mathcal{D}$. Some examples are discussed.

In Chapter 6 we consider additive functions on all the ideals of $\mathcal{D}$. We prove the Ideal Hardy-Ramanujan Theorem by using the results of Chapter 5 and the correspondence between ideals of $\mathcal{D}$, in a given class, and certain elements of $\mathcal{D}$. We then discuss a version of the Turán-Kubilius inequality for ideals and the consequent improved version of the Ideal Hardy-Ramanujan Theorem. Some speculations about directions of further study are then given.

Finally, a word about notation and presentation. We will use the O-notation of Landau freely and, occasionally, the $\ll$-notation of Vinogradov. The symbol # will always mean the number of elements in a finite set. Theorems, corollaries and lemmas are numbered consecutively within a chapter (thus, Corollary 4.3 would follow Lemma 4.2 in Chapter 4). Within a chapter, a numbered line is referred to as, say, (12) and a numbered line in another chapter is referred to as, say, (2.4) (if we want line (4) of Chapter 2). The symbol ■ will be used to mark the end of a proof.

CHAPTER 2

PRELIMINARIES

In this chapter we will collect some definitions, notations and results needed for our later work. Standard definitions and basic results of ideal theory and algebraic number theory will be taken from Stewart and Tall [1] (especially from Chapters 2 and 5). Results on probability will mainly be taken from Halmos [1] and Rényi [1].

The notation introduced here will remain throughout the thesis.

1. Algebraic Number Fields.

Let $\mathbb{Z}$ and $\mathbb{R}$ denote, respectively, the sets of integers and real numbers.

Let $\mathbb{K}$ be a fixed algebraic number field of degree s, that is, a finite extension of the field of rational numbers, of degree s.

Let $\mathcal{D}$ be the ring of integers of $\mathbb{K}$, that is, the set of complex numbers in $\mathbb{K}$ which are the zeros of monic polynomials with coefficients in $\mathbb{Z}$. The elements of $\mathcal{D}$ are called **algebraic integers** and the elements of $\mathbb{Z}$ **rational integers.**

There are s distinct one-to-one ring homomorphisms which map $\mathbb{K}$ into the complex numbers and which are the identity on rational numbers. These maps are the **conjugate maps** and they either map $\mathbb{K}$ into $\mathbb{R}$ or occur in complex conjugate pairs. We list them as

$$\sigma_1, \sigma_2, \cdots, \sigma_{r_1}, \sigma_{r_1+1}, \overline{\sigma}_{r_1+1}, \cdots, \sigma_{r_1+r_2}, \overline{\sigma}_{r_1+r_2}$$

where $\sigma_i(\mathbb{K}) \subseteq \mathbb{R}$ for $i = 1, \cdots, r_1$ only, and the bar denotes complex conjugation. We then have

$$s = r_1 + 2r_2.$$

For $a \in \mathbb{K}$ we define the **norm map** from $\mathbb{K}$ to $\mathbb{R}$ by

$$N(a) = \prod_{i=1}^{s} \sigma_i(a). \qquad (1)$$

We note that if $a \in \mathcal{D}$ is an algebraic integer then $N(a) \in \mathbb{Z}$; and also that the following multiplicative property holds: for a, $b \in \mathbb{K}$

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

If $\mathcal{I}$ is any non-zero ideal of $\mathcal{D}$ with $\mathcal{I} \neq \mathcal{D}$ then $\mathcal{I}$ has a decomposition into the product of prime ideals

$$\mathcal{I} = \mathcal{P}_1^{\alpha_1} \cdots \mathcal{P}_r^{\alpha_r}$$

where $\alpha_1, \cdots, \alpha_r$ are positive rational integers and this decomposition is unique except for the order of the factors (we may also write $\mathcal{D}$ itself as $\mathcal{P}^0$).

Throughout this thesis $\mathcal{P}$ and $\mathcal{Q}$ (with or without subscripts) will denote prime ideals.

If $\mathcal{I}$ and $\mathcal{L}$ are ideals we say

$$\mathcal{L}|\mathcal{I} \text{ if } \mathcal{I} = \mathcal{L}\mathcal{K}$$

for some ideal $\mathcal{K}$, and note that

$$\mathcal{L}|\mathcal{I} \text{ if and only if } \mathcal{I} \subseteq \mathcal{L}.$$

The distinct cosets of $\mathcal{I}$ in $\mathcal{D}$,

$$\mathcal{D}/\mathcal{I} = \{a_1 + \mathcal{I}, \cdots, a_t + \mathcal{I}\},$$

say, form a finite additive group and each algebraic integer $d \in \mathcal{D}$ belongs to one and only one $a_i + \mathcal{I}$ for $i = 1, \cdots, t$. We call

$$\{a_1, \cdots, a_t\}$$

a **set, or system, of representatives mod $\mathcal{I}$**, and if d belongs to $a_i + \mathcal{I}$ we say that d is **congruent to a_i mod $\mathcal{I}$** .

The number, t, of such representatives mod $\mathcal{I}$ is denoted by **the norm of $\mathcal{I}$, $N(\mathcal{I})$**, that is,

$$N(\mathcal{I}) = \#(\mathcal{D}/\mathcal{I}).$$

For $d \in \mathcal{D}$, we let **$<d>$ denote the principal ideal of $\mathcal{D}$ generated by d.** The new concept of norm generalizes that in (1) since

$$N(<d>) = |N(d)|$$

and for ideals $\mathcal{I}$ and $\mathcal{L}$ of $\mathcal{D}$,

$$N(\mathcal{I}\mathcal{L}) = N(\mathcal{I})N(\mathcal{L}).$$

We need a few results on ideals. **We will always assume an ideal is a non-zero ideal of $\mathcal{D}$.**

Lemma 2.1

i) (Chinese Remainder Theorem). Let $\mathcal{P}_1, \cdots, \mathcal{P}_n$ be distinct prime ideals. Let $\alpha_1, \cdots, \alpha_n$ be non-negative rational integers and $b_1, \cdots, b_n$ elements of $\mathcal{D}$. Then there is an algebraic integer $d \in \mathcal{D}$ such that

$$d - b_i \in \mathcal{P}_i^{\alpha_i}$$

for all $i = 1, \cdots, n$.

ii) If $\mathcal{P}$ is a prime ideal of $\mathcal{D}$ then $\mathcal{P}$ contains exactly one rational prime p and

$$N(\mathcal{P}) = p^n$$

for some integer n with $1 \leq n \leq s$.

If $\mathcal{I}$ is any ideal of $\mathcal{D}$ then $N(\mathcal{I}) \in \mathcal{I}$.

iii) There are at most s prime ideals of given norm, t.

Proof

i) See Goldstein [1], Theorem 2.2.13 or Narkiewicz [1], Corollary 3 to Proposition 1.6.

ii) See Stewart and Tall [1], Theorem 5.11.

iii) This result is easily deduced from the index equation of ramification theory (see Goldstein [1], Theorem 5.1.3 or Narkiewicz [1], Theorem 4.1) but to avoid introducing notation unnecessary in the sequel we present a proof here.

From part ii) we may as well assume $t = p^n$ for a rational prime p and $1 \leq n \leq s$. Let

$$<p> = \mathcal{P}_1^{e_1} \cdots \mathcal{P}_r^{e_r} \tag{2}$$

be the decomposition of $<p>$ into prime factors with $e_1, \cdots, e_r \geq 1$. As $p \in \mathcal{P}_i$ for each $i = 1, \cdots, r$ then part ii) gives $N(\mathcal{P}_i) = p^{f_i}$ for some

$1 \leq f_i \leq s$. Thus, if we take norms in (2) we have

$$p^s = p^{e_1 f_1 + \cdots + e_r f_r}$$

and therefore, as $e_i f_i \geq 1$ for $i = 1, \cdots, r$, we have $r \leq s$. Finally, if $\mathcal{P}$ is any prime ideal with $N(\mathcal{P}) = p^n$ then $p \in \mathcal{P}$ and so $\mathcal{P}$ is one of the prime factors in (2) and there are at most s of these.

■

We now introduce the concept of **integral basis**. Any (non-zero) ideal $\mathcal{I}$ of $\mathcal{D}$ (including $\mathcal{D}$ itself) has an integral basis with s elements. That is, there exist $b_1, \cdots, b_s \in \mathcal{I}$ such that any $d \in \mathcal{I}$ can be expressed uniquely as

$$d = \alpha_1 b_1 + \cdots + \alpha_s b_s \tag{3}$$

for $\alpha_1, \cdots, \alpha_s \in \mathbb{Z}$.

Suppose $d_1, \cdots, d_s$ is an integral basis for $\mathcal{D}$. We define the **discriminant** of $\mathcal{D}$ (or $\mathbb{K}$) to be the square of the determinant of the matrix formed by taking the conjugates of the basis,

$$\delta = (\det[\sigma_i(d_j)])^2$$

where $i, j = 1, \cdots, s$. The discriminant, δ, is a non-zero rational integer which is independent of the choice of basis $d_1, \cdots, d_s$ of $\mathcal{D}$.

It is possible to choose an integral basis for an ideal $\mathcal{I}$ which is not too large compared with $N(\mathcal{I})$.

Lemma 2.2

There is a real number $c > 0$, dependent on $\mathbb{K}$, such that if $\mathcal{I}$ is any ideal of $\mathcal{D}$, then $\mathcal{I}$ has an integral basis $b_1, \cdots, b_s$ with

$$|\sigma_i(b_j)| \leq cN(\mathcal{I})^{1/s} \qquad (i, j = 1, \cdots, s).$$

Proof

The important point here is that c is independent of the ideal $\mathcal{I}$ chosen.

Rieger [1] gives the reference Hasse [1], page 406, but the result is not explicitly stated there. However, it is contained in the more general result of Mahler [1] (see Theorem 1) where a constant c is produced which depends explicitly on the degree, s, the number of complex conjugates, r_2 and the discriminant, δ (see equations (26) and (4) of Mahler [1]) (see also Luthar [1] and McFeat [1]).

■

Finally, in this section we introduce the concept of additive functions. We say two (non-zero) ideals $\mathcal{I}$, $\mathcal{L}$ are **relatively prime** if their prime factorizations have no common factors (that is if $\mathcal{P}|\mathcal{I}$ then $\mathcal{P}\not|\,\mathcal{L}$ and vice-versa).

Let f be a function from the set of ideals of $\mathcal{D}$ to the real numbers. We say that **f is additive on the ideals of $\mathcal{D}$** if, for relatively prime ideals $\mathcal{I}$ and $\mathcal{L}$,

$$f(\mathcal{I}\mathcal{L}) = f(\mathcal{I}) + f(\mathcal{L}). \tag{4}$$

For such an f we may write,

$$f(\mathcal{I}) = \sum_{\mathcal{P}^r||\mathcal{I}} f(\mathcal{P}^r) \tag{5}$$

where $\mathcal{P}^r||\mathcal{I}$ means $\mathcal{P}^r|\mathcal{I}$ but $\mathcal{P}^{r+1}\not|\,\mathcal{I}$. We say that **$f$ is strongly additive on the ideals of $\mathcal{D}$** if, as well as (4) above, we have

$$f(\mathcal{P}^r) = f(\mathcal{P}) \tag{6}$$

for any prime ideal $\mathcal{P}$ and positive integer r. In this case we have

$$f(\mathcal{I}) = \sum_{\mathcal{P}|\mathcal{I}} f(\mathcal{P}). \tag{7}$$

Equations (5) and (7) could be taken as defining additive and strongly additive functions. Note that from (4) we have $f(\mathcal{D}) = 0$ for additive functions f.

We say a function h from the ideals of $\mathcal{D}$ to $\mathbb{R}$ is **multiplicative** if

$$h(\mathcal{I}\mathcal{L}) = h(\mathcal{I})h(\mathcal{L})$$

for relatively prime ideals, $\mathcal{I}$ and $\mathcal{L}$. We have an equation like (5) above,

$$h(\mathcal{I}) = \prod_{\mathcal{P}^r||\mathcal{I}} h(\mathcal{P}^r).$$

These concepts of additive, strongly additive and multiplicative functions are the most natural to consider. They agree with the classical definitions (see Chapter 1 Section 1) when $s = 1$ and $\mathcal{D} = \mathbb{Z}$, since all ideals of $\mathbb{Z}$ are principal and we may identify an element of $\mathbb{Z}$ with the ideal it generates. In an algebraic number field $\mathbb{K}$ it is possible for an algebraic integer $d \in \mathcal{D}$ not to have a unique factorization and therefore the natural objects to consider for the definition of additive functions, and so on, are the ideals of $\mathcal{D}$.

2. Some Arithmetic Estimates for Prime Ideals of $\mathcal{D}$.

Lemma 2.3

Let $z \geq 2$ and $X \geq 3$ be real numbers. Then:

i)

$$\sum_{N(\mathcal{P}) \leq z} \frac{1}{N(\mathcal{P})} = \log\log z + O(1),$$

where the sum runs over all prime ideals $\mathcal{P}$ with $N(\mathcal{P}) \leq z$. (A similar convention will apply when such sums are used later.)

ii) Let $\epsilon > -1$. There is a constant c_ϵ dependent only on ϵ and s such that

$$\sum_{N(\mathcal{P}) \leq z} N(\mathcal{P})^\epsilon \leq c_\epsilon \frac{z^{1+\epsilon}}{\log z}.$$

iii)

$$\frac{1}{X}\left(\sum_{N(\mathcal{P}) \leq X^s} \frac{1}{N(\mathcal{P})}\right)^{1/2} \left(\sum_{N(\mathcal{P}) \leq X^s} \frac{1}{N(\mathcal{P})^{1-2/s}}\right)^{1/2} = O(1).$$

iv)

$$\frac{1}{X}\left(\sum_{\mathcal{P} \neq \mathcal{Q},\, N(\mathcal{P})N(\mathcal{Q}) \leq X^s} \frac{1}{N(\mathcal{P})^{1-2/s} N(\mathcal{Q})^{1-2/s}}\right)^{1/2} = O(1),$$

where the double sum runs over all ordered pairs of prime ideals $(\mathcal{P}, \mathcal{Q})$ with $\mathcal{P} \neq \mathcal{Q}$ and $N(\mathcal{P})N(\mathcal{Q}) \leq X^s$. (A similar convention will apply when such sums are used later.)

In each of the above inequalities the constant implied by the O-notation depends on $\mathbb{K}$ but not on X or z.

Proof

i) In fact a stronger result is true. For some constant B, dependent on $\mathbb{K}$,

$$\sum_{N(\mathcal{P})\leq z} \frac{1}{N(\mathcal{P})} = \log\log z + B + O\left(\frac{1}{\log z}\right).$$

This result is proven in Narkiewicz [1] (Lemma 9.2). Alternatively see Fluch [1] or de Kroon [1].

ii) We use the corresponding result for rational primes (see Prachar [2] Satz 4.2): For some constant c'_ϵ, dependent on ϵ,

$$\sum_{p\leq z} p^\epsilon \leq c'_\epsilon \frac{z^{1+\epsilon}}{\log z}. \tag{8}$$

(For $\epsilon = 0$ this is a weak form of the Prime Number Theorem.)

From Lemma 2.1 parts ii) and iii) we have, upon grouping prime ideals of norm a prime, a prime square and so on,

$$\sum_{N(\mathcal{P})\leq z} N(\mathcal{P})^\epsilon \leq s \sum_{j=1}^{s} \sum_{p^j\leq z} p^{j\epsilon}. \tag{9}$$

Suppose, firstly, that $\epsilon \geq 0$. Then for p in the above sum we have $p^{j\epsilon} \leq z^\epsilon$ and so

$$\begin{aligned} \sum_{N(\mathcal{P})\leq z} N(\mathcal{P})^\epsilon &\leq sz^\epsilon \sum_{j=1}^{s} \sum_{p^j\leq z} 1 \\ &\leq s^2 z^\epsilon \sum_{p\leq z} 1 \leq s^2 c'_o z^\epsilon \frac{z}{\log z} \end{aligned} \tag{10}$$

(this last inequality from (8) with $\epsilon = 0$).

On the other hand if $-1 < \epsilon < 0$, then $p^{j\epsilon} \leq p^\epsilon$ for $j = 1, \cdots, s$ and so (9) becomes, again using (8),

$$\sum_{N(\mathcal{P})\leq z} N(\mathcal{P})^\epsilon \leq s^2 \sum_{p\leq z} p^\epsilon \leq s^2 c'_\epsilon \frac{z^{1+\epsilon}}{\log z}.$$

Combining this with (10) gives the result for any $\epsilon > -1$.

iii) We let $z = X^s$ and use part i) (combining main and error term) and part ii) with $\epsilon = 2/s - 1 > -1$. We then see that the left hand side of iii) is

$$O\left(\frac{1}{X}(\log\log X^s)^{1/2}\left(\frac{(X^s)^{2/s}}{\log X^s}\right)^{1/2}\right)$$

$$= O\left(\left(\frac{\log\log X^s}{\log X^s}\right)^{1/2}\right) = O(1).$$

All O- constants depend on $\mathbb{K}$ but not on X.

iv) If $N(\mathcal{P})N(Q) \le X^s$ then one factor, $N(\mathcal{P})$ say, is $\le X^{s/2}$. Let S denote the double sum in iv). We have, crudely,

$$\begin{aligned} S &= \sum_{\mathcal{P}\neq Q,\, N(\mathcal{P})N(Q)\le X^s} \frac{1}{N(\mathcal{P})^{1-2/s}N(Q)^{1-2/s}} \\ &\le 2 \sum_{N(\mathcal{P})\le X^{s/2}} \frac{1}{N(\mathcal{P})^{1-2/s}} \sum_{N(Q)\le X^s/N(\mathcal{P})} \frac{1}{N(Q)^{1-2/s}}. \qquad (11) \end{aligned}$$

We first estimate the inner sum of (11) using part ii) with $z = X^s/N(\mathcal{P})$ and $\epsilon = 2/s - 1$, and using the fact that $N(\mathcal{P}) \le X^{s/2}$ implies

$$\log\frac{X^s}{N(\mathcal{P})} \ge \log X^{s/2}.$$

We obtain,

$$S = O\left(\frac{(X^s)^{2/s}}{\log X^{s/2}} \sum_{N(\mathcal{P})\le X^{s/2}} \frac{1}{N(\mathcal{P})^{1-2/s}}\cdot\frac{1}{N(\mathcal{P})^{2/s}}\right),$$

and, hence, using part i),

$$S = O\left(X^2\frac{\log\log X^{s/2}}{\log X^{s/2}}\right).$$

Finally, then,

$$\frac{1}{X}S^{1/2} = O\left(\left(\frac{\log\log X^{s/2}}{\log X^{s/2}}\right)^{1/2}\right) = O(1),$$

which establishes iv). All O-constants depend on $\mathbb{K}$ but not on X.

This completes the proof of Lemma 2.3.

$\blacksquare$

We now introduce the following convention: A sum $\sum_{\mathcal{P}}$, over all prime ideals, will mean the limit as $z \to \infty$ of the partial sums, $\sum_{N(\mathcal{P}) \leq z}$.

3. Some Probability.

Let Ω be any probability space with σ-field B and measure P such that $P(\Omega) = 1$.

A collection $\mathcal{E}$ of real-valued measurable functions on Ω is **independent** if for any finite sub-collection $g_1, \cdots, g_n \in \mathcal{E}$ and real numbers $\lambda_1, \cdots, \lambda_n$ we have

$$\begin{aligned} &P(x \in \Omega : g_1(x) \leq \lambda_1, \cdots, g_n(x) \leq \lambda_n) \\ &\qquad = P(x \in \Omega : g_1(x) \leq \lambda_1) \cdots P(x \in \Omega : g_n(x) \leq \lambda_n). \qquad (12) \end{aligned}$$

For a real-valued measurable function g on Ω, the **expectation of g** is

$$E(g) = \int_\Omega g \, dP.$$

The main results we will need from probability theory are contained in the following two lemmas.

Lemma 2.4

i) (Tchebycheff's Inequality). For any real-valued measurable function g on Ω and for any real $\lambda > 0$,

$$P(x \in \Omega : |g(x)| \geq \lambda) \leq \frac{1}{\lambda^2} E(g^2).$$

ii) (Kolmogorov's Three Series Theorem). Let $\{g_k\}$ be a sequence of independent functions on Ω. Let

$$g_k^o(x) = \begin{cases} g_k(x) & \text{if } |g_k(x)| < 1 \\ 0 & \text{if } |g_k(x)| \geq 1. \end{cases}$$

Then, the series

$$\sum_{k=1}^{\infty} g_k(x)$$

converges almost everywhere (a.e.) on Ω if and only if the following three series converge:

$$\sum_{k=1}^{\infty} P(x \in \Omega : |g_k(x)| \geq 1),$$

$$\sum_{k=1}^{\infty} E(g_k^\circ),$$

$$\sum_{k=1}^{\infty} (\, E((g_k^\circ)^2) - (E(g_k^\circ))^2 \,).$$

Proof

i) This is a special case of Rényi [1], Theorem 2.11.1.

ii) See Halmos [1], Theorem E of Section 46.

■

Let $\{g_k\}$ be a sequence of real-valued measurable functions on Ω. We say that g_k **converges in probability to** g and write

$$g_k(x) \xrightarrow{P} g(x)$$

if, for every $\lambda > 0$,

$$P(x \in \Omega : |g_k(x) - g(x)| > \lambda) \to 0 \tag{13}$$

as $k \to \infty$.

A function $G : \mathbb{R} \to \mathbb{R}$ is called a **distribution function** if it is non-decreasing and is left continuous, that is, for any λ

$$G(\lambda) = \lim_{\epsilon \to 0^+} G(\lambda - \epsilon),$$

and if it is normalized, that is

$$\lim_{\lambda \to \infty} G(\lambda) = 1, \quad \lim_{\lambda \to -\infty} G(\lambda) = 0. \tag{14}$$

A real-valued measurable function g on Ω is **purely discrete** if it takes, almost everywhere, only a countable set of values. That is, if for some countable set $A \subseteq \mathbb{R}$,

$$P(x \in \Omega : g(x) \in A) = 1.$$

Lemma 2.5 (Lévy)

Let $\{g_k\}$ be a sequence of independent and purely discrete functions on Ω such that

$$g(x) = \sum_{k=1}^{\infty} g_k(x)$$

converges almost everywhere on Ω. Define the maximum jump of g_k to be

$$J_k = \sup_{\lambda} P(x \in \Omega : g_k(x) = \lambda).$$

Then the distribution function of g, $P(x \in \Omega : g(x) < \lambda)$, is continuous for all λ if and only if

$$\sum_{k=1}^{\infty} (1 - J_k)$$

diverges.

Proof

See Elliott [1], Lemmas 1.22 and 1.18 or Lévy [1], Theorem XIII.

■

CHAPTER 3

LATTICE ESTIMATES AND THE TURÁN-KUBILIUS INEQUALITY

This chapter is mainly devoted to obtaining an extension of the classical Turán-Kubilius inequality (see Chapter 1 Section 1 above) to one for strongly additive functions on the ideals of the algebraic integers $\mathcal{D}$ (of fixed number field $\mathbb{K}$). This inequality will provide one of the tools used in Chapter 5 to link the frequency of additive functions with the probability spaces to be developed in Chapter 4. We will prove this inequality in Section 4, below, by using the argument of Elliott[1] (Chapter 4) and a main estimate (Theorem 3.3 below) which gives the number of a special set of representatives mod $<n>$ (that is, modulo the principal ideal generated by a rational integer n) which lie in an ideal, $\mathcal{I}$. In Sections 2 and 3, below, we will view ideals as lattices in $\mathbb{R}^s$ and the special representatives as the lattice points inside a parallelotope in $\mathbb{R}^s$. We use a volume estimate, to be discussed in Section 1 below, for the number of such points to prove Theorem 3.3, the main estimate.

The following notation will remain throughout this chapter.

i) For an $s \times s$ real matrix $D = [d_{ij}]$,

$$|D| = \max\{|d_{ij}| : i, j = 1, \cdots, s\}.$$

Note that then, for any $s \times s$ matrices, D_1 and D_2,

$$|D_1 D_2| \leq s|D_1||D_2|.$$

ii) For a vector u in $\mathbb{R}^s$ with components $u_1, \cdots, u_s$,

$$||u|| = \sqrt{u_1^2 + \cdots + u_s^2} = \text{Euclidean length of } u.$$

iii) $\mathbb{Z}^s$ is the integer lattice in $\mathbb{R}^s$.

iv) ∂B denotes the boundary of a set, B, in $\mathbb{R}^s$.

v) $V_s(B)$ denotes the s-dimensional volume of a set, B, in $\mathbb{R}^s$.

1. Lattice Points in Parallelotopes.

Let M be an $s \times s$ real matrix with row vectors $m_1, \cdots, m_s$. Let L be a parallelotope in $\mathbb{R}^s$ determined by M. Specifically, let

$$\begin{aligned} L &= \{u \in \mathbb{R}^s : -1/2 < m_i \cdot u \leq 1/2,\ i = 1, \cdots, s\} \\ &= \{u \in \mathbb{R}^s : Mu \in (-1/2, 1/2] \times \cdots \times (-1/2, 1/2]\}. \end{aligned}$$

We suppose that $\det M \neq 0$ and note that L has s-dimensional volume

$$V_s(L) = \frac{1}{|\det M|}.$$

Furthermore, any parallelotope

$$\{u \in \mathbb{R}^s : |m_i \cdot u| \leq \lambda_i,\ i = 1, \cdots, s\}$$

with $\lambda_1, \cdots, \lambda_s$ positive real numbers has s-dimensional volume

$$\frac{2^s \lambda_1 \cdots \lambda_s}{|\det M|}.$$

Now we proceed to estimate the number of lattice points of an arbitrary translation of $\mathbb{Z}^s$ which lie in L. The argument is similar to that of Lang [1] (see Chapter 5 Section 2, especially Theorem 2) but we will be more concerned with parallelotopes in $\mathbb{R}^s$ and with the exact nature of any O-constants. A more detailed version of the argument of Lang [1] is found in Marcus [1] (in the proof of Lemma 2 in Chapter 6).

Theorem 3.1

Let $\beta > 0$ be a given real number and M a real $s \times s$ non-singular matrix such that

$$|M| \leq \beta,$$

where $|M|$ is as defined by i) above. Let L be the parallelotope

$$\{u \in \mathbb{R}^s : -1/2 < m_i \cdot u \leq 1/2,\ i = 1, \cdots, s\}$$

where $m_1, \cdots, m_s$ are the rows of M and let, for $a \in \mathbb{R}^s$

$$\Lambda = a + \mathbb{Z}^s$$

be any arbitrary translation of the integer lattice in $\mathbb{R}^s$. Then the number of points, $\#\{\Lambda \cap L\}$, of Λ lying in L satisfies

$$\left| \#\{\Lambda \cap L\} - \frac{1}{|\det M|} \right| \leq \gamma \cdot \frac{|M|}{|\det M|}$$

where $\gamma = \max\{2^{s+1}s^2, 2^{s+1}s(1/2 + s\beta)^{s-1}\}$ does not depend on M or on a.

Proof

For each $b \in \Lambda$ let

$$C_b = \{b + y : y \in (0,1] \times \cdots \times (0,1]\}$$

be the half-open box of volume one in $\mathbb{R}^s$ determined by b. If $b \in \Lambda \cap L$ then C_b either lies in the interior, int$L$, of L or intersects the boundary, ∂L, of L. Therefore we have

$$\begin{aligned}\#\{b \in \Lambda : C_b \subseteq \text{int}L\} \leq V_s(L) &\leq \#\{b \in \Lambda : C_b \subseteq \text{int}L\} \\ &+ \#\{b \in \Lambda : C_b \cap \partial L \neq \phi\}.\end{aligned}$$

Clearly, $\#\{\Lambda \cap L\}$ is also bounded by the terms on the left and right of this inequality and so we have

$$\begin{aligned}|\#\{\Lambda \cap L\} - V_s(L)| &\leq \#\{b \in \Lambda : C_b \cap \partial L \neq \phi\} \\ &= V_s(\cup C_b : b \in \Lambda, C_b \cap \partial L \neq \phi).\end{aligned}$$

The diameter of any C_b is $\sqrt{s}$ and so if $C_b \cap \partial L \neq \phi$ the Euclidean distance $\mathrm{d}(u, \partial L)$ of any point $u \in C_b$ to ∂L can be no greater than $\sqrt{s}$. Therefore

$$|\#\{\Lambda \cap L\} - V_s(L)| \leq V_s(E)$$

where

$$E = \{u \in \mathbb{R}^s : d(u, \partial L) \le \sqrt{s}\}.$$

We know that $V_s(L) = 1/|\det M|$ and so it remains to estimate $V_s(E)$.

We recall that the perpendicular (that is smallest) distance from any point $y \in \mathbb{R}^s$ to the hyperplane $m_i \cdot u = \lambda$ is

$$\frac{|m_i \cdot y - \lambda|}{||m_i||} \qquad (i = 1, \cdots, s),$$

and so, suppose that we expand the parallelotope L up and down in the direction perpendicular to each bounding hyperplane $m_i \cdot u = \pm 1/2$ by a distance $\sqrt{s}$. That is, we consider two new parallelotopes,

$$\begin{aligned} L^+ &= \{u : |m_i \cdot u| \le 1/2 + \sqrt{s}||m_i||,\ i = 1, \cdots, s\}, \\ L^- &= \{u : |m_i \cdot u| \le 1/2 - \sqrt{s}||m_i||,\ i = 1, \cdots, s\}. \end{aligned}$$

It is easy to check that parallel faces of L and L^+ (and of L and L^- when $L^- \ne \phi$) are $\sqrt{s}$ apart. It is also easy to see that

$$E \subseteq L^+.$$

The estimate of $V_s(E)$ involves two possible cases.

Case 1: Suppose that L is narrow in the direction perpendicular to one of its bounding hyperplanes - the hyperplane $m_1 \cdot u = 1/2$ without loss of generality. So in this case we suppose

$$\frac{1}{2||m_1||} \le \sqrt{s}. \tag{1}$$

It then follows that any point of L is within $\sqrt{s}$ of the hyperplane $m_1 \cdot u = 1/2$ or the hyperplane $m_1 \cdot u = -1/2$, and we have

$$L \subseteq E \subseteq L^+.$$

Therefore

$$\begin{aligned} V_s(E) &\leq V_s(L^+) \\ &= \frac{2^s(1/2+||m_1||\sqrt{s})(1/2+||m_2||\sqrt{s})\cdots(1/2+||m_s||\sqrt{s})}{|\det M|} \end{aligned}$$

We estimate the first factor above, $1/2+||m_1||\sqrt{s}$, by using (1) and also $||m_1|| \leq \sqrt{s}\,|M|$, and the other factors by using

$$||m_i|| \leq \sqrt{s}|M| \leq \sqrt{s}\beta \qquad (i = 2, \cdots, s).$$

Therefore we obtain,

$$V_s(E) \leq \frac{2^s.2s|M|(1/2+s\beta)^{s-1}}{|\det M|}. \tag{2}$$

Case 2: Here we suppose that each bounding hyperplane of L, $m_i \cdot u = \pm 1/2$ $(i = 1, \cdots, s)$ is far from the origin. That is we suppose, for each $i = 1, \cdots, s$, that

$$\frac{1}{2||m_i||} > \sqrt{s}, \tag{3}$$

so that L^- has a positive volume and

$$E \subseteq L^+ \backslash L^-.$$

Therefore

$$V_s(E) \leq V_s(L^+) - V_s(L^-).$$

We write

$$||m|| = \max\{||m_1||, \cdots, ||m_s||\}$$

and, calculating the volumes $V_s(L^+)$ and $V_s(L^-)$, we see that

$$V_s(E) \leq \frac{2^s}{|\det M|}((1/2+||m||\sqrt{s})^s - (1/2-||m||\sqrt{s})^s).$$

The bracketed term above has the form $a^s - b^s = (a-b)(a^{s-1} + ba^{s-2} + \cdots + b^{s-1})$ where $a - b = 2\sqrt{s}\,||m||$ and by (3), $0 < b < a < 1$. Therefore we obtain

$$\begin{aligned} V_s(E) &\leq \frac{2^s}{|\det M|} 2\sqrt{s}\,||m||s \\ &\leq \frac{2^{s+1}s^2|M|}{|\det M|}, \end{aligned} \qquad (4)$$

since $||m|| \leq \sqrt{s}\,|M|$.

Finally, combining (2) and (4) we obtain, in either Case 1 or Case 2, the estimate

$$V_s(E) \leq \gamma. \frac{|M|}{|\det M|}$$

where

$$\gamma = \max\{2^{s+1}s^2, 2^{s+1}s(1/2 + s\beta)^{s-1}\}$$

which depends only on β and s.

This completes the proof.

■

We note that it is possible to express this result as

$$\#\{\Lambda \cap L\} = V_s(L) + O(V_{s-1}(\partial L)),$$

where V_{s-1} is $(s-1)$-dimensional volume. Examination of the proof of Case 2, above, also shows that we did not use the condition $|M| \leq \beta$ and so the theorem remains true if we replace this condition by the conditions

$$\frac{1}{2||m_i||} > \sqrt{s} \qquad (i = 1, \cdots, s),$$

and use $\gamma = 2^{s+1}s^2$. These conditions express the fact that each bounding hyperplane of L is further from its centre than the diameter of a unit cube.

2. Standard Representatives and the Constant K.

Let $d_1, \cdots, d_s$ be an integral basis of the algebraic integers $\mathcal{D}$. Suppose $n \geq 1$ is a rational integer and $<n>$ is the principal ideal of $\mathcal{D}$ generated by n. It is easy to see that the algebraic integers

$$b_i' = nd_i \qquad\qquad (i = 1, \cdots, s),$$

form an integral basis for $<n>$ and the n^s non-zero elements

$$\alpha_1 d_1 + \cdots + \alpha_s d_s \qquad (\alpha_i = 1, \cdots, n, \;\; i = 1, \cdots, s), \qquad (5)$$

are a system of representatives mod $<n>$. **We call these the standard representatives mod $<n>$ with respect to the basis $d_1, \cdots, d_s$ of $\mathcal{D}$. It is important to note that these representatives cannot be zero.** Unless otherwise stated the phrase "representatives mod$<n>$" will always refer to these standard ones.

Results similar to the above are true more generally. If $\mathcal{I}$ is an ideal of $\mathcal{D}$ there is a "triangular" integral basis of $\mathcal{I}$:

$$\begin{aligned} i_1' &= c_{11}d_1 \\ i_2' &= c_{21}d_1 + c_{22}d_2 \\ &\vdots \\ i_s' &= c_{s1}d_1 + \cdots + c_{ss}d_s \end{aligned}$$

where the c_{ij} are rational integers and $c_{ii} > 0$ $(i, j = 1, \cdots, s)$. The algebraic integers

$$\alpha_1 d_1 + \cdots + \alpha_s d_s \qquad (\alpha_i = 1, \cdots, c_{ii}, \;\; i = 1, \cdots, s),$$

are a system of representatives mod $\mathcal{I}$ and the number of these is $N(\mathcal{I}) = c_{11}c_{22}\cdots c_{ss}$. For a proof of these assertions about $\mathcal{I}$ see, for example, Weiss [1] Proposition 4-8-16 (Weiss uses $\alpha_i = 0, \cdots, c_{ii} - 1$ but this is irrelevant).

In our future work we will only be interested in standard representatives mod $<n>$ and not in the more general situation.

It is possible to limit the size of any ideal which contains a standard representative mod<n>.

Lemma 3.2

Let $d_1, \cdots, d_s$ be an integral basis of $\mathcal{D}$, $\mathcal{L}$ an ideal of $\mathcal{D}$ and $n \geq 1$ a rational integer. Suppose that $\mathcal{L}$ contains a standard representative mod <n> with respect to $d_1, \cdots, d_s$, that is, an element of the form

$$d = \alpha_1 d_1 + \cdots + \alpha_s d_s, \tag{6}$$

where $\alpha_i \in \{1, \cdots, n\}$ for $i = 1, \cdots, s$. Then

$$N(\mathcal{L}) \leq Kn^s$$

where

$$K = s^s J^s \geq 1 \tag{7}$$

and

$$J = \max\{|\sigma_i(d_j)| : i, j = 1, \cdots, s\}$$

is the maximum modulus of the conjugates of the basis elements.

Proof

From (6), since $d \neq 0$ and $d \in \mathcal{L}$ we have

$$\begin{aligned} N(\mathcal{L}) &\leq N(<d>) = |\prod_{i=1}^{s} \sigma_i(d)| \\ &\leq \prod_{i=1}^{s} \sum_{j=1}^{s} |\alpha_j||\sigma_i(d_j)| \\ &\leq (snJ)^s. \end{aligned}$$

We note that $J \geq 1$ (and in fact $J > 1$ unless each of $d_1, \cdots, d_s$ is a root of unity, see Narkiewicz [1] Theorem 2.1).

■

3. The Main Estimate.

In this section we assume the results of Stewart and Tall [1] about the embedding of algebraic numbers in Euclidean space (especially Chapters 6, 8 and Chapter 9 Section 2).

We write, as in Chapter 2,

$$s = r_1 + 2r_2$$

and list the conjugate functions in a fixed order as

$$\sigma_1, \cdots, \sigma_{r_1}, \sigma_{r_1+1}, \overline{\sigma}_{r_1+1}, \cdots, \sigma_{r_1+r_2}, \overline{\sigma}_{r_1+r_2}$$

where $\sigma_1, \cdots, \sigma_{r_1}$ are real and the rest complex (the bar denotes complex conjugation).

We may embed $\mathbb{K}$ in $\mathbb{R}^s$ by the following map : for $a \in \mathbb{K}$ let

$$\sigma(a) = (u_1, \cdots, u_{r_1}, u_{r_1+1}, v_{r_1+1}, \cdots, u_{r_1+r_2}, v_{r_1+r_2})$$

where

$$\begin{aligned} u_i &= \sigma_i(a) && \text{if } i = 1, \cdots, r_1, \\ &= \mathrm{Re}(\sigma_i(a)) && \text{if } i = r_1 + 1, \cdots, r_1 + r_2, \\ v_i &= \mathrm{Im}(\sigma_i(a)) && \text{if } i = r_1 + 1, \cdots, r_1 + r_2. \end{aligned}$$

The map σ is a ring homomorphism with the extra property that for $a \in \mathbb{K}$ and a rational number r,

$$\sigma(ra) = r\sigma(a).$$

An integral basis $d_1, \cdots, d_s$ for $\mathcal{D}$ becomes, under σ, a basis $\sigma(d_1), \cdots, \sigma(d_s)$ for $\mathbb{R}^s$ over $\mathbb{R}$. Also, under σ the algebraic integers $\mathcal{D}$ become the lattice $\sigma(\mathcal{D})$ generated by $\sigma(d_1), \cdots, \sigma(d_s)$. That is, the lattice,

$$\sigma(\mathcal{D}) = \{\alpha_1\sigma(d_1) + \cdots + \alpha_s\sigma(d_s) : \alpha_i \in \mathbb{Z}, i = 1, \cdots, s\}$$

whose fundamental domain,

$$H = \{t_1\sigma(d_1) + \cdots + t_s\sigma(d_s) : 0 < t_i \leq 1, i = 1, \cdots, s\} \tag{8}$$

has volume

$$V_s(H) = 2^{-r_2}|\delta|^{1/2}$$

where δ is the discriminant of $\mathbb{K}$. (Stewart and Tall [1] have $0 \leq t_i < 1$ above but this is irrelevant). The standard representatives mod<n> with respect to $d_1, \cdots, d_s$, for a rational integer $n \geq 1$, become the points in the lattice $\sigma(\mathcal{D})$ of the form

$$\alpha_1\sigma(d_1) + \cdots + \alpha_s\sigma(d_s) \qquad (\alpha_i = 1, \cdots, n,\ i = 1, \cdots, s)$$

that is the elements of $\sigma(\mathcal{D})$ in nH.

Furthermore, if $\mathcal{I}$ is an ideal of $\mathcal{D}$ with integral basis $b_1, \cdots, b_s$ then $\mathcal{I}$ maps to a sublattice of $\sigma(\mathcal{D})$ which is generated by $\sigma(b_1), \cdots, \sigma(b_s)$ and whose fundamental domain has volume

$$2^{-r_2}|\delta|^{1/2}N(\mathcal{I}). \tag{9}$$

We call this sublattice $\sigma(\mathcal{I})$.

We are now ready to state our main estimate.

Theorem 3.3

Let $d_1, \cdots, d_s$ be an integral basis of $\mathcal{D}$ and H the fundamental domain for the lattice $\sigma(\mathcal{D})$ in $\mathbb{R}^s$ defined by

$$H = \{t_1\sigma(d_1) + \cdots + t_s\sigma(d_s) : 0 < t_i \leq 1, i = 1, \cdots, s\}.$$

Let n be a positive integer and let $\mathcal{I}$ be an ideal of $\mathcal{D}$ such that the lattice $\sigma(\mathcal{I})$ in $\mathbb{R}^s$ intersects nH in at least one point. (So $\#\{\sigma(\mathcal{I}) \cap nH\}$, which denotes the number of such points, is also the number of elements of $\mathcal{I}$ which are standard representatives mod<n> with respect to $d_1, \cdots, d_s$ in the sense of definition (5) above.) Then

$$\#\{\sigma(\mathcal{I}) \cap nH\} = \frac{n^s}{N(\mathcal{I})} + O\left(\frac{n^{s-1}}{N(\mathcal{I})^{1-1/s}}\right),$$

where the O-constant depends on the basis $d_1, \cdots, d_s$ of $\mathcal{D}$ and the constant c from Lemma 2.2 but not on n or $\mathcal{I}$.

Proof

Choose an integral basis $b_1, \cdots, b_s$ for $\mathcal{I}$ such that

$$|\sigma_i(b_j)| \leq cN(\mathcal{I})^{1/s}, \qquad (i, j = 1, \cdots, s),$$

where c is the constant in Lemma 2.2 (and is therefore independent of $\mathcal{I}$). Define two $s \times s$ real matrices,

$$\begin{aligned} T &= [\sigma(b_1), \cdots, \sigma(b_s)], \\ A &= [\sigma(d_1), \cdots, \sigma(d_s)], \end{aligned}$$

where vectors are written as columns and $d_1, \cdots, d_s$ is the given integral basis for $\mathcal{D}$.

We note that T maps $\mathbb{Z}^s$, the integers of $\mathbb{R}^s$, onto the lattice $\sigma(\mathcal{I})$ and, by (9) above,

$$|\det T| = V_s(\text{fundamental domain of } \sigma(\mathcal{I})) = 2^{-r_2}|\delta|^{1/2}N(\mathcal{I}).$$

Also, as each element of T is $\sigma_i(b_j)$, or the real or imaginary part of such, for some i and j, we have

$$|T| \leq cN(\mathcal{I})^{1/s}.$$

Furthermore, A maps the unit cube $\mathcal{U} = (0,1] \times \cdots \times (0,1]$ of $\mathbb{R}^s$ onto H and so, as above

$$|\det A| = 2^{-r_2}|\delta|^{1/2}.$$

We wish to estimate

$$\#\{nH \cap \sigma(\mathcal{I})\} = \#\{T^{-1}(nA\mathcal{U}) \cap \mathbb{Z}^s\},$$

so we let L be the parallelotope $T^{-1}nA\mathcal{U}$ in $\mathbb{R}^s$. That is,

$$L = \{u : Mu \in \mathcal{U}\},$$

where $M = \frac{1}{n}A^{-1}T$. Now, using the special property of our basis $b_1, \cdots, b_s$,

$$\begin{aligned} |M| = \frac{1}{n}|A^{-1}T| &\leq \frac{s}{n}|A^{-1}||T| \\ &\leq \frac{s}{n}|A^{-1}|cN(\mathcal{I})^{1/s} \end{aligned} \tag{10}$$

and also

$$|\det M| = \frac{|\det T|}{n^s|\det A|} = \frac{N(\mathcal{I})}{n^s}. \tag{11}$$

From our hypothesis and Lemma 3.2 we have $N(\mathcal{I}) \leq Kn^s$ (with K as in Lemma 3.2.) Therefore, from (10),

$$|M| \leq s|A^{-1}|cK^{1/s}$$

and consequently we can use Theorem 3.1 with $\beta = s|A^{-1}|cK^{1/s}$ (we note that $\mathcal{U}$ is a translation of $(-1/2, 1/2] \times \cdots \times (-1/2, 1/2]$ and this theorem is independent of any translation). Therefore, using Theorem 3.1 and then (10) and (11) we obtain,

$$\begin{aligned} |\#\{nH \cap \sigma(\mathcal{I})\} - \frac{n^s}{N(\mathcal{I})}| &\leq \gamma \cdot \frac{|M|}{|\det M|} \\ &\leq \frac{\gamma\, s\, |A^{-1}|cN(\mathcal{I})^{1/s}}{nN(\mathcal{I})/n^s} \\ &= \gamma s|A^{-1}|c.\frac{n^{s-1}}{N(\mathcal{I})^{1-1/s}}. \end{aligned}$$

This proves the result with O-constant of (see Theorem 3.1)

$$s|A^{-1}|c. \max\{2^{s+1}s^2, 2^{s+1}s(1/2 + s^2|A^{-1}|cK^{1/s})^{s-1}\} \tag{12}$$

which is dependent only on s, c and the integral basis $d_1, \cdots, d_s$ of $\mathcal{D}$ and not on n or $\mathcal{I}$.

■

4. The Turán-Kubilius Inequality.

We are now ready to state the main result of this chapter.

Theorem 3.4 (Turán-Kubilius Inequality)

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$, and K be the constant from Lemma 3.2, namely

$$K = (s \max\{|\sigma_i(d_j)| : i, j = 1, \cdots, s\})^s.$$

Let f be a real-valued strongly additive function on the ideals of $\mathcal{D}$ (see (2.4) and (2.6)) and define for rational integers n,

$$\begin{aligned} A(f,n) &= \sum_{N(\mathcal{P}) \leq Kn^s} \frac{f(\mathcal{P})}{N(\mathcal{P})} \\ B(f,n) &= \sum_{N(\mathcal{P}) \leq Kn^s} \frac{(f(\mathcal{P}))^2}{N(\mathcal{P})} \end{aligned}$$

(as usual $\mathcal{P}$ denotes prime ideals). Then, there is a constant c' which depends on $d_1, \cdots, d_s$, on K and on the constant c from Lemma 2.2 (but not on n or f) such that , for $n \geq 3$,

$$\frac{1}{n^s} \sum_d (f(<d>) - A(f,n))^2 \leq c' B(f,n)$$

where the sum is over all $d \in \mathcal{D}$ which are standard representatives mod$<n>$ with respect to $d_1, \cdots, d_s$ (as defined by (5) in Section 2 above) or, equivalently, over all $d \in \mathcal{D}$ such that $\sigma(d) \in nH$ (where H is the fundamental domain for $\mathcal{D}$ in $\mathbb{R}^s$ defined by (8) in Section 3 above).

Proof

The proof proceeds as in Elliott [1] (Lemma 4.1). We use A and B as abbreviations for $A(f,n)$ and $B(f,n)$. We let, during this proof,

$$\sum_{\mathcal{P}}{}^{*}$$

denote the sum over prime ideals $\mathcal{P}$ with $N(\mathcal{P}) \leq Kn^s$ and also

$$\sum_{\mathcal{P} \neq \mathcal{Q}}{}^{*},$$

denote the double sum over all pairs of prime ideals $(\mathcal{P}, Q)$ with $\mathcal{P} \neq Q$ and $N(\mathcal{P})N(Q) \leq Kn^s$.

Firstly we assume f is a non-negative function.

We need to estimate the sum

$$S = \sum_d (f(<d>) - A)^2 = \sum_d (f(<d>))^2 - 2A\sum_d f(<d>) + n^s A^2. \qquad (13)$$

Since f is strongly additive and $d \neq 0$ in any sum we may write

$$f(<d>) = \sum_{\mathcal{P}|<d>} f(\mathcal{P}) = \sum_{d\in\mathcal{P}} f(\mathcal{P})$$

and, consequently, the first sum in (13), above, is

$$\begin{aligned} S_1 &= \sum_d (\sum_{d\in\mathcal{P}} f(\mathcal{P}))^2 \\ &= \sum_d (\sum_{d\in\mathcal{P}} (f(\mathcal{P}))^2 + \sum_{\mathcal{P}\neq Q, d\in\mathcal{P}Q} f(\mathcal{P})f(Q)). \qquad (14) \end{aligned}$$

We use Lemma 3.2 and interchange the order of summation to obtain, using the notation introduced above,

$$S_1 = \sum_{\mathcal{P}}{}^{*} ((f(\mathcal{P}))^2 \sum_{d\in\mathcal{P}} 1) + \sum_{\mathcal{P}\neq Q}{}^{*} (f(\mathcal{P})f(Q) \sum_{d\in\mathcal{P}Q} 1).$$

Now we use our main estimate, Theorem 3.3, with $\mathcal{I} = \mathcal{P}$ and then with $\mathcal{I} = \mathcal{P}Q$ to obtain, with O-constants independent of f and n,

$$\begin{aligned} S_1 \leq n^s \Bigg(B + \frac{1}{n} O \left(\sum_{\mathcal{P}}{}^{*} \frac{(f(\mathcal{P}))^2}{N(\mathcal{P})^{1-1/s}} \right) \\ + \sum_{\mathcal{P}\neq Q}{}^{*} \frac{f(\mathcal{P})f(Q)}{N(\mathcal{P})N(Q)} + \frac{1}{n} O \left(\sum_{\mathcal{P}\neq Q}{}^{*} \frac{f(\mathcal{P})f(Q)}{(N(\mathcal{P})N(Q))^{1-1/s}} \right) \Bigg). \qquad (15) \end{aligned}$$

We now examine the two O-terms in (15). Firstly,

$$\begin{aligned} \frac{1}{n} \sum_{\mathcal{P}}{}^{*} \frac{(f(\mathcal{P}))^2}{N(\mathcal{P})^{1-1/s}} &= \sum_{\mathcal{P}}{}^{*} \frac{(f(\mathcal{P}))^2}{N(\mathcal{P})} \frac{N(\mathcal{P})^{1/s}}{n} \\ &\leq K^{1/s} B \qquad (16) \end{aligned}$$

because for $\mathcal{P}$ in the sum, $N(\mathcal{P})^{1/s} \leq K^{1/s}n$. Secondly, using the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\frac{1}{n}\sum_{\mathcal{P}\neq Q}{}^{*} \frac{f(\mathcal{P})f(Q)}{(N(\mathcal{P})N(Q))^{1-1/s}} \\
&\leq \left(\sum_{\mathcal{P}\neq Q}{}^{*} \frac{(f(\mathcal{P}))^2(f(Q))^2}{N(\mathcal{P})N(Q)}\right)^{1/2} \frac{1}{n}\left(\sum_{\mathcal{P}\neq Q}{}^{*} \frac{1}{(N(\mathcal{P})N(Q))^{1-2/s}}\right)^{1/2} \\
&\leq (B^2)^{1/2}O(1), \qquad (17)
\end{aligned}$$

where in the last step we have used Lemma 2.3 iv) with $X = K^{1/s}n$. (Note that $K \geq 1$). The O-constant in (17) will depend only on K and s. We substitute (16) and (17) into (15) and estimate the remaining double sum in (15) by A^2 to obtain

$$S_1 \leq n^s(A^2 + O(B)). \qquad (18)$$

In a similar way we may estimate the second sum in (13) using Theorem 3.3. We obtain,

$$\begin{aligned}
S_2 &= A\sum_{d} f(<d>) = A\sum_{\mathcal{P}}{}^{*}(\, f(\mathcal{P})\sum_{d\in\mathcal{P}} 1\,) \\
&\geq n^s A\left(A - \frac{1}{n}O\left(\sum_{\mathcal{P}}{}^{*} \frac{f(\mathcal{P})}{N(\mathcal{P})^{1-1/s}}\right)\right).
\end{aligned}$$

Substituting this estimate and (18) into (13) we obtain,

$$\begin{aligned}
S &= S_1 - 2S_2 + n^sA^2 \\
&\leq n^s\left(O(B) + A^2 - 2A^2 + \frac{2A}{n}O\left(\sum_{\mathcal{P}}{}^{*} \frac{f(\mathcal{P})}{N(\mathcal{P})^{1-1/s}}\right) + A^2\right). \qquad (19)
\end{aligned}$$

We see that the A^2 terms cancel and we may estimate the second O-term in (19) by the Cauchy-Schwarz inequality in a way similar to the proof of (17). We obtain, for this term,

$$\begin{aligned}
&\frac{A}{n}\sum_{\mathcal{P}}{}^{*} \frac{f(\mathcal{P})}{N(\mathcal{P})^{1-1/s}} \\
&\leq \left(\sum_{\mathcal{P}}{}^{*} \frac{(f(\mathcal{P}))^2}{N(\mathcal{P})}\right) \frac{1}{n}\left(\sum_{\mathcal{P}}{}^{*} \frac{1}{N(\mathcal{P})}\right)^{1/2} \left(\sum_{\mathcal{P}}{}^{*} \frac{1}{N(\mathcal{P})^{1-2/s}}\right)^{1/2} \\
&= BO(1),
\end{aligned}$$

where in the last step we have used Lemma 2.3 iii) (with $X = K^{1/s}n$). Substituting this estimate into (19) we obtain, for an O-constant independent of n and f,

$$S \leq n^s O(B),$$

which proves the Turán-Kubilius inequality in the present case (where f is non-negative).

For a more general, real, f we proceed as in Elliott [1] and write

$$f(\mathcal{I}) = g(\mathcal{I}) - h(\mathcal{I})$$

where g and h are strongly additive, non-negative functions defined by

$$g(\mathcal{P}) = \begin{cases} f(\mathcal{P}) & \text{if} \quad f(\mathcal{P}) \geq 0 \\ 0 & \text{if} \quad f(\mathcal{P}) < 0 \end{cases}$$

and

$$h(\mathcal{P}) = \begin{cases} 0 & \text{if} \quad f(\mathcal{P}) \geq 0 \\ -f(\mathcal{P}) & \text{if} \quad f(\mathcal{P}) < 0. \end{cases}$$

We may use the Turán-Kubilius inequality, just proven, for g and h to obtain it for f since

$$A(f,n) = A(g,n) - A(h,n)$$

and

$$B(f,n) = B(g,n) + B(h,n).$$

This completes the proof.

■

We need the following Corollary in Chapter 5.

Corollary 3.5

Let $d_1, \cdots, d_s$, K, f, $A(f,n)$, $B(f,n)$ and H be as in Theorem 3.4. Let $n > m \geq 3$ be integers and $\lambda > 0$ a real number. Then

$$\frac{1}{n^s}\#\{d \in \mathcal{D} : \sigma(d) \in nH \text{ and } |\sum_{N(\mathcal{P})>Km^s, d\in\mathcal{P}} f(\mathcal{P})| > \lambda\}$$
$$\leq \frac{4}{\lambda^2}(A(f,n) - A(f,m))^2 + \frac{4c'}{\lambda^2}(B(f,n) - B(f,m)),$$

where c' is the constant from Theorem 3.4.

Instead of counting $\sigma(d) \in nH$ we could, equivalently, count the $d \in \mathcal{D}$ which are standard representatives mod$<n>$ with respect to $d_1, \cdots, d_s$.

Proof

From the triangle inequality, the left hand side of the desired inequality can be estimated as

$$\leq \frac{1}{n^s}\#\{d : \sigma(d) \in nH \text{ and } |h_m(<d>) - A(f,n) + A(f,m)| > \lambda/2\}$$
$$+\frac{1}{n^s}\#\{d : \sigma(d) \in nH \text{ and } |A(f,n) - A(f,m)| > \lambda/2\} \qquad (20)$$

where, for ideals $\mathcal{I}$,

$$h_m(\mathcal{I}) = \sum_{N(\mathcal{P})>Km^s, \mathcal{P}|\mathcal{I}} f(\mathcal{P}).$$

We note that h_m is strongly additive and since $n > m$,

$$\begin{aligned} A(h_m, n) &= \sum_{N(\mathcal{P})\leq Kn^s} \frac{h_m(\mathcal{P})}{N(\mathcal{P})} = \sum_{Km^s<N(\mathcal{P})\leq Kn^s} \frac{f(\mathcal{P})}{N(\mathcal{P})} \\ &= A(f,n) - A(f,m). \end{aligned}$$

Similarly

$$B(h_m, n) = B(f,n) - B(f,m).$$

To estimate the first term in (20), above, we use the Tchebycheff inequality (Lemma 2.4, i)) on the finite space of n^s elements $d \in \mathcal{D}$ with $\sigma(d) \in nH$ (or the n^s standard representatives mod$<n>$), and then use Theorem 3.4 on the function h_m. Thus, the first term in (20) is

$$\begin{aligned} &\leq \frac{1}{(\lambda/2)^2}\frac{1}{n^s}\sum_d (h_m(<d>) - A(f,n) + A(f,m))^2 \\ &\leq \frac{4}{\lambda^2}c'(B(f,n) - B(f,m)). \qquad (21) \end{aligned}$$

The second term in (20) is easier to estimate. Again we use the Tchebycheff inequality to estimate the second term of (20) as

$$\leq \frac{4}{\lambda^2}\frac{1}{n^s}\sum_d (A(f,n) - A(f,m))^2$$

$$= \frac{4}{\lambda^2}(A(f,n) - A(f,m))^2. \tag{22}$$

Substituting (21) and (22) into (20) gives the desired result.

■

It is possible to prove the Turán-Kubilius inequality for additive (and not just strongly additive) functions f. In this case $A(f,n)$ and $B(f,n)$ should be replaced by

$$C(f,n) = \sum_{N(\mathcal{P}^r) \leq Kn^s} \frac{f(\mathcal{P}^r)}{N(\mathcal{P}^r)}$$

and

$$D(f,n) = \sum_{N(\mathcal{P}^r) \leq Kn^s} \frac{(f(\mathcal{P}^r))^2}{N(\mathcal{P}^r)}.$$

The constant c' will not be the same as in Theorem 3.4 but will still only depend on $d_1, \cdots, d_s, K$ and c.

CHAPTER 4

THE POLYADIC INTEGERS OF AN ALGEBRAIC NUMBER FIELD

In this chapter we will discuss an extension of the ring of integers, $\mathcal{D}$, of a fixed algebraic number field $\mathbb{K}$, to form a metric probability space. In Sections 1 and 2 properties of the $\mathcal{P}$-adic completions of $\mathbb{K}$ are stated and used to construct a space, Ω, which is the Cartesian product of these. In Section 3, a special topology with the ideals of $\mathcal{D}$ as basic open sets around 0 is used to give another complete space $\overline{\mathcal{D}}$. Both spaces are completions of $\mathcal{D}$ and both have probability measures. In Section 4 the equivalence of $\overline{\mathcal{D}}$ and Ω from a topological and measure theoretic point of view is established. Section 5 deals with the independence of functions on Ω and shows how to extend an additive function to Ω.

Once the equivalence of the spaces Ω and $\overline{\mathcal{D}}$ has been established we shall call both the **space of polyadic integers of** $\mathbb{K}$, and use which ever formulation is most useful in any given circumstance.

We will assume the basic results of ideal theory (as in Stewart and Tall [1]) and take the necessary properties of $\mathcal{P}$-adic valuations from Goldstein [1] especially from Chapter 3, Sections 1 and 2 (Note that Goldstein develops $\mathcal{P}$-adic valuations from ideal theory). Taylor [1] and Halmos [1] supply the necessary properties of the Haar measure and measurable functions and Dugundji [1] basic topological ideas.

1. Basic Properties of the $\mathcal{P}$-adic Valuation.

Let $\mathcal{P}$ be a prime ideal of $\mathcal{D}$ and $d \in \mathcal{D}$ a non-zero algebraic integer of $\mathbb{K}$. We may write the principal ideal generated by d as a non-negative power of $\mathcal{P}$ times a finite product of positive powers of other prime ideals,

$$<d>= \mathcal{P}^{\alpha}\mathcal{P}_1^{\alpha_1}\mathcal{P}_2^{\alpha_2}\cdots\mathcal{P}_l^{\alpha_l},$$

and then define the $\mathcal{P}$-adic value of d as,

$$|d|_{\mathcal{P}} = 1/N(\mathcal{P})^{\alpha}.$$

This defines a non-Archimedean valuation on $\mathcal{D}$ (once we set $|0|_{\mathcal{P}} = 0$) which easily extends to $\mathbb{K}$.

In the topology induced by the metric $d(a,b) = |a-b|_{\mathcal{P}}$, $\mathbb{K}$ is Hausdorff and the field operations are continuous - that is, $\mathbb{K}$ is a metric field (and consequently $\mathcal{D}$ a metric ring).

We may complete $\mathbb{K}$ with respect to the $\mathcal{P}$-adic valuation to obtain the complete field $\mathbb{K}_{\mathcal{P}}$.

Define

$$\begin{aligned} \mathcal{D}_{\mathcal{P}} &= \{x \in \mathbb{K}_{\mathcal{P}} : |x|_{\mathcal{P}} \leq 1\}, \\ \overline{\mathcal{P}} &= \{x \in \mathbb{K}_{\mathcal{P}} : |x|_{\mathcal{P}} < 1\} \end{aligned}$$

and let $\pi \in \mathcal{D}_{\mathcal{P}}$ be such that $|\pi|_{\mathcal{P}} = 1/N(\mathcal{P})$ (any element of $\mathcal{P}\backslash\mathcal{P}^2$ will do).

The following results are proven in Goldstein [1] (see Theorems 9,12,13,14 and 21 of Chapter 3 Section 1):

a) Every ideal of $\mathcal{D}_{\mathcal{P}}$ is of the form $\overline{\mathcal{P}}^{\alpha}$ ($\alpha \geq 0$ a rational integer). Furthermore, $\overline{\mathcal{P}}^{\alpha} = \pi^{\alpha}\mathcal{D}_{\mathcal{P}}$ so that $\mathcal{D}_{\mathcal{P}}$ is a principal ideal domain.

b) $\mathcal{D}_{\mathcal{P}}$ is the closure of $\mathcal{D}$ and $\overline{\mathcal{P}}$ of $\mathcal{P}$ in the $\mathcal{P}$-adic topology.

c) $\mathcal{D}_{\mathcal{P}}$ is a compact open subring of $\mathbb{K}_{\mathcal{P}}$ and all its ideals $\overline{\mathcal{P}}^{\alpha}$ are compact and open.

d) The factor rings $\mathcal{D}_{\mathcal{P}}/\overline{\mathcal{P}}^{\alpha}$ and $\mathcal{D}/\mathcal{P}^{\alpha}$ are isomorphic and so $\mathcal{D}_{\mathcal{P}}/\overline{\mathcal{P}}^{\alpha}$ is finite (with $N(\mathcal{P})^{\alpha}$ elements).

e) Any element $x \in \mathcal{D}_{\mathcal{P}}$ can be expressed uniquely in series form as

$$x = \sum_{j=l}^{\infty} b_j \pi^j, \qquad b_j \in A,$$

where $l \geq 0$ and A is a system of representatives mod $\mathcal{P}$ (that is of the cosets of $\mathcal{D}/\mathcal{P}$ in $\mathcal{D}$). Furthermore, if $b_l \neq 0$, $|x|_{\mathcal{P}} = 1/N(\mathcal{P})^l$.

f) In $\mathcal{D}_{\mathcal{P}}$, the $\mathcal{P}$-adic valuation takes only the values $1/N(\mathcal{P})^{\alpha}(\alpha \geq 0)$ that it takes in $\mathcal{D}$.

g) The set $\{x \in \mathbb{K}_{\mathcal{P}} : |x|_{\mathcal{P}} \leq 1/N(\mathcal{P})^{\alpha}\} = \overline{\mathcal{P}}^{\alpha}$ is open, closed and compact (this is contained in b) and c) above).

The set $\mathcal{D}_{\mathcal{P}}$ is called the set of **$\mathcal{P}$-adic integers of $\mathbb{K}$** and, since it is a compact metric group (under addition), it has a unique, translation - invariant, complete, normalized measure (the Haar measure) on a σ-field containing the Borel sets of $\mathcal{D}_{\mathcal{P}}$ (that is, containing the σ-field generated by the open sets of the topology). **This measure we will call $M_{\mathcal{P}}$** . Note that by the normalization, $M_{\mathcal{P}}(\mathcal{D}_{\mathcal{P}}) = 1$.

The next lemma gives the measure of a typical ball in the topology.

Lemma 4.1

Let $\alpha \geq 0$ be a rational integer and $a \in \mathcal{D}_{\mathcal{P}}$, then

$$M_{\mathcal{P}}(x \in \mathcal{D}_{\mathcal{P}} : |x - a|_{\mathcal{P}} \leq 1/N(\mathcal{P})^{\alpha}) = 1/N(\mathcal{P})^{\alpha}.$$

Proof

Firstly note that we may as well assume $a = 0$ by the translation invariance of the Haar measure, $M_{\mathcal{P}}$. The set in question is then $\overline{\mathcal{P}}^{\alpha}$ and we seek $M_{\mathcal{P}}(\overline{\mathcal{P}}^{\alpha})$. By result d) mentioned above, we may write $\mathcal{D}_{\mathcal{P}}$ as the disjoint union of cosets of $\overline{\mathcal{P}}^{\alpha}$,

$$\mathcal{D}_{\mathcal{P}} = \overline{\mathcal{P}}^{\alpha} \cup (a_2 + \overline{\mathcal{P}}^{\alpha}) \cup \cdots \cup (a_t + \overline{\mathcal{P}}^{\alpha}),$$

where $t = N(\mathcal{P})^{\alpha}$.

The Haar measure, $M_{\mathcal{P}}$, is translation invariant and so each of the cosets in the above union has the same measure which therefore must be $1/t$ because $\mathcal{D}_{\mathcal{P}}$ has total measure 1.

■

2. The Polyadic Integers of $\mathbb{K}$ - First Version.

We take the Cartesian product of the countably many $\mathcal{P}$-adic spaces, $\mathcal{D}_{\mathcal{P}}$, to define the **space of polyadic integers of $\mathbb{K}$** ,

$$\Omega = \prod_{\mathcal{P}} \mathcal{D}_{\mathcal{P}}.$$

In this definition a fixed order of the prime ideals: $\mathcal{P}_1, \mathcal{P}_2, \cdots$ is assumed and we write $\underset{\sim}{x} \in \Omega$ as $\underset{\sim}{x} = (x_{\mathcal{P}_i})$.

This space, Ω, we endow with the product topology and note that, from Tychonoff's Theorem, it is compact because each $\mathcal{D}_{\mathcal{P}}$ is compact. Furthermore, it is a metric ring with the metric inducing the product topology.

One such metric may be briefly described as follows. Let

$$d_{\mathcal{P}_i}(x_{\mathcal{P}_i}, y_{\mathcal{P}_i}) = \min\{1/i, |x_{\mathcal{P}_i} - y_{\mathcal{P}_i}|_{\mathcal{P}_i}\}.$$

This metric induces the $\mathcal{P}_i$-adic topology on $\mathcal{D}_{\mathcal{P}_i}$. Now set

$$D(\underset{\sim}{x}, \underset{\sim}{y}) = \sup_i\{d_{\mathcal{P}_i}(x_{\mathcal{P}_i}, y_{\mathcal{P}_i})\}.$$

D is a metric inducing the product topology on Ω. (See Dugundji [1] Chapter IX, Corollaries 3.3 and 7.3 for details).

We can now say that Ω has a Haar measure and that this is the same as the product measure inherited from the $\mathcal{D}_{\mathcal{P}}$ (the Borel σ-field on Ω being the product σ-field). **We call this measure M .**

As in Lemma 4.1 we find the measure of a typical ball in the topology. Firstly we need a definition.

Let $\mathcal{I}$ be an ideal of $\mathcal{D}$ and $\mathcal{P}_1, \mathcal{P}_2, \cdots$ the list of prime ideals above. For certain non-negative integers $\alpha_1, \cdots, \alpha_l$ (some of which may be zero) we have

$$\mathcal{I} = \mathcal{P}_1^{\alpha_1} \cdots \mathcal{P}_l^{\alpha_l}.$$

For $\underset{\sim}{x} = (x_{\mathcal{P}_i}) \in \Omega$ we say,

$$\mathcal{I} | \underset{\sim}{x} \quad \text{if} \quad |x_{\mathcal{P}_i}|_{\mathcal{P}_i} \leq 1/N(\mathcal{P}_i)^{\alpha_i} \quad \text{for each} \quad i = 1, \cdots, l. \tag{1}$$

Lemma 4.2

For any ideal $\mathcal{I}$ of $\mathcal{D}$,

$$M(\underset{\sim}{x} \in \Omega : \mathcal{I} | \underset{\sim}{x}) = 1/N(\mathcal{I}).$$

Proof

Write $\mathcal{I} = \mathcal{P}_1^{\alpha_1} \cdots \mathcal{P}_l^{\alpha_l}$ as above. The desired measure can be calculated as a product measure using Lemma 4.1 since

$$\{\underset{\sim}{x} \in \Omega : \mathcal{I}|\underset{\sim}{x}\} = \overline{\mathcal{P}_1}^{\alpha_1} \times \overline{\mathcal{P}_2}^{\alpha_2} \times \cdots \times \overline{\mathcal{P}_l}^{\alpha_l} \times \mathcal{D}_{\mathcal{P}_{l+1}} \times \mathcal{D}_{\mathcal{P}_{l+2}} \times \cdots .$$

Therefore,

$$\begin{aligned} M(\underset{\sim}{x} \in \Omega : \mathcal{I}|\underset{\sim}{x}) &= \prod_{i=1}^{l} M_{\mathcal{P}_i}(\overline{\mathcal{P}_i}^{\alpha_i}) \\ = \prod_{i=1}^{l} 1/N(\mathcal{P}_i)^{\alpha_i} &= 1/N(\mathcal{I}). \end{aligned}$$

■

Many other measure theoretic and topological properties of Ω can also be proven by this product technique, but at times it is easier to consider Ω as the completion of $\mathcal{D}$ with respect to a certain metric. This is the aim of the next two sections.

The space, Ω, defined above is closely connected with the ring of adeles of $\mathbb{K}$. (See Goldstein [1], Chapter 3 Section 2). The formulation is topologically much simpler, however, as we are taking the product of compact spaces and so the resulting space is also compact. The completions of $\mathbb{K}$ for Archimedean valuations (which are basically those defined as the absolute value of conjugates of algebraic numbers) do not appear in our product space.

3. The Polyadic Integers of $\mathbb{K}$ - Second Version.

The construction which we will now give is an adaptation of that of Novoselov [1] (as amplified in Babu [1]) to the integers of $\mathbb{K}$ (in place of the rational integers). More topological details will be included as the space to be constructed here is not as familiar as that considered in Section 2.

The collection of sets

$$\{a + \mathcal{I} : a \in \mathcal{D} \text{ and } \mathcal{I} \text{ an ideal of } \mathcal{D}\},$$

may serve as a basis for a topology on $\mathcal{D}$ (that is, define open sets as the unions of such sets) because any element $d \in \mathcal{D}$ is in such a set ($d \in d + \mathcal{D}$) and the intersection of two such sets, if it is not empty, contains a third (if $d \in (a + \mathcal{I}) \cap (b + \mathcal{I}')$ we check that $d + \mathcal{I}\mathcal{I}' \subseteq (a + \mathcal{I}) \cap (b + \mathcal{I}')$).

Let the topology generated by this basis be $\mathcal{N}$. We now have a topological space $(\mathcal{D}, \mathcal{N})$ which is second countable (that is, has a countable basis) because the set of algebraic integers and the set of ideals are both countable. It is also possible to show that this space is regular and hence, by Urysohn's Theorem, metrizable, but it is useful to exhibit the metric directly (compare with Babu [1], Chapter 1 Section 2).

We list the non-zero ideals of $\mathcal{D}$ in some arbitrary, but fixed, order: $\mathcal{D}, \mathcal{I}_2, \mathcal{I}_3 \cdots$. For two algebraic integers $a, b \in \mathcal{D}$ define

$$d(a,b) = \sum_{n=2}^{\infty} \frac{\Psi(a-b, \mathcal{I}_n)}{2^n}$$

where

$$\Psi(c, \mathcal{I}) = \begin{cases} 0 & \text{if} \quad c \in \mathcal{I} \\ 1 & \text{if} \quad c \notin \mathcal{I} \,. \end{cases}$$

Theorem 4.3

The function $d(a, b)$ is a translation invariant metric which generates the topology $\mathcal{N}$. With this metric, $\mathcal{D}$ is a metric ring.

Proof

Clearly $d(a, b)$ is translation invariant as

$$d(a+c, b+c) = \sum_{n=2}^{\infty} \frac{\Psi(a+c-b-c, \mathcal{I}_n)}{2^n} = d(a,b).$$

Only two of the defining properties of a metric are non-trivial to check.

Suppose that $d(a, b) = 0$, then $\Psi(a-b, \mathcal{I}_n) = 0$ for all $\mathcal{I}_n$ and so $a-b \in \mathcal{I}_n$ for all $\mathcal{I}_n$. Hence $a - b = 0$.

Now, note that if $\Psi(a-c, \mathcal{I}_n) = 1$ then one of $\Psi(a-b, \mathcal{I}_n)$ or $\Psi(c-b, \mathcal{I}_n)$ equals 1 (for if both equal zero then $a - b \in \mathcal{I}_n$ and $c - b \in \mathcal{I}_n$ which implies $a - c \in \mathcal{I}_n$). Hence,

$$\Psi(a-c,\mathcal{I}_n) \leq \Psi(a-b,\mathcal{I}_n) + \Psi(c-b,\mathcal{I}_n),$$

which gives the triangle inequality for $d(a,b)$.

Now we must show that the open balls of $\mathcal{D}$,

$$B(a,\epsilon) = \{x \in \mathcal{D} : d(a,x) < \epsilon\}$$

give the same topology as the basis $\{a+\mathcal{I}\}$. To do this we must check that each topology is finer than the other.

Firstly, let $c \in B(a,\epsilon)$ so that $d(c,a) < \epsilon$. Choose an integer l such that

$$1/2^l + d(c,a) < \epsilon,$$

and let

$$\mathcal{I} = \mathcal{I}_2 \cdots \mathcal{I}_{l+1}.$$

Then, for $x \in c+\mathcal{I}$,

$$d(x,c) = \sum_{n=l+2}^{\infty} \frac{\Psi(x-c,\mathcal{I}_n)}{2^n} < 1/2^l,$$

and so

$$\begin{aligned} d(x,a) &\leq d(x,c) + d(c,a) \\ &< 1/2^l + d(c,a) < \epsilon. \end{aligned}$$

Therefore,

$$c \in c+\mathcal{I} \subseteq B(a,\epsilon)$$

and so the topology generated by $\{a+\mathcal{I}\}$ is finer than that generated by the open balls.

The converse conclusion is easier to prove and so the topologies are the same.

To show that $\mathcal{D}$ is a metric ring we need to show that if $a_n \to a$ and $b_n \to b$ (as $n \to \infty$) with respect to the metric d then $a_n b_n \to ab$ and $a_n \pm b_n \to a \pm b$ (as $n \to \infty$). This follows easily from the following lemma which, in view of the equality of the topologies just discussed, is a trivial restatement of the fact that ideals are open sets containing 0.

Lemma 4.4

Let $\{a_n\}$ be a sequence of elements in $\mathcal{D}$.

i) If $a_n \to 0$ as $n \to \infty$ then for any ideal $\mathcal{L}$ there is an integer N_0, such that $a_n \in \mathcal{L}$ when $n \geq N_0$.

ii)Conversely: If for every ideal $\mathcal{L}$ there is an integer N_0 such that $a_n \in \mathcal{L}$ when $n \geq N_0$, then $a_n \to 0$ as $n \to \infty$.

This completes the proof of Theorem 4.3.

■

We now have a metric space $(\mathcal{D}, \mathcal{N})$ which we can complete. We will show $\mathcal{D}$ is totally bounded. It then follows that the completion, $\overline{\mathcal{D}}$ is totally bounded and so, compact.

Lemma 4.5

$\mathcal{D}$ is totally bounded. That is, for every $\epsilon > 0$, there is a finite covering of $\mathcal{D}$ by balls of radius ϵ.

Proof

Let l be a positive integer such that $1/2^l < \epsilon$ and put $\mathcal{I} = \mathcal{I}_2 \cdots \mathcal{I}_{l+1}$. We may write, for some elements $\{a_1, \cdots, a_t\}$ of $\mathcal{D}$,

$$\mathcal{D} = (a_1 + \mathcal{I}) \cup (a_2 + \mathcal{I}) \cup \cdots \cup (a_t + \mathcal{I}).$$

If $b \in \mathcal{D}$ then $b \in a + \mathcal{I}$ for some $a \in \{a_1, \cdots, a_t\}$, so that $b \in a + \mathcal{I}_i$ for each $i = 2, \cdots, l+1$. Therefore $d(b, a) < 1/2^l < \epsilon$ and so $b \in B(a, \epsilon)$.

Therefore

$$\mathcal{D} = B(a_1, \epsilon) \cup \cdots \cup B(a_t, \epsilon).$$

■

We can now say that $\overline{\mathcal{D}}$ has a unique, normalized, complete translation invariant Haar-measure on a σ-field containing its Borel sets. **This measure we call P .**

Throughout the rest of this section we assume we have a fixed ideal, $\mathcal{I}$, of $\mathcal{D}$ and that

$$\{a_1, \cdots, a_t\}$$

is a system of representatives mod $\mathcal{I}$ (where $t = N(\mathcal{I})$), so that we may write $\mathcal{D}$ as the disjoint union of cosets,

$$\mathcal{D} = (a_1 + \mathcal{I}) \cup \cdots \cup (a_t + \mathcal{I}).$$

We now proceed to find the measure of a typical basic open set, $a + \mathcal{I}$.

Lemma 4.6

If $x \in \overline{\mathcal{D}}$, there is a unique $a \in \{a_1, \cdots, a_t\}$ such that

$$x \in a + \overline{\mathcal{I}},$$

where $\overline{\mathcal{I}}$, the completion of $\mathcal{I}$, is an ideal of $\overline{\mathcal{D}}$. **We call a the unique representative of x mod $\mathcal{I}$ from $\{a_1, \cdots, a_t\}$.**

The result may be expressed by saying $\mathcal{D}/\mathcal{I}$ is isomorphic to $\overline{\mathcal{D}}/\overline{\mathcal{I}}$ (compare result d) of Section 1).

Proof

It is easy, using limits, to check that $\overline{\mathcal{I}}$ is an ideal of $\overline{\mathcal{D}}$.

Suppose that $\{x_n\}$ is a sequence of elements of $\mathcal{D}$ with $x_n \to x$ as $n \to \infty$. Since

$$\{x_n : n \geq 1\} \subseteq (a_1 + \mathcal{I}) \cup \cdots \cup (a_t + \mathcal{I}),$$

there is an $a \in \{a_1, \cdots, a_t\}$ and a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that for all n_k, $x_{n_k} \in a + \mathcal{I}$. Let us say

$$x_{n_k} = a + i_{n_k}, \qquad i_{n_k} \in \mathcal{I}.$$

The space $\overline{\mathcal{D}}$ is compact and so the sequence $\{i_{n_k}\}$ has a convergent subsequence which converges to, say, $i \in \overline{\mathcal{I}}$. We can then say that $\{x_{n_k}\}$ has a subsequence converging to $a + i$. This subsequence should also, of course, converge to x. Therefore $x = a + i$ for some $i \in \overline{\mathcal{I}}$.

As for uniqueness: Suppose a and a' are representatives of x mod $\mathcal{I}$. It then follows that $a - a' \in \overline{\mathcal{I}}$ and so there is a sequence $\{i_n\}$ of elements in $\mathcal{I}$ such that $i_n - (a - a') \to 0$ as $n \to \infty$. According to Lemma 4.4, there is an N_0 such that $i_n - (a - a') \in \mathcal{I}$ when $n \geq N_0$. We know that $i_n \in \mathcal{I}$ and so $a - a' \in \mathcal{I}$ which means that $a = a'$, since $a, a' \in \{a_1, \cdots, a_t\}$.

This completes the proof.

■

Lemma 4.7

For any $a \in \overline{\mathcal{D}}$,

$$P(x \in a + \overline{\mathcal{I}}) = 1/N(\mathcal{I}).$$

Proof

This result should be compared with Lemmas 4.1 and 4.2.

Lemma 4.6 allows us to write $\overline{\mathcal{D}}$ as the disjoint union of cosets of $\overline{\mathcal{I}}$,

$$\overline{\mathcal{D}} = (a_1 + \overline{\mathcal{I}}) \cup (a_2 + \overline{\mathcal{I}}) \cup \cdots \cup (a_t + \overline{\mathcal{I}}),$$

where $t = N(\mathcal{I})$. As in the Proof of Lemma 4.1 we note that each of these cosets has the same measure, which must therefore be $1/t = 1/N(\mathcal{I})$.

■

In later chapters we use properties of $\overline{\mathcal{D}}$ (and Ω) to establish results about the frequency of certain sets of algebraic integers. The following Lemma shows the connection between these concepts.

Lemma 4.8

Let g be a real valued function on $\overline{\mathcal{D}}$ which is **periodic mod $\mathcal{I}$** (that is if $x - y \in \overline{\mathcal{I}}$ then $g(x) = g(y)$) and let A be a set of real numbers. For any system of representatives $\{a_1, \cdots, a_t\}$ mod $\mathcal{I}$, where $t = N(\mathcal{I})$, we have,

$$\begin{aligned} P(g(x) \in A) &= \frac{1}{N(\mathcal{I})} \#\{a_j : g(a_j) \in A, j = 1, \cdots, t\}, \\ \int g(x) dP &= \frac{1}{N(\mathcal{I})} (g(a_1) + \cdots + g(a_t)). \end{aligned}$$

These expressions are independent of the particular choice of representatives mod $\mathcal{I}$.

Proof

As in Lemma 4.7 we may write $\overline{\mathcal{D}}$ as the disjoint union of cosets and we notice that on each coset $a_j + \overline{\mathcal{I}}$, g takes the constant value $g(a_j)$. This means that g is measurable and the two desired equations easily follow.

■

We have seen that the two spaces Ω and $\overline{\mathcal{D}}$ have properties in common (compare Lemma 4.7 with Lemmas 4.1 and 4.2) and, in fact, they are equivalent spaces. That is to say, there is a map $\phi : \overline{\mathcal{D}} \to \Omega$ such that both ϕ and ϕ^{-1} are ring homeomorphisms which preserve measure. The maps ϕ and ϕ^{-1} preserve ring, topological and measure properties. The proof of this equivalence is the object of the next section.

4. The Equivalence of Ω and $\overline{\mathcal{D}}$.

We need a few topological and measure theoretical notions.

Let $\mathcal{A}_i (i = 1, 2)$ be two topological rings with measure (that is, $\mathcal{A}_i$ has a basis for a topology T_i in which the ring operations are continuous, and a measure M_i with σ-field F_i).

Let $\phi : \mathcal{A}_1 \to \mathcal{A}_2$ be a ring homomorphism. We say:

ϕ is *measure preserving* if for any $B \in F_2$, we have $\phi^{-1}(B) \in F_1$ and $M_2(B) = M_1(\phi^{-1}(B))$,

ϕ is a *homeomorphism* if ϕ is invertible and both ϕ and ϕ^{-1} are continuous (it suffices to check this at 0),

ϕ is *uniformly continuous* if for any basic open set $U \subseteq \mathcal{A}_2$ containing 0, there is a basic open set $V \subseteq \mathcal{A}_1$ containing 0 so that, when $a - b \in V$ we have $\phi(a) - \phi(b) \in U$.

This last definition is taken from Husain [1] (see Definition 3 of Section 22).

We also need some notation. Let

$$\Omega = \prod_{\mathcal{P}} \mathcal{D}_{\mathcal{P}},$$

$\Delta = \{(d, d, \cdots) : d \in \mathcal{D}\}$, the diagonal of $\prod_{\mathcal{P}} \mathcal{D}$,

$\mathcal{T}_{\mathcal{P}}$ = the $\mathcal{P}$-adic topology restricted to $\mathcal{D}$,

$\mathcal{T}$ = the product of the $\mathcal{P}$-adic topologies, $\mathcal{T}_{\mathcal{P}}$, on $\prod_{\mathcal{P}} \mathcal{D}$, restricted to Δ,

$\mathcal{N}$ = the topology on $\mathcal{D}$ generated by the basis $\{a + \mathcal{I}\}$ as described in Section 3,

$\phi : \mathcal{D} \to \Delta$ be the map $\phi(d) = (d, d, \cdots)$.

Lemma 4.9 (Strong Chinese Remainder Theorem)
Δ is dense in Ω.

Proof

Let $\underset{\sim}{z} = (z_{\mathcal{P}_i})$ be an element of Ω. We need only show that any basic open set around $\underset{\sim}{z}$ contains an element of Δ. Such a basic open set may be described as

$$U = U_1 \times U_2 \times \cdots \times U_l \times \mathcal{D}_{\mathcal{P}_{l+1}} \times \mathcal{D}_{\mathcal{P}_{l+2}} \times \cdots,$$

where, for certain non-negative integers $\alpha_1, \cdots, \alpha_l$ (some of which may be zero), we have for each $i = 1, \cdots, l$

$$U_i = \{x \in \mathcal{D}_{\mathcal{P}_i} : |x - z_{\mathcal{P}_i}|_{\mathcal{P}_i} \le 1/N(\mathcal{P}_i)^{\alpha_i}\}.$$

We know that $\mathcal{D}$ is dense in any $\mathcal{D}_{\mathcal{P}}$ and therefore for each $i = 1, \cdots, l$ there is an algebraic integer $d_i \in \mathcal{D}$ such that,

$$|d_i - z_{\mathcal{P}_i}|_{\mathcal{P}_i} \le 1/N(\mathcal{P}_i)^{\alpha_i}.$$

From the Chinese Remainder Theorem (see Lemma 2.1) there is a single algebraic integer $d \in \mathcal{D}$ such that, for each $i = 1, \cdots, l$

$$|d - d_i|_{\mathcal{P}_i} \le 1/N(\mathcal{P}_i)^{\alpha_i}.$$

Therefore, remembering that the $\mathcal{P}$-adic valuations are non-Archimedean, for each $i = 1, \cdots, l$ we have,

$$\begin{aligned}|d - z_{\mathcal{P}_i}|_{\mathcal{P}_i} &\leq \max\{|d - d_i|_{\mathcal{P}_i}, |d_i - z_{\mathcal{P}_i}|_{\mathcal{P}_i}\} \\ &\leq 1/N(\mathcal{P}_i)^{\alpha_i}.\end{aligned}$$

We now have $d \in U_i$ $(i = 1, \cdots, l)$ and hence $(d, d, \cdots) \in U$.

■

It is clear that both ϕ and ϕ^{-1} are ring isomorphisms but topological properties are also preserved.

Lemma 4.10

ϕ is a uniformly continuous map from $(\mathcal{D}, \mathcal{N})$ onto $(\Delta, \mathcal{T})$,
ϕ^{-1} is a uniformly continuous map from $(\Delta, \mathcal{T})$ onto $(\mathcal{D}, \mathcal{N})$.
Consequently, ϕ is a homeomorphism.

Proof

Let U be a basic open set around 0 in $\mathcal{T}$. For some non-negative integers $\alpha_1, \cdots, \alpha_l$ we may write,

$$U = \{(d, d, \cdots) : |d|_{\mathcal{P}_i} \leq 1/N(\mathcal{P}_i)^{\alpha_i}, i = 1, \cdots, l\}.$$

Now set

$$\mathcal{I} = \mathcal{P}_1^{\alpha_1}\mathcal{P}_2^{\alpha_2} \cdots \mathcal{P}_l^{\alpha_l}.$$

Then

$$\begin{aligned}\phi^{-1}(U) &= \{d \in \mathcal{D} : |d|_{\mathcal{P}_i} \leq 1/N(\mathcal{P}_i)^{\alpha_i}, i = 1, \cdots, l\} \\ &= \{d \in \mathcal{D} : d \in \mathcal{P}_i^{\alpha_i}, i = 1, \cdots, l\} \\ &= \mathcal{I}\end{aligned}$$

which is a basic open set around 0 in $\mathcal{N}$.

This proves the continuity of ϕ. In fact as ϕ is invertible we also have $U = \phi(\mathcal{I})$ and this is enough to show the uniform continuity of ϕ, for if $a - b \in \mathcal{I}$ then $\phi(a) - \phi(b) = \phi(a - b) \in U$.

The proof that ϕ^{-1} is uniformly continuous is similar.

■

Now, we complete $\mathcal{D}$ with respect to $\mathcal{N}$ to get $\overline{\mathcal{D}}$ and Δ with respect to $\mathcal{T}$ to obtain Ω (by Lemma 4.9):

Lemma 4.11

ϕ extends to a ring isomorphism from $\overline{\mathcal{D}}$ to Ω which again is a homeomorphism.

Proof

For convenience, we remember that the topologies of $\mathcal{D}$ and Δ are generated by metrics and so from Dugundji [1] (Chapter XIV Theorem 5.3), ϕ can be extended to a homeomorphism from $\overline{\mathcal{D}}$ to Ω.

The ring isomorphism properties of ϕ come from its continuity. For example, if $x, y \in \overline{\mathcal{D}}$ we have $x_n \longrightarrow x$ and $y_n \longrightarrow y$ as $n \longrightarrow \infty$ for some sequences $\{x_n\}$ and $\{y_n\}$ in $\mathcal{D}$. Thus,

$$\begin{aligned}\phi(x+y) &= \phi(\lim_{n\to\infty}(x_n+y_n)) = \lim_{n\to\infty}\phi(x_n+y_n)\\ &= \lim_{n\to\infty}\phi(x_n) + \lim_{n\to\infty}\phi(y_n) = \phi(x)+\phi(y).\end{aligned}$$

In fact the extension is defined in terms of such limits. Note that for this extension procedure to work, ϕ and ϕ^{-1} have to be uniformly continuous. Dugundji [1] has an example where the extension of a homeomorphism is not a homeomorphism.

■

The last Lemma tells us that we can regard $\overline{\mathcal{D}}$ and Ω as topologically the same and the next tells us that they are the same in measure.

Lemma 4.12

The ring homeomorphisms ϕ and ϕ^{-1}, from Lemma 4.11, are measure preserving on Borel sets.

Proof

We shall prove the result for ϕ as the proof for ϕ^{-1} is the same.

Let B be a Borel set of Ω. It then follows, because ϕ is continuous, that $\phi^{-1}(B)$ is a Borel set in $\overline{\mathcal{D}}$ and so ϕ is measurable.

Consider the function defined on Borel sets, B, of Ω by

$$L(B) = P(\phi^{-1}(B))$$

where P is the Haar measure on $\overline{\mathcal{D}}$. It is easy to check that this is a translation invariant, normalized measure (that is $L(\Omega) = 1$) because ϕ^{-1} is an isomorphism and P has these properties. By uniqueness, then, L must be the Haar measure, M, on the Borel sets of Ω. Thus, for Borel sets, B, of Ω,

$$M(B) = P(\phi^{-1}(B)).$$

■

Thus we have shown that we may regard Ω and $\overline{\mathcal{D}}$ as the same space from a measure and topological point of view and we call both the **space of polyadic integers of $\mathbb{K}$. From now on, we will identify these two spaces when convenient, and use whichever formulation is most suitable in any given circumstance. We will use P for the measure on Ω.**

It is also worth noting that the two concepts $\mathcal{I}|\underset{\sim}{x}$ (as defined by (1) in Section 2, above) and $x \in \overline{\mathcal{I}}$ (where $\overline{\mathcal{I}}$ is the closure of $\mathcal{I}$ in $\overline{\mathcal{D}}$) coincide because they are defined topologically.

Lemma 4.13

Let $\mathcal{I} = \mathcal{P}_1^{\alpha_1} \cdots \mathcal{P}_l^{\alpha_l}$ and let $\underset{\sim}{x} \in \Omega$ with $\underset{\sim}{x} = \phi(x)$ for $x \in \overline{\mathcal{D}}$. Then $\mathcal{I}|\underset{\sim}{x}$ if and only if $x \in \overline{\mathcal{I}}$.

Proof

Consider the following open set,

$$U = \{\underset{\sim}{z} \in \Omega : |z_{\mathcal{P}_i}|_{\mathcal{P}_i} \leq 1/N(\mathcal{P}_i)^{\alpha_i}, i = 1, \cdots, l\}.$$

There is a sequence of elements $\{\underset{\sim}{x}_n\} = \{(x_n, x_n, \cdots)\}$ of Δ converging to $\underset{\sim}{x}$ by Lemma 4.9.

Suppose $\mathcal{I}|\underset{\sim}{x}$. Then $\underset{\sim}{x} \in U$ and so $\underset{\sim}{x}_n \in U$ for $n \geq$ some N_0. In that case $\mathcal{I}|\underset{\sim}{x}_n$ and so $x_n \in \mathcal{I}$ for $n \geq N_0$. Therefore as

$$x = \phi^{-1}(\underset{\sim}{x}) = \lim_{n\to\infty} \phi^{-1}(\underset{\sim}{x}_n) = \lim_{n\to\infty} x_n,$$

we have $x \in \overline{\mathcal{I}}$.

The reverse conclusion is similar (note that U is closed).

■

5. The Extension of Additive Functions to Ω.

Suppose we have a sequence $\{g_{\mathcal{P}}\}$ of Borel measurable functions from Ω to the reals which are almost everywhere finite and with the property that, for any $\underset{\sim}{x} \in \Omega$, the value of $g_{\mathcal{P}}(\underset{\sim}{x})$ depends only on $x_{\mathcal{P}}$, the $\mathcal{P}$-th component of $\underset{\sim}{x}$. In other words, the value of $g_{\mathcal{P}}(\underset{\sim}{x})$ is "independent" of all coordinates of $\underset{\sim}{x}$ except $x_{\mathcal{P}}$. As we might suspect we have the following result.

Lemma 4.14

The functions $\{g_{\mathcal{P}}\}$ described above are independent functions on the probability space Ω (in the sense of (2.12)).

Proof

For any prime ideal $\mathcal{P}$, we may define, unambiguously, a function from $\mathcal{D}_{\mathcal{P}}$ to the real numbers by

$$h_{\mathcal{P}}(x_{\mathcal{P}}) = g_{\mathcal{P}}(\underset{\sim}{x}),$$

where $\underset{\sim}{x}$ is any element of Ω with a $\mathcal{P}$-th component of $x_{\mathcal{P}}$.

Suppose we have n functions from the sequence $\{g_{\mathcal{P}}\}$ and n real numbers $\lambda_1, \cdots, \lambda_n$. For notational convenience we suppose that these functions are $g_{\mathcal{P}_1}, \cdots, g_{\mathcal{P}_n}$ (the proof being similar in other cases).

Let

$$U = \{\underset{\sim}{x} \in \Omega : g_{\mathcal{P}_1}(\underset{\sim}{x}) \leq \lambda_1, \cdots, g_{\mathcal{P}_n}(\underset{\sim}{x}) \leq \lambda_n\}$$

and, for $i = 1, \cdots, n$, let

$$S_i = \{x_{\mathcal{P}_i} \in \mathcal{D}_{\mathcal{P}_i} : h_{\mathcal{P}_i}(x_{\mathcal{P}_i}) \leq \lambda_i\}.$$

It is easy to see that

$$U = S_1 \times \cdots \times S_n \times \mathcal{D}_{\mathcal{P}_{n+1}} \times \cdots$$

and, for $i = 1, \cdots, n$ that

$$\{\underset{\sim}{x} \in \Omega : g_{\mathcal{P}_i}(\underset{\sim}{x}) \leq \lambda_i\} = \mathcal{D}_{\mathcal{P}_1} \times \cdots \times \mathcal{D}_{\mathcal{P}_{i-1}} \times S_i \times \mathcal{D}_{\mathcal{P}_{i+1}} \times \cdots.$$

Therefore, using P for the measure on Ω,

$$\begin{aligned} P(U) &= \prod_{i=1}^{n} M_{\mathcal{P}_i}(S_i) \\ &= \prod_{i=1}^{n} P(\underset{\sim}{x} \in \Omega : g_{\mathcal{P}_i}(\underset{\sim}{x}) \leq \lambda_i), \end{aligned}$$

which gives the independence of the functions $g_{\mathcal{P}_1}, \cdots, g_{\mathcal{P}_n}$.

■

We will now concentrate on a real valued additive function, f, on the ideals of $\mathcal{D}$ (see (2.4)). We may write, for any ideal $\mathcal{I}$,

$$f(\mathcal{I}) = \sum_{\mathcal{P}^r || \mathcal{I}} f(\mathcal{P}^r)$$

where $\mathcal{P}^r || \mathcal{I}$ means that $\mathcal{P}^r | \mathcal{I}$ but $\mathcal{P}^{r+1} \not| \mathcal{I}$.

In Ω the concept $\mathcal{P}^r || \underset{\sim}{x}$ is defined also. Let $\mathcal{P}$ be a prime ideal and $\underset{\sim}{x} \in \Omega$. Let the $\mathcal{P}$-th component of $\underset{\sim}{x}$ be $x_{\mathcal{P}}$. If $x_{\mathcal{P}} \neq 0$ we say, using (1) in Section 2 above,

$$\mathcal{P}^r || \underset{\sim}{x} \quad \text{if} \quad \mathcal{P}^r | \underset{\sim}{x} \quad \text{but} \quad \mathcal{P}^{r+1} \not| \underset{\sim}{x}$$

or equivalently

$$\mathcal{P}^r || \underset{\sim}{x} \quad \text{if} \quad |x_{\mathcal{P}}|_{\mathcal{P}} = 1/N(\mathcal{P})^r. \tag{2}$$

It is important to note that $P(\underset{\sim}{x} \in \Omega : x_{\mathcal{P}} = 0) = 0$ and so, for almost all $\underset{\sim}{x} \in \Omega$, $\mathcal{P}^r || \underset{\sim}{x}$ for some $r \geq 0$. We now define some functions on Ω associated with f. We define,

$$\overline{f}_{\mathcal{P}}(\underset{\sim}{x}) = \begin{cases} f(\mathcal{P}^r) & \text{if } \mathcal{P}^r \| \underset{\sim}{x}, \\ 0 & \text{if } x_{\mathcal{P}} = 0. \end{cases} \tag{3}$$

Note that for almost all $\underset{\sim}{x} \in \Omega$, $\overline{f}_{\mathcal{P}}(\underset{\sim}{x})$ is given by $f(\mathcal{P}^r)$ for some r. If $\mathcal{P}^0 \| \underset{\sim}{x}$, then $\overline{f}_{\mathcal{P}}(\underset{\sim}{x}) = f(\mathcal{D}) = 0$ and also if $\underset{\sim}{d} = (d, d, \cdots)$, for non-zero $d \in \mathcal{D}$, then

$$\overline{f}_{\mathcal{P}}(\underset{\sim}{d}) = f(\mathcal{P}^r) \quad \text{if} \quad \mathcal{P}^r \| <d> .$$

We also define, using the convention introduced in Chapter 2 Section 2,

$$\overline{f}(\underset{\sim}{x}) = \sum_{\mathcal{P}} \overline{f}_{\mathcal{P}}(\underset{\sim}{x}) \tag{4}$$

and note that, for non-zero $d \in \mathcal{D}$,

$$\overline{f}(\underset{\sim}{d}) = \sum_{\mathcal{P}^r \| <d>} f(\mathcal{P}^r) = f(<d>).$$

For a general $\underset{\sim}{x} \in \Omega$ there is no guarantee that the series $\overline{f}(\underset{\sim}{x})$ will converge because there may be infinitely many primes $\mathcal{P}$ that divide $\underset{\sim}{x}$. In the next theorem we show that $\overline{f}(\underset{\sim}{x})$ converges almost everywhere (a.e.) on Ω under certain growth conditions on $f(\mathcal{P})$. The proof follows Novoselov [1] Proposition 46.

Theorem 4.15

Let f be a real valued additive function on the ideals of $\mathcal{D}$ (in the sense of (2.4)) and suppose the two series below converge:

$$\sum_{\mathcal{P}} \frac{f'(\mathcal{P})}{N(\mathcal{P})}$$

and

$$\sum_{\mathcal{P}} \frac{(f'(\mathcal{P}))^2}{N(\mathcal{P})},$$

where

$$f'(\mathcal{P}) = \begin{cases} f(\mathcal{P}) & \text{if} \quad |f(\mathcal{P})| < 1 \\ 1 & \text{if} \quad |f(\mathcal{P})| \geq 1. \end{cases}$$

Then

$$\overline{f}(\underset{\sim}{x}) = \sum_{\mathcal{P}} \overline{f}_{\mathcal{P}}(\underset{\sim}{x})$$

converges a.e. on Ω where $\overline{f}_{\mathcal{P}}(\underset{\sim}{x})$ is defined by (3) above. Furthermore, for non-zero $d \in \mathcal{D}$, we have $\overline{f}((d, d, \cdots)) = f(<d>)$.

Proof

The second assertion has already been proven.

For $\underset{\sim}{x} \in \Omega$ we saw in (2) and (3), above, that $\overline{f}_{\mathcal{P}}(\underset{\sim}{x})$ depends only on $x_{\mathcal{P}}$, the $\mathcal{P}$-th component of $\underset{\sim}{x}$, and so by Lemma 4.14, the functions $\{\overline{f}_{\mathcal{P}}\}$ are independent (trivially they are measurable). Kolmogorov's Three Series Theorem (see Lemma 2.4) tells us that $\overline{f}(\underset{\sim}{x})$ converges a.e. if and only if the following three series converge:

i) $\sum_{\mathcal{P}} P(\underset{\sim}{x} \in \Omega : |\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| \geq 1)$,

ii) $\sum_{\mathcal{P}} E(\overline{f}^{o}_{\mathcal{P}})$,

iii) $\sum_{\mathcal{P}} (E((\overline{f}^{o}_{\mathcal{P}})^2) - (E(\overline{f}^{o}_{\mathcal{P}}))^2)$,

where

$$\overline{f}^{o}_{\mathcal{P}}(\underset{\sim}{x}) = \begin{cases} \overline{f}_{\mathcal{P}}(\underset{\sim}{x}) & \text{if} \quad |\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| < 1 \\ 0 & \text{if} \quad |\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| \geq 1, \end{cases}$$

is the truncated function associated with $\overline{f}_{\mathcal{P}}(\underset{\sim}{x})$. We may say that if $\mathcal{P}^r || \underset{\sim}{x}$,

$$\overline{f}^{o}_{\mathcal{P}}(\underset{\sim}{x}) = f^{o}(\mathcal{P}^r) = \begin{cases} f(\mathcal{P}^r) & \text{if} \quad |f(\mathcal{P}^r)| < 1 \\ 0 & \text{if} \quad |f(\mathcal{P}^r)| \geq 1. \end{cases}$$

The convergence of the three series above will follow from the convergence of the two series in the hypothesis once we have established the following three equations. For O-constants not dependent on $\mathcal{P}$ we have:

iv) $E(\overline{f}^o_{\mathcal{P}}) = \dfrac{f^o(\mathcal{P})}{N(\mathcal{P})} + O\left(\dfrac{1}{N(\mathcal{P})^2}\right)$,

v) $E((\overline{f}^o_{\mathcal{P}})^2) = \dfrac{(f^o(\mathcal{P}))^2}{N(\mathcal{P})} + O\left(\dfrac{1}{N(\mathcal{P})^2}\right)$,

vi) $P(\underset{\sim}{x} \in \Omega : |\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| \geq 1) = \begin{cases} \frac{1}{N(\mathcal{P})} + O\left(\frac{1}{N(\mathcal{P})^2}\right) & \text{if} \quad |f(\mathcal{P})| \geq 1 \\ O\left(\frac{1}{N(\mathcal{P})^2}\right) & \text{otherwise.} \end{cases}$

The proofs of these equations are similar. We will prove vi) and iv). From Lemma 4.2 we have (denoting measure on Ω by P)

$$\begin{aligned} P(|\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| \geq 1) &= \sum_{|f(\mathcal{P}^r)|\geq 1} P(\underset{\sim}{x} \in \Omega : \mathcal{P}^r || \underset{\sim}{x}) \\ &= \sum_{|f(\mathcal{P}^r)|\geq 1} \frac{1}{N(\mathcal{P})^r} - \frac{1}{N(\mathcal{P})^{r+1}}. \end{aligned}$$

If $|f(\mathcal{P})|\geq$1 then $r = 1$ is included in this sum and so

$$\begin{aligned} \mid P(|\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| \geq 1) - 1/N(\mathcal{P}) \mid &\leq \frac{1}{N(\mathcal{P})^2} + \sum_{|f(\mathcal{P}^r)|\geq 1}^{r>1} \frac{1}{N(\mathcal{P})^r}\left(1 - \frac{1}{N(\mathcal{P})}\right) \\ &\leq \frac{2}{N(\mathcal{P})^2}. \end{aligned}$$

On the other, hand if $|f(\mathcal{P})| < 1$ then $r = 1$ is not included and we just get $P(|\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| \geq 1) = O(1/N(\mathcal{P})^2)$. This proves vi).

The proof of iv) is similar,

$$\begin{aligned} E(\overline{f}^o_{\mathcal{P}}) &= \sum_{r=1}^{\infty} f^o(\mathcal{P}^r)\frac{1}{N(\mathcal{P})^r}\left(1 - \frac{1}{N(\mathcal{P})}\right) \\ &= \frac{f^o(\mathcal{P})}{N(\mathcal{P})} + O\left(\frac{1}{N(\mathcal{P})^2}\right), \end{aligned}$$

since $|f^\circ(\mathcal{P}^r)| < 1$.

We now turn our attention to proving the convergence of the three series i), ii) and iii). From vi) we have

$$\begin{aligned}
&\sum_{\mathcal{P}} P(|\overline{f}_{\mathcal{P}}(\underset{\sim}{x})| \geq 1) \\
&= \sum_{|f(\mathcal{P})|\geq 1} \left(\frac{1}{N(\mathcal{P})} + O\left(\frac{1}{N(\mathcal{P})^2}\right)\right) + \sum_{|f(\mathcal{P})|<1} O\left(\frac{1}{N(\mathcal{P})^2}\right) \\
&= \sum_{|f(\mathcal{P})|\geq 1} \frac{1}{N(\mathcal{P})} + O\left(\sum_{\mathcal{P}} \frac{1}{N(\mathcal{P})^2}\right).
\end{aligned}$$

The first term here is

$$\sum_{|f(\mathcal{P})|\geq 1} \frac{(f'(\mathcal{P}))^2}{N(\mathcal{P})}$$

which is bounded (it is bounded above by the second series of the hypothesis). The second term is trivially bounded. This establishes the convergence of series i). Next, from iv) we have

$$\begin{aligned}
\sum_{\mathcal{P}} E(\overline{f}^\circ_{\mathcal{P}}) &= \sum_{\mathcal{P}} \frac{f^\circ(\mathcal{P})}{N(\mathcal{P})} + O\left(\sum_{\mathcal{P}} \frac{1}{N(\mathcal{P})^2}\right) \\
&= \sum_{|f(\mathcal{P})|<1} \frac{f'(\mathcal{P})}{N(\mathcal{P})} + O(1)
\end{aligned}$$

since $f^\circ(\mathcal{P}) = f'(\mathcal{P})$ for $|f(\mathcal{P})| < 1$ and $f^\circ(\mathcal{P}) = 0$ otherwise. This last series also converges by the convergence of the series in the hypothesis. This gives the convergence of ii).

The convergence of iii) follows in a similar manner.

This completes the proof of Theorem 4.15.

■

Later we will use the previous theorem to establish the Erdös-Wintner Theorem which says that, given the convergence of the two series in Theorem 4.15, the additive function f has a limiting distribution (in a sense to be made precise). The Erdös-Wintner Theorem will be proven by extending f to Ω (via Theorem 4.15) and using the connection between frequency and measure on Ω (as expressed by Lemma 4.8 for example).

CHAPTER 5

LIMITING DISTRIBUTION OF ADDITIVE FUNCTIONS

In this chapter we will combine the two previous areas of study, the arithmetic estimates on $\mathcal{D}$ and the space Ω, to prove some results about the distribution of additive functions defined on the set of ideals of $\mathcal{D}$. In Section 1 we will define a concept of frequency with respect to an integral basis of $\mathcal{D}$ and prove several versions of the Hardy-Ramanujan Theorem in the set of algebraic integers, $\mathcal{D}$ (see Chapter 1 Section 1 for the classical version). In Section 2 we will introduce a special sequence of numbers $\{N_k\}$ which will provide a tool for transferring frequency concepts to the probability space Ω. Several lemmas connecting these concepts will be established. In Section 3 we will prove an analogue of the theorem of Erdös-Wintner for $\mathcal{D}$ (see Chapter 1 Section 1 for the classical version).

The results and proofs will be adapted from those found in Novoselov [1] and Babu [1] for the case $s = 1$ and $\mathcal{D} = \mathbb{Z}$ with some simplifications as noted. Some probabilistic results will be taken from Rényi [1].

1. Frequency and the Hardy-Ramanujan Theorem.

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and n a positive rational integer. Let $\mathcal{R} \subseteq \mathcal{D}$ be the collection of standard representatives mod$<n>$ with respect to the basis $d_1, \cdots, d_s$ as defined in (3.5). That is

$$\mathcal{R} = \{\alpha_1 d_1 + \cdots + \alpha_s d_s : \alpha_i = 1, \cdots, n, \ i = 1, \cdots, s\} \tag{1}$$

For a polyadic integer $x \in \Omega$ we let (in accordance with Lemma 4.6) **$R_n(x)$ be the unique standard representative of x mod $<n>$ with respect to $d_1, \cdots, d_s$**, that is $R_n(x)$ is the unique element of $\mathcal{R}$ such that

$$x - R_n(x) \in \overline{<n>} \tag{2}$$

where $\overline{<n>}$ is the closure of $<n>$ in Ω. Note that $R_n(x) \neq 0$.

We define the **frequency of $\mathcal{A} \subseteq \mathcal{D}$ with respect to $d_1, \cdots, d_s$** as

$$\begin{aligned}\nu_n(\mathcal{A}) &= \frac{1}{n^s}\#\{d \in \mathcal{R} : d \in \mathcal{A}\} \\ &= \frac{1}{n^s}\#\{\sigma(d) \in nH \cap \sigma(\mathcal{D}) : d \in \mathcal{A}\} \qquad (3)\end{aligned}$$

where H is the fundamental domain for $\sigma(\mathcal{D})$ in $\mathbb{R}^s$ as defined by (3.8). We also define the **upper and lower densities of $\mathcal{A} \subseteq \mathcal{D}$ with respect to $d_1, \cdots, d_s$** as

$$\begin{aligned}\overline{\pi}(\mathcal{A}) &= \limsup_{n\to\infty} \nu_n(\mathcal{A}) \\ \underline{\pi}(\mathcal{A}) &= \liminf_{n\to\infty} \nu_n(\mathcal{A}) \qquad (4)\end{aligned}$$

and when these both exist and are equal we speak of the **density of $\mathcal{A} \subseteq \mathcal{D}$ with respect to $d_1, \cdots, d_s$,**

$$\pi(\mathcal{A}) = \lim_{n\to\infty} \nu_n(\mathcal{A}).$$

The set function π has some of the properties a probability measure should have. For example, if $\mathcal{A}$ and $\mathcal{C}$ are disjoint subsets of $\mathcal{D}$ for which $\pi(\mathcal{A})$ and $\pi(\mathcal{C})$ exist, then $\pi(\mathcal{A} \cup \mathcal{C})$ exists and equals $\pi(\mathcal{A}) + \pi(\mathcal{C})$. In other words, π is finitely additive. However, π is not countably additive and, even worse, it is possible to find (even in the case $s = 1$ when $\mathcal{D} = \mathbb{Z}$) two sets $\mathcal{A}$ and $\mathcal{C}$ for which $\pi(\mathcal{A})$ and $\pi(\mathcal{C})$ exist but $\pi(\mathcal{A} \cap \mathcal{C})$ and $\pi(\mathcal{A} \cup \mathcal{C})$ do not (see Kubilius [1] p.23 or Babu [1] Chapter 1 Section 1). This means the subsets of $\mathcal{D}$ for which π is defined do not even form a field of sets. To use the techniques of probability theory, therefore, we need to find a probability space with a measure P which mimics π in some sense. This is the reason for constructing the space Ω in Chapter 4 (Lemma 4.8 already shows how the frequency ν_n is connected with measure and integral). The paper de Kroon [1] seems to ignore these points. For example, to prove his results he uses the Central Limit Theorem with the "probability measure"

$$P_1(E) = \lim_{z\to\infty} \frac{\#\{\mathcal{I} \in E : N(\mathcal{I}) \leq z\}}{\#\{\mathcal{I} : N(\mathcal{I}) \leq z\}}$$

where E is a set of ideals of $\mathcal{D}$. As we saw above, even in the case $s = 1$ when $\mathcal{D} = \mathbb{Z}$, this function is not well behaved.

Already we are in a position to prove a Hardy-Ramanujan Theorem in $\mathcal{D}$. We will present two proofs because this further illustrates the connection between ν_n and the measure P on the space of polyadic integers, Ω, of $\mathbb{K}$, and shows that the Turán-Kubilius inequality (Theorem 3.4) can be regarded as a sort of Tchebycheff inequality and the functions $A(f, n)$, $B(f, n)$ in that inequality as a mean and variance.

Theorem 5.1 (Hardy-Ramanujan)

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$. There is a constant c', dependent upon $d_1, \cdots, d_s$ such that, for all strongly additive functions, f, on the ideals of $\mathcal{D}$ (as defined by (2.6)), for all rational integers $n \geq 3$ and real numbers $\lambda > 0$ we have

$$\nu_n\{\, d : |f(<d>) - A(f, n)| \geq \lambda\sqrt{B(f, n)}\,\} \leq \frac{c'}{\lambda^2},$$

where $A(f, n)$ and $B(f, n)$ are as defined for the Turán-Kubilius inequality (Theorem 3.4) and ν_n is the frequency with respect to $d_1, \cdots, d_s$ (as defined in (3)).

Furthermore, if $\theta(n)$ is any function of n such that $\theta(n) \to \infty$ as $n \to \infty$, then

$$\nu_n\{\, d : |f(<d>) - A(f, n)| \geq \theta(n)\sqrt{B(f, n)}\,\} \to 0$$

as $n \to \infty$.

Proof

We use A and B as abbreviations for $A(f, n)$ and $B(f, n)$.

Version 1: We use the Tchebycheff inequality (see Lemma 2.4) on the finite space of standard representatives mod$<n>$ (as defined in (1)) and then the Turán-Kubilius inequality (see Theorem 3.4). Therefore,

$$\begin{aligned} \nu_n\{d : |f(<d>) - A| \geq \lambda\sqrt{B}\} &\leq \frac{1}{B\lambda^2}\cdot\frac{1}{n^s}\sum_d (f(<d>) - A)^2 \\ &\leq \frac{1}{B\lambda^2}\cdot c'B = \frac{c'}{\lambda^2}. \end{aligned}$$

Here the sum is over the d which are standard representatives mod$<n>$.

Version 2: For $x \in \Omega$ let $R_n(x)$ denote the unique standard representative of x mod $<n>$ with respect to $d_1, \cdots, d_s$ as in (2) above. Define

$$f_n(x) = \sum_{\mathcal{P} | <R_n(x)>} f(\mathcal{P}),$$

and

$$g_n(x) = f_n(x) - A.$$

From the Tchebycheff inequality on the polyadic space Ω we get

$$P(x \in \Omega : |g_n(x)| \geq \lambda\sqrt{B}) \leq \frac{1}{B\lambda^2} \int g_n^2(x) dP.$$

We note that $f_n(x)$, $g_n(x)$ and $g_n^2(x)$ are periodic mod $<n>$ in the sense of Lemma 4.8 and so, from that lemma (with $\mathcal{I} = <n>$),

$$\nu_n\{d : |g_n(d)| \geq \lambda\sqrt{B}\} \leq \frac{1}{B\lambda^2} \cdot \frac{1}{n^s} \sum_d g_n^2(d), \tag{5}$$

where the sum is over the standard representatives d mod $<n>$ as in (1). For such a d we have $R_n(d) = d$ and so

$$g_n(d) = \sum_{\mathcal{P} | <d>} f(\mathcal{P}) - A = f(<d>) - A.$$

We substitute this into (5) and again use the Turán-Kubilius inequality to obtain the first result.

The second result comes from putting $\lambda = \theta(n)$.

■

The second version of the proof above is the analogue in $\mathcal{D}$ of the argument in Novoselov [1] (see Example 1 of Section 6) for the rational integers. Our proof is a little neater, however, as we have isolated the Turán-Kubilius inequality and Novoselov develops the relevant estimates as he needs them during the proof.

The case $\omega(\mathcal{I}) = \sum_{\mathcal{P}|\mathcal{I}} 1$, the number of prime ideals dividing $\mathcal{I}$, holds special interest. In this case we have, from Lemma 2.3,

$$A(\omega, n) = B(\omega, n) = \sum_{N(\mathcal{P}) \leq Kn^s} \frac{1}{N(\mathcal{P})} = \log\log n^s + O(1), \tag{6}$$

(where we have absorbed K into the O-constant). We may obtain a more classical version of the result of Theorem 5.1 as follows.

Corollary 5.2

Let the notation of Theorem 5.1 apply. There are constants c'' and n_o depending on the basis $d_1, \cdots, d_s$ of $\mathcal{D}$, such that, for any $n \geq n_o$ and real $\lambda > 0$,

$$\nu_n\{\, d : |\omega(<\!d\!>) - \log\log n^s| \geq \lambda\sqrt{\log\log n^s}\,\} \leq \frac{c''}{\lambda^2}.$$

Also, if $\theta(n) \to \infty$ as $n \to \infty$, then

$$\nu_n\{\, d : |\omega(<\!d\!>) - \log\log n^s| \geq \theta(n)\sqrt{\log\log n^s}\,\} \to 0$$

as $n \to \infty$.

Proof

In view of the estimates in (6) we may choose n_o such that for $n \geq n_o$,

$$\sqrt{\log\log n^s} - 1/2 \cdot \sqrt{B(\omega, n)} \geq |A(\omega, n) - \log\log n^s|.$$

Firstly suppose $\lambda > 1$. If we have a $d \in \mathcal{D}$ such that,

$$|\omega(<\!d\!>) - \log\log n^s| \geq \lambda\sqrt{\log\log n^s},$$

then, for $n \geq n_o$,

$$\begin{aligned} |\omega(<\!d\!>) - A(\omega, n)| &\geq |\omega(<\!d\!>) - \log\log n^s| - |A(\omega, n) - \log\log n^s| \\ &\geq \lambda(\sqrt{\log\log n^s} - |A(\omega, n) - \log\log n^s|) \\ &\geq \lambda/2 \cdot \sqrt{B(\omega, n)}. \end{aligned}$$

Therefore, using the Hardy-Ramanujan Theorem just proven,

$$\nu_n\{\, d : |\omega(<\!d\!>) - \log\log n^s| \geq \lambda\sqrt{\log\log n^s}\,\}$$

$$\begin{aligned} &\leq \nu_n\{\, d : |\omega(<\!d\!>) - A(\omega, n)| \geq \lambda/2 \cdot \sqrt{B(\omega, n)}\,\} \\ &\leq \frac{c'}{(\lambda/2)^2}. \end{aligned}$$

If $\lambda \leq 1$ then $1/\lambda^2 \geq 1$ and it suffices to choose $c'' = 1$. Therefore, the first result is proven with $c'' = \max\{1, 4c'\}$. It is possible to improve this constant by improving the "1/2" in the first inequality of the proof. This would require increasing n_o however.

The second result comes from putting $\lambda = \theta(n)$.

■

Much stronger versions of Corollary 5.2 exist. For example, see Rieger [2] (a discussion of this result is in Chapter 1 Section 3, above).

2. The Sequence $\{N_k\}$ and some Frequency Results.

Let $\{N_k\}$ be a fixed sequence of positive rational integers with the following properties:

i) $N_k < N_{k+1}$,

ii) $N_{k+1}/N_k \to 1$ as $k \to \infty$,

iii) $N_k \to 0$ in Ω as $k \to \infty$ (that is, for any ideal $\mathcal{I}$ there is a k_o such that $<N_k> \subseteq \mathcal{I}$ for all $k \geq k_o$).

For example, we may choose the sequence defined in Novoselov [1] or Babu [1]:

$$N_k = (k - s(n) + n + 2)n!$$

if $s(n) \leq k < s(n+1)$ where $s(n) = 1^2 + 2^2 + \cdots + n^2$. The properties i) and ii) above are easy to check, and we note that for each rational integer n, there is a k_n such that $n | N_k$ for all $k \geq k_n$. If we then choose n to be in the ideal $\mathcal{I}$ ($n = N(\mathcal{I})$ for example) we have $<N_k> \subseteq <n> \subseteq \mathcal{I}$ for $k \geq k_n$. This gives property iii) above.

We now define a set of measurable functions on Ω, the space of polyadic integers of $\mathbb{K}$ with probability measure P.

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and let $\mathcal{S}$ **be the set of measurable functions, g, from Ω to $\mathbb{R}$ such that**

$$g(R_{N_k}(x)) \xrightarrow{P} g(x), \tag{7}$$

where $\xrightarrow{P}$ denotes convergence in probability (see (2.13)) and $R_{N_k}(x)$ is the representative of x mod$<N_k>$ with respect to $d_1, \cdots, d_s$ in the sense of (2), above. The set $\mathcal{S}$ clearly depends on the particular choice of basis $d_1, \cdots, d_s$ but our major result (see the Erdös-Wintner Theorem 5.7, below) is valid for any choice of basis. It can also be shown (along the lines of Novoselov [1] Proposition 10) that the set $\mathcal{S}$ does not depend on the particular choice of the sequence $\{N_k\}$ with the properties i), ii) and iii) above. We have the following extension of Lemma 4.8.

Lemma 5.3

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$. Let $h(d)$ be any non-negative function from $\mathcal{D}$ to the real numbers, and A any set of real numbers. Then

$$\limsup_{n\to\infty} \frac{1}{n^s} \sum_d h(d) = \limsup_{k\to\infty} \int h(R_{N_k}(x))dP$$

and

$$\overline{\pi}(d : h(d) \in A) = \limsup_{k\to\infty} P(x \in \Omega : h(R_{N_k}(x)) \in A),$$

where $\sum_d$ denotes the sum over the standard representatives mod $<n>$ with respect to $d_1, \cdots, d_s$ as defined in (1) above and $\overline{\pi}$ is as in (4) above.

Furthermore, the above equations will still hold if we replace lim sup with lim inf and $\overline{\pi}$ with $\underline{\pi}$.

Proof

Suppose that $N_k \leq n < N_{k+1}$. Therefore, the standard representatives mod$< N_k >$ are also standard representatives mod $<n>$, which are also standard representatives mod$<N_{k+1}>$. (In the language of (3.8) we have, $N_kH \subseteq nH \subseteq N_{k+1}H$). Therefore,

$$\left(\frac{N_k}{N_{k+1}}\right)^s \frac{1}{N_k^s} \sum_{d_k} h(d_k) \leq \frac{1}{n^s} \sum_d h(d) \leq \left(\frac{N_{k+1}}{N_k}\right)^s \frac{1}{N_{k+1}^s} \sum_{d_{k+1}} h(d_{k+1})$$

where $\sum_{d_k}, \sum_d, \sum_{d_{k+1}}$ denote summation over the standard representatives mod$<N_k>$, $<n>$, $<N_{k+1}>$ respectively.

Now, $h(R_{N_k}(x))$ is periodic mod$<N_k>$ in the sense of Lemma 4.8 and so, from that lemma,

$$\frac{1}{{N_k}^s}\sum_{d_k} h(d_k) = \int h(R_{N_k}(x))dP.$$

The same expression holds with $k+1$ in place of k. If we substitute these expressions into the above inequality and take lim sup we obtain the first result.

The second result is proven the same way (or use the first result on the characteristic function of $\{d : h(d) \in A\}$).

The corresponding results with lim inf and $\underline{\pi}$ are proved similarly.

■

The following results are of a more probabilistic nature and are proven in exactly the same way as in Novoselov [1] or Babu [1] for the case $s = 1$ and $\mathcal{D} = \mathbb{Z}$.

Lemma 5.4

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and $\mathcal{S}$ the corresponding set of measurable functions (see (7) above). Then

i) If $\{g_n\}$is a sequence of functions in $\mathcal{S}$ and g is a real-valued measurable function on Ω, then any two of the following implies the third:

a) $g_n \xrightarrow{P} g$,

b) $g \in \mathcal{S}$,

c) $\lim_{n\to\infty} \overline{\pi}(d : |g(d) - g_n(d)| > \lambda) = 0$,

for all $\lambda > 0$, where $\overline{\pi}$ is as in (4) above.

ii) $\mathcal{S}$ is closed under arithmetic operations. That is, if h, $g \in \mathcal{S}$ and $a, b \in \mathbb{R}$ then the following are also in $\mathcal{S}$: $ah + bg$, gh, $a + h$ and h/g (this last provided g is bounded away from 0 on Ω).

Proof

See Novoselov [1], Propositions 7 and 15 or Babu [1] Lemmas 1.5, 1.6 and 1.7.

■

Part i) of this Lemma tells us that, for functions in $\mathcal{S}$, $\overline{\pi}$ mimics the probability P on Ω. The next lemma gives the fundamental connection between the probability P on Ω and the limiting frequency of functions on $\mathcal{D}$.

Lemma 5.5

Let g be a real-valued measurable function on Ω and let, for real λ,

$$G(\lambda) = P(x \in \Omega : g(x) < \lambda).$$

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and $\mathcal{S}$ the corresponding set of measurable functions (see (7) above).

If $g \in \mathcal{S}$ then

$$G(\lambda) = \lim_{n\to\infty} \nu_n\{d : g(d) < \lambda\}$$

for any point λ, of continuity of G. Here ν_n denotes the frequency with respect to $d_1, \cdots, d_s$ as in (3) above. In other words, functions in $\mathcal{S}$ have limiting distributions on $\mathcal{D}$ which equal their distribution functions on Ω.

Proof

For $g \in \mathcal{S}$ we have $g(R_{N_k}(x)) \xrightarrow{P} g(x)$ and so we have convergence of distribution functions (see Rényi [1] Theorem 4.2.1),

$$\lim_{k\to\infty} P(x \in \Omega : g(R_{N_k}(x)) < \lambda) = G(\lambda),$$

if λ is a point of continuity of G. From Lemma 5.3, as $\overline{\pi} = \underline{\pi}$, then

$$\lim_{n\to\infty} \nu_n\{d : g(d) < \lambda\}$$

exists and equals $G(\lambda)$ for such a λ.

■

The above lemma provides a criterion for deciding whether a function on Ω has a limiting distribution, in some sense, when restricted to $\mathcal{D}$ (remember Ω is the completion of $\mathcal{D}$). Usually the problem is the other way around. We start with a function defined on $\mathcal{D}$ and ask when it has a limiting distribution on $\mathcal{D}$. The above lemma could be used, were it possible to extend our function on $\mathcal{D}$ to a measurable function on Ω and guarantee that the extension is in $\mathcal{S}$. In the next section we will describe one possible extension and explore its consequences.

3. The Erdös-Wintner Theorem.

We recall some definitions. Let f be a real-valued additive function on the ideals of $\mathcal{D}$ (as in (2.4)). Let $x \in \Omega$ and let $\mathcal{P}$ be a fixed prime ideal. As in Chapter 4 Section 5 we put $x_{\mathcal{P}}$ for the $\mathcal{P}$-th component of x, and we say,

$$\mathcal{P}^r || x \text{ if } |x_{\mathcal{P}}|_{\mathcal{P}} = 1/N(\mathcal{P})^r \tag{8}$$

and

$$\overline{f}_{\mathcal{P}}(x) = \begin{cases} f(\mathcal{P}^r) & \text{if } \mathcal{P}^r || x, \\ 0 & \text{if } x_{\mathcal{P}} = 0. \end{cases} \tag{9}$$

Lemma 5.6

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and $\mathcal{S}$ the corresponding set of measurable functions (as defined in (7)). For any finite collection $Q_1, \cdots, Q_n$ of prime ideals,

$$\overline{f}_{Q_1}(x) + \cdots + \overline{f}_{Q_n}(x) \in \mathcal{S}.$$

Proof

From Lemma 5.4 part ii) it suffices to show that $\overline{f}_{\mathcal{P}}(x) \in \mathcal{S}$ for any prime ideal $\mathcal{P}$.

Let $x \in \Omega$ and suppose that $\mathcal{P}^r || x$ for some $r \geq 0$ (which is the case for almost all $x \in \Omega$). From the definition of the sequence $\{N_k\}$ we may find a k_o such that

$$<N_k> \subseteq \mathcal{P}^{r+1} \quad \text{when } k \geq k_o.$$

Thus, since $R_{N_k}(x) = (R_{N_k}(x), R_{N_k}(x), \cdots) \in \Omega$ and

$$x - R_{N_k}(x) \in \overline{<N_k>}$$

we have, for $k \geq k_o$,

$$|x_{\mathcal{P}} - R_{N_k}(x)|_{\mathcal{P}} \leq 1/N(\mathcal{P})^{r+1} < |x_{\mathcal{P}}|_{\mathcal{P}}.$$

The $\mathcal{P}$-adic valuation is non-Archimedean so we have, for $k \geq k_o$,

$$|x_{\mathcal{P}}|_{\mathcal{P}} = |R_{N_k}(x)|_{\mathcal{P}} = 1/N(\mathcal{P})^r,$$

and therefore $\mathcal{P}^r || R_{N_k}(x)$. Thus, for $k \geq k_o$,

$$\overline{f}_{\mathcal{P}}(x) = \overline{f}_{\mathcal{P}}(R_{N_k}(x)).$$

This means that $\overline{f}_{\mathcal{P}}(R_{N_k}(x))$ tends point-wise to $\overline{f}_{\mathcal{P}}(x)$ for almost all $x \in \Omega$ which, in turn, implies convergence in the probability measure P (see Rényi [1], Theorem 4.2.4). Therefore $\overline{f}_{\mathcal{P}}(x) \in \mathcal{S}$.

■

We are now ready for the major result of this chapter.

Theorem 5.7 (Erdös-Wintner)

Let f be a real-valued additive function on the ideals of $\mathcal{D}$ (as in (2.4)). Suppose the following two series converge:

$$\sum_{\mathcal{P}} \frac{f'(\mathcal{P})}{N(\mathcal{P})},$$

$$\sum_{\mathcal{P}} \frac{(f'(\mathcal{P}))^2}{N(\mathcal{P})} \qquad (10)$$

where

$$f'(\mathcal{P}) = \begin{cases} f(\mathcal{P}) & \text{if } |f(\mathcal{P})| < 1 \\ 1 & \text{if } |f(\mathcal{P})| \geq 1. \end{cases}$$

Then, f has a limiting distribution on the principal ideals of $\mathcal{D}$. That is, there is a distribution function F (in the sense of (2.14)) such that, for any integral basis $d_1, \cdots, d_s$ of $\mathcal{D}$ and any point of continuity λ, of F, we have

$$F(\lambda) = \lim_{n \to \infty} \nu_n \{d : f(<d>) < \lambda\}$$

where ν_n is the frequency with respect to $d_1, \cdots, d_s$ as defined in (3) above.

Furthermore, F is continuous for all λ if and only if the following series diverges,

$$\sum_{f(\mathcal{P}) \neq 0} \frac{1}{N(\mathcal{P})}. \qquad (11)$$

Proof

The proof is modelled on Babu [1], Theorem 1.1.

For $x \in \Omega$ put

$$\overline{f}(x) = \sum_{\mathcal{P}} \overline{f}_{\mathcal{P}}(x). \tag{12}$$

Theorem 4.15 gives the convergence of $\overline{f}$ a.e. on Ω and for non-zero $d \in \mathcal{D}$, we have $\overline{f}(d) = f(<d>)$. Let

$$F(\lambda) = P(x \in \Omega : \overline{f}(x) < \lambda) \tag{13}$$

and let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and $\mathcal{S}$ the corresponding set of measurable functions defined by (7) above. We will show $\overline{f} \in \mathcal{S}$ and then Lemma 5.5 gives the first result.

To show $\overline{f} \in \mathcal{S}$ it is sufficient, by Lemmas 5.6 and 5.4 part i) and using the fact that a.e. convergence implies convergence in probability, to show that for any $\lambda > 0$,

$$\overline{\pi}(d : |\overline{f}(d) - \sum_{N(\mathcal{P}) \leq Km^s} \overline{f}_{\mathcal{P}}(d)| > \lambda) \to 0 \tag{14}$$

as $m \to \infty$, where $\overline{\pi}$ is the upper density with respect to $d_1, \cdots, d_s$ as defined by (4) above, and K is the constant from Lemma 3.2.

Define a strongly additive function f^* on the ideals of $\mathcal{D}$ by

$$f^*(\mathcal{I}) = \sum_{\mathcal{P}|\mathcal{I}} f'(\mathcal{P})$$

and sets W, Y_m by

$$\begin{aligned} W &= \{\mathcal{P} : |f(\mathcal{P})| \geq 1\}, \\ Y_m &= \{d \in \mathcal{D} : \text{ either } d \in \mathcal{P}^2 \text{ for some } \mathcal{P} \text{ with } N(\mathcal{P}) > Km^s \\ &\qquad \text{or } d \in Q \text{ for some } Q \in W \text{ with } N(Q) > Km^s\}. \end{aligned}$$

It can be seen that for $d \in Y_m^c$, the complement of Y_m, we have

$$\begin{aligned} |\overline{f}(d) - \sum_{N(\mathcal{P}) \leq Km^s} \overline{f}_{\mathcal{P}}(d)| &= |f^*(<d>) - \sum_{N(\mathcal{P}) \leq Km^s, d \in \mathcal{P}} f^*(\mathcal{P})| \\ &= |\sum_{N(\mathcal{P}) > Km^s, d \in \mathcal{P}} f^*(\mathcal{P})|. \end{aligned}$$

Therefore, for integers $n > m > 0$ and real $\lambda > 0$, we have

$$\begin{aligned}
&\nu_n\{d : |\overline{f}(d) - \sum_{N(\mathcal{P})\leq Km^s} \overline{f}_{\mathcal{P}}(d)| > \lambda\} \\
&\leq \nu_n\{d : d \in Y_m\} + \nu_n\{d : d \in Y_m^c \text{ and } |\overline{f}(d) - \sum_{N(\mathcal{P})\leq Km^s} \overline{f}_{\mathcal{P}}(d)| > \lambda\} \\
&\leq \nu_n\{d : d \in Y_m\} + \nu_n\{d : |\sum_{N(\mathcal{P})>Km^s,\, d\in\mathcal{P}} f^*(\mathcal{P})| > \lambda\}. \qquad (15)
\end{aligned}$$

The second term in (15) is easy to estimate. Using Corollary 3.5 and noting that $f^*(\mathcal{P}) = f'(\mathcal{P})$, so that (in the notation of that corollary), $A(f^*, n) = A(f', n)$, $B(f^*, n) = B(f', n)$ and so on, we have the second term of (15)

$$\leq \frac{4}{\lambda^2}(A(f', n) - A(f', m))^2 + \frac{4c'}{\lambda^2}(B(f', n) - B(f', m)). \qquad (16)$$

For the first term in (15), we use Lemma 3.2 and then Theorem 3.3 (on $\mathcal{I} = \mathcal{P}^2$ and then $\mathcal{I} = Q \in W$ and combining main and error terms) to obtain

$$\begin{aligned}
\nu_n\{d : d \in Y_m\} &\leq \sum_{Km^s<N(\mathcal{P})\leq\sqrt{Kn^s}} \nu_n\{d : d \in \mathcal{P}^2\} \\
&\qquad + \sum_{Km^s<N(Q)\leq Kn^s, Q\in W} \nu_n\{d : d \in Q\} \\
&\ll \sum_{Km^s<N(\mathcal{P})\leq\sqrt{Kn^s}} \frac{1}{N(\mathcal{P})^2} + \sum_{Km^s<N(Q)\leq Kn^s, Q\in W} \frac{1}{N(Q)} \\
&\leq \sum_{Km^s<N(\mathcal{P})} \frac{1}{N(\mathcal{P})^2} + \sum_{Km^s<N(Q)} \frac{(f'(Q))^2}{N(Q)}, \qquad (17)
\end{aligned}$$

where in the last step we have used $(f'(Q))^2 = 1$ for $Q \in W$. The constant implied by $\ll$ does not depend on $\mathcal{P}$, Q, n or m.

We use estimates (16) and (17) in (15) and let $n \to \infty$ and then $m \to \infty$. The convergence of the series (10) in the hypothesis gives (14) and the first part of the theorem is proven.

To prove the second assertion we use Lévy's Theorem (see Lemma 2.5).

The measurable functions $\overline{f}_{\mathcal{P}}(x)$ are purely discrete, independent, and take the values $f(\mathcal{P}^r)$ with probability $\frac{1}{N(\mathcal{P})^r}\left(1-\frac{1}{N(\mathcal{P})}\right)$. The maximum jump of $\overline{f}_{\mathcal{P}}(x)$ can be seen to be,

$$J_{\mathcal{P}} = \begin{cases} 1+O(1/N(\mathcal{P})^2) & \text{if } f(\mathcal{P})=0 \\ 1-1/N(\mathcal{P})+O(1/N(\mathcal{P})^2) & \text{if } f(\mathcal{P})\neq 0. \end{cases}$$

Therefore $\sum\limits_{\mathcal{P}}(1-J_{\mathcal{P}})$ diverges if and only if $\sum\limits_{f(\mathcal{P})\neq 0} 1/N(\mathcal{P})$ diverges. From Levy's Theorem, then, the second result follows.

This completes the proof of Theorem 5.7.

■

We now discuss a few examples of the application of this theorem. They are adapted from the examples in Elliott [1] Chapter 5, pages 188-189.

Let $\zeta(\mathcal{I})$ denote the norm sum of the finite number of divisors of the ideal $\mathcal{I}$,

$$\zeta(\mathcal{I}) = \sum_{\mathcal{L}|\mathcal{I}} N(\mathcal{L}).$$

This function is multiplicative because $N(\mathcal{L})$ is multiplicative, and we have

$$\zeta(\mathcal{I}) = \prod_{\mathcal{P}^r||\mathcal{I}} \zeta(\mathcal{P}^r)$$

where

$$\zeta(\mathcal{P}^r) = 1+N(\mathcal{P})+\cdots+N(\mathcal{P})^r = \frac{N(\mathcal{P})^{r+1}-1}{N(\mathcal{P})-1}.$$

We examine the additive function $f(\mathcal{I}) = \log(\zeta(\mathcal{I})/N(\mathcal{I}))$. We have, for any prime ideal $\mathcal{P}$,

$$0 < f(\mathcal{P}) = \log\frac{1+N(\mathcal{P})}{N(\mathcal{P})} \leq \frac{1}{N(\mathcal{P})}$$

and so the two series (10) of Theorem 5.7 converge and the series (11) diverges. We put $z = e^{\lambda}$ and deduce that, for any integral basis $d_1,\cdots,d_s$ of $\mathcal{D}$ the function

$$F'(z) = \lim_{n\to\infty} \nu_n\{d : \zeta(<d>) < z|N(d)|\}$$

exists and is continuous for all $z > 0$. Furthermore, we have the same limit function F', no matter what the choice of $d_1, \cdots, d_s$ may be.

A further example may be provided by using an analogue of the Euler totient function,

$$\varphi(\mathcal{I}) = \prod_{\mathcal{P}^r \| \mathcal{I}} \varphi(\mathcal{P}^r)$$

where

$$\varphi(\mathcal{P}^r) = N(\mathcal{P})^r - N(\mathcal{P})^{r-1}.$$

Again, using Theorem 5.7, we deduce that

$$F''(z) = \lim_{n \to \infty} \nu_n\{d : \varphi(<d>) < z|N(d)|\}$$

exists, is continuous for all $z > 0$, and is independent of the choice of basis $d_1, \cdots, d_s$ of $\mathcal{D}$.

The above examples show that, if $h(\mathcal{I})$ is a multiplicative function on the ideals of $\mathcal{D}$, then, as $\log |h(\mathcal{I})|$ is additive, we can deduce some information about the distribution of $h(\mathcal{I})$ and the Erdös-Wintner Theorem gives criteria for the existence of a limiting distribution of $h(\mathcal{I})$. There are many general results of this sort. For example we have the following (compare with Babu [1] Theorem 6.2).

Corollary 5.8

Let h be a positive real multiplicative function on the ideals of $\mathcal{D}$, such that for some $\mu > 1$ the series

$$\sum_1 \frac{1}{N(\mathcal{P})}, \ \sum_2 \frac{\log^2 h(\mathcal{P})}{N(\mathcal{P})}, \ \sum_2 \frac{\log h(\mathcal{P})}{N(\mathcal{P})}$$

converge, where $\sum_2$ denotes the sum over $\mathcal{P}$ with $1/\mu < h(\mathcal{P}) < \mu$ and $\sum_1$ denotes the sum over the remaining $\mathcal{P}$. Then h has a distribution function of the sort described in Theorem 5.7.

Proof

Put $f(\mathcal{I}) = \log h(\mathcal{I})$. The Erdös-Wintner Theorem gives $\overline{f} \in \mathcal{S}$ and then $\overline{h} = e^{\overline{f}} \in \mathcal{S}$ for any possible $\mathcal{S}$.

■

In this chapter we have been examining the distribution of additive functions $f(\mathcal{I})$ in terms of their behaviour on principal ideals. That is, we have been investigating for sets of reals numbers, A, the frequency

$$\frac{1}{n^s}\#\{d : f(<d>) \in A\},$$

where the d which are counted are of a special sort (standard representatives mod $<n>$). There are many results concerning the distribution

$$\frac{1}{z}\#\{\mathcal{I} : N(\mathcal{I}) \leq z,\ f(\mathcal{I}) \in A\}$$

of f among all its ideals. It is possible to prove some results of this sort using the methods of this chapter. This is discussed in the next chapter.

CHAPTER 6

SOME EXTENSIONS

In this chapter we present some extensions of the results of the previous chapters and consider some further directions of possible research. In Section 1 we will prove a Hardy-Ramanujan Theorem for a general bounded set in $\mathbb{R}^s$. Then, we will turn our attention to obtaining results for the distribution of additive functions on all the ideals of $\mathcal{D}$ (not just the principal ones). For this purpose, in Section 2 we will present some known results concerning the correspondence between ideals of $\mathcal{D}$ in a given ideal class and certain special elements of $\mathcal{D}$. In Section 3 we will prove Prachar's version of the Hardy-Ramanujan Theorem (see Theorem 6.2 below). In Section 4 we will discuss a Turán-Kubilius inequality for ideals and a consequent strengthening of Theorem 6.2, below. We will then indicate some possible areas of future research.

1. The Hardy-Ramanujan Theorem for Bounded Sets in $\mathbb{R}^s$.

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and K the constant corresponding to $d_1, \cdots, d_s$ as defined in Lemma 3.2. As in Chapter 3 Section 3, let $H = H(d_1, \cdots, d_s)$ be the fundamental domain for $\sigma(\mathcal{D})$ in $\mathbb{R}^s$ defined by

$$H(d_1, \cdots, d_s) = \{t_1\sigma(d_1) + \cdots + t_s\sigma(d_s) : 0 < t_i \leq 1,\ i = 1, \cdots, s\}. \quad (1)$$

Let f be a real-valued additive function on the ideals of $\mathcal{D}$. In accordance with Theorem 3.4 we put

$$\begin{aligned} A(n) = A(f,n) &= \sum_{N(\mathcal{P}) \leq Kn^s} \frac{f(\mathcal{P})}{N(\mathcal{P})} \\ B(n) = B(f,n) &= \sum_{N(\mathcal{P}) \leq Kn^s} \frac{(f(\mathcal{P}))^2}{N(\mathcal{P})}. \end{aligned} \quad (2)$$

In Theorem 5.1, above, we showed that, if f is strongly additive, and $H = H(d_1, \cdots, d_s)$ is as in (1), then

$$\frac{1}{n^s} \#\{\, d \in \mathcal{D} : \sigma(d) \in nH \text{ and } |f(<d>) - A(n)| \geq \theta(n)\sqrt{B(n)}\,\}$$

tends to zero as $n \to \infty$ for any unbounded function $\theta(n)$, of n. We wish to replace the fundamental domain H with a more general subset of $\mathbb{R}^s$.

Theorem 6.1

Let $d_1, \cdots, d_s$ be an integral basis for $\mathcal{D}$ and f a strongly additive function on the ideals of $\mathcal{D}$ (as defined in (2.6) and (2.4)). Let E be any bounded subset of $\mathbb{R}^s$ and $\theta(n)$ any real-valued function of n such that $\theta(n) \to \infty$ as $n \to \infty$. Then, there is a real number $\mu > 0$, which depends on E and $d_1, \cdots, d_s$ but not on n, θ or f such that,

$$\frac{1}{n^s} \#\{\, d \in \mathcal{D} : \sigma(d) \in n\mu^{-1}E \text{ and } |f(<d>) - A(n)| \geq \theta(n)\sqrt{B(n)}\,\}$$

tends to zero as $n \to \infty$, where $A(n)$ and $B(n)$ are as in (2) above.

Proof

Let I be a subset of $\{1, \cdots, s\}$. Using the notation in (1) above, let

$$H_I = H(d_1^*, \cdots, d_s^*)$$

where

$$d_i^* = \begin{cases} d_i & \text{if } i \in I \\ -d_i & \text{if } i \notin I. \end{cases}$$

Since $\sigma(d_1), \cdots, \sigma(d_s)$ is a basis for $\mathbb{R}^s$ over $\mathbb{R}$, any $y \in \mathbb{R}^s$ can be written in the form

$$y = t_1\sigma(d_1) + \cdots + t_s\sigma(d_s). \tag{3}$$

If such a point, y, belongs to the "quadrant" of $\mathbb{R}^s$ with $t_i \geq 0$ (for $i \in I$) and $t_i \leq 0$ (for $i \notin I$) it belongs to $\mu\overline{H_I}$ for some $\mu > 0$, where $\overline{H_I}$ is the closure of H_I. The set E is bounded and so, there exists a positive μ, depending on E and $d_1, \cdots, d_s$, such that

$$E \subseteq \mu \bigcup_I \overline{H_I}$$

and, therefore,

$$n\mu^{-1}E \subseteq \bigcup_I n\overline{H_I} = \bigcup_I (nH_I \cup nJ_I), \tag{4}$$

where

$$J_I = \overline{H_I} \backslash H_I = \{y \in H_I : (3) \text{ holds with some } t_i = 0,\ i = 1, \cdots, s\}.$$

Counting points, we have,

$$\begin{aligned} &\frac{1}{n^s} \#\{d \in \mathcal{D} : \sigma(d) \in n\overline{H_I} \text{ and } |f(<\!d\!>) - A(n)| \geq \theta(n)\sqrt{B(n)}\,\} \\ &\leq \frac{1}{n^s} \#\{d \in \mathcal{D} : \sigma(d) \in nH_I \text{ and } |f(<\!d\!>) - A(n)| \geq \theta(n)\sqrt{B(n)}\,\} \\ &\qquad + \frac{1}{n^s} \#\{d \in \mathcal{D} : \sigma(d) \in nJ_I\}. \end{aligned} \tag{5}$$

By applying the Hardy-Ramanujan Theorem (Theorem 5.1) with the basis $d_1^*, \cdots, d_s^*$, defined above, we see that, for some c_I depending on $d_1^*, \cdots, d_s^*$ (but not on n, f or θ) the first term on the right of (5) is

$$\leq \frac{c_I}{(\theta(n))^2}. \tag{6}$$

Furthermore, from the definition of J_I we see that the lattice points in nJ_I are of the form (3), above, with $t_i = 0, \cdots, n$ if $i \in I$ and $t_i = -n, \cdots, 0$ if $i \notin I$, and some $t_i = 0\,(i = 1, \cdots, s)$. The number of such points is

$$\leq s(n+1)^{s-1}. \tag{7}$$

If we substitute (6) and (7) into (5) and let $n \to \infty$ we get the required result with $n\overline{H_I}$ in place of $n\mu^{-1}E$. Finally, we use (4) and sum over the 2^s possible subsets I, to get the desired result.

■

It is possible to use this result to obtain Prachar's version of the Hardy-Ramanujan Theorem (see Theorem 6.2 below), but we need a way of deducing results about ideals of $\mathcal{D}$ from results about elements of $\mathcal{D}$. A method for doing this, developed by Hecke, is described in the next section.

2. A Fundamental Domain for Units.

For this section we will assume some basic results about ideal classes, fractional ideals and fundamental units for $\mathcal{D}$. In particular, we use the finiteness of the number of ideal classes and of the number of roots of unity in $\mathcal{D}$. All relevant results may be found in Chapters 9 and 12 of Stewart and Tall [1].

Let w be the number of roots of unity of $\mathcal{D}$, $\mathcal{C}$ a fixed ideal class and $\mathcal{L}$ a fixed (integral) ideal in the inverse class $\mathcal{C}^{-1}$. In that case, we have the following one-to-one correspondence:

For each (integral) ideal $\mathcal{I} \in \mathcal{C}$ there is a unique principal ideal $<d> \subseteq \mathcal{L}$ such that $\mathcal{I}\mathcal{L} = <d>$ and, conversely, for each $<d> \subseteq \mathcal{L}$ there is a unique (integral) ideal $\mathcal{I} = \mathcal{L}^{-1} <d> \in \mathcal{C}$ (where $\mathcal{L}^{-1} = \{y \in \mathbb{K} : y\mathcal{L} \subseteq \mathcal{D}\}$ is the inverse fractional ideal of $\mathcal{L}$). Note that if $\mathcal{I}$ corresponds to $<d>$ then $N(\mathcal{I}) \leq z$ if and only if $|N(d)| \leq N(\mathcal{L})z$.

In Section 3 below, we will want to count the number of $\mathcal{I} \in \mathcal{C}$ with certain properties. This will be accomplished by counting the number of $d \in \mathcal{L}$ corresponding to $\mathcal{I}$, but, since $d \in \mathcal{L}$ is determined from $<d> \subseteq \mathcal{L}$ only up to multiplication by units, we need the concept of a fundamental domain for multiplication by units.

From this point on we will identify $\mathcal{D}$ with its embedding $\sigma(\mathcal{D})$, which is a lattice in $\mathbb{R}^s$. Therefore, we will write $\mathcal{D}$ for $\sigma(\mathcal{D})$ and regard $d \in \mathcal{D}$ as being a lattice point and an ideal as a sublattice.

We can define a norm on $\mathbb{R}^s$ which agrees with the usual one on $\mathcal{D}$ as follows. For $y \in \mathbb{R}^s$ of the form,

$$y = (u_1, \cdots, u_{r_1}, u_{r_1+1}, v_{r_1+1}, \cdots, u_{r_1+r_2}, v_{r_1+r_2})$$

we define

$$N(y) = u_1 \cdots u_{r_1}(u_{r_1+1}^2 + v_{r_1+1}^2) \cdots (u_{r_1+r_2}^2 + v_{r_1+r_2}^2).$$

Let $\zeta_1, \cdots, \zeta_{r_1+r_2-1}$ be a fundamental system of units of $\mathcal{D}$ and let U be the free multiplicative group generated by them. Then, every unit of $\mathcal{D}$ is of the form $u\rho$ where $u \in U$ and ρ is a root of unity, so that, U is isomorphic to the factor group of the group of units modulo the group of roots of unity in $\mathcal{D}$.

There is a set $T \subseteq \mathbb{R}^s$, whose construction will be described later, with the following properties:

a) For each non-zero $d \in \mathcal{D}$, there is a unique $\tau \in T$ such that $d = u\tau$ for some unit $u \in U$.

b) $zT = T$ for every $z > 0$.

c) If we put $T(z) = \{\tau \in T : |N(\tau)| \leq z\}$ we have,

$$T(z) = z^{1/s}T(1).$$

d) $T(1) = \{\tau \in T : |N(\tau)| \leq 1\}$ is bounded.

Property a) says that T is a fundamental domain for multiplication by units in U. In view of these properties and the one-to-one correspondence discussed earlier, we have the following one-to-w correspondence (where w is the number of roots of unity).

Main Correspondence

Let $\mathcal{C}$ be an ideal class and $\mathcal{L}$ a fixed integral ideal in the inverse class $\mathcal{C}^{-1}$. Let $T \subseteq \mathbb{R}^s$ have the properties a) to d) above and let $z > 0$.

For each (integral) ideal $\mathcal{I} \in \mathcal{C}$ with $N(\mathcal{I}) \leq z$, there are w corresponding points $d \in \mathcal{D}$ such that,

$$d \in \mathcal{L} \cap (N(\mathcal{L})z)^{1/s}T(1).$$

Conversely, for each $d \in \mathcal{L} \cap (N(\mathcal{L})z)^{1/s}T(1)$, there is a unique (integral) ideal $\mathcal{I} \in \mathcal{C}$ with $N(\mathcal{I}) \leq z$.

In this correspondence, $\mathcal{I}\mathcal{L} = <d>$.

■

We shall now briefly outline a method for constructing a set, T, with properties a) to d) above. A fuller discussion is given in Marcus [1], Chapter 6 (and for a more compact discussion see Lang [1], Chapter 6, Section 3). Given this construction, it is not difficult to obtain the properties a) to d) above.

Consider the following log map from $\mathbb{R}^{*s}$ (the points in $\mathbb{R}^s$ with non-zero coordinates) to $\mathbb{R}^{r_1+r_2}$. For $y \in \mathbb{R}^{*s}$ of the form,

$$y = (u_1, \cdots, u_{r_1}, u_{r_1+1}, v_{r_1+1}, \cdots, u_{r_1+r_2}, v_{r_1+r_2}),$$

define

$$\log y = (l_1, \cdots, l_{r_1+r_2}),$$

where

$$\begin{aligned} l_i &= \log \frac{|u_i|}{|N(y)|^{1/s}} \quad \text{if } i = 1, \cdots, r_1 \\ &= \log \frac{(u_i^2 + v_i^2)}{|N(y)|^{2/s}} \quad \text{if } i = r_1 + 1, \cdots, r_1 + r_2. \end{aligned}$$

Under this map, any point in $\mathbb{R}^{*s}$ maps to the hyperplane in $\mathbb{R}^{r_1+r_2}$ defined by

$$l_1 + \cdots + l_{r_1+r_2} = 0. \tag{8}$$

Furthermore, the units of $\mathcal{D}$ (viewed in $\mathbb{R}^s$) map to a lattice in the hyperplane (8) and the fundamental units, $\zeta_1, \cdots, \zeta_{r_1+r_2-1}$, map to a basis for this lattice over $\mathbb{Z}$. We take a fundamental domain, F, for this lattice of units in the hyperplane (8), and set

$$T = \{\tau \in \mathbb{R}^{*s} : \log \tau \in F\}$$

which is the pre-image of F under the log map. As we said before, the properties a) to d) above are not difficult to check for this choice of T.

3. An Ideal form for the Hardy-Ramanujan Theorem.

Let f be a function on the ideals of $\mathcal{D}$ and $z > 0$ a real number. We define two new summatory functions associated with f,

$$\begin{aligned} A'(z) &= \sum_{N(\mathcal{P}) \leq z} \frac{f(\mathcal{P})}{N(\mathcal{P})} \\ B'(z) &= \sum_{N(\mathcal{P}) \leq z} \frac{(f(\mathcal{P}))^2}{N(\mathcal{P})}, \end{aligned} \tag{9}$$

(so in the notation of (2) above, $B(n) = B'(Kn^s)$ and $A(n) = A'(Kn^s)$).

In this section we will assume that f has the following properties:

i) f is strongly additive,
ii) $|f(\mathcal{P})| \leq 1$ for any prime ideal $\mathcal{P}$,
iii) $B^{'}(z) \to \infty$ as $z \to \infty$,
(the constant 1 in ii) could be replaced by any other positive constant).

Suppose $m(z)$ is a positive function of z, such that

$$\frac{(m(z))^s}{z} \to \alpha > 0 \text{ as } z \to \infty. \tag{10}$$

If we use the estimate

$$\sum_{N(\mathcal{P})\leq z} \frac{1}{N(\mathcal{P})} = \log\log z + O(1)$$

from Lemma 2.3 and properties i), ii) and iii) above, we obtain for $m(z)$ as in (10), and for any constant $K \geq 1$, the following estimates:

iv) $A^{'}(K(m(z))^s) = A^{'}(z) + O(1)$ as $z \to \infty$,

v) $\dfrac{B^{'}(z)}{B^{'}(K(m(z))^s)} \to 1$ as $z \to \infty$,

vi) For any ideals $\mathcal{I}$ and $\mathcal{L}$ of $\mathcal{D}$,

$$|f(\mathcal{I}\mathcal{L}) - f(\mathcal{I})| \leq \sum_{\mathcal{P}|\mathcal{L}} 1 = \omega(\mathcal{L}).$$

In these equations the O-constants depend on the field $\mathbb{K}$ and the constant K but not on z or f.

We are now ready to prove a version of the Hardy-Ramanujan Theorem in ideal form (see also Section 4 below).

Theorem 6.2

Let f be a strongly additive function on the ideals of $\mathcal{D}$ such that $|f(\mathcal{P})| \leq 1$ for all prime ideals and $B^{'}(z) \to \infty$ as $z \to \infty$. Let $\psi(z)$ be any increasing function of real z such that $\psi(z) \to \infty$ as $z \to \infty$. Then,

$$\frac{1}{z}\#\{\mathcal{I} : N(\mathcal{I}) \leq z \text{ and } |f(\mathcal{I}) - A^{'}(z)| \geq \psi(z)\sqrt{B^{'}(z)}\,\}$$

tends to zero as $z \to \infty$, where $A^{'}(z)$ and $B^{'}(z)$ are as defined in (9) above.

Proof

Let $\mathcal{C}$ be any ideal class and $\mathcal{L}$ a fixed integral ideal in the inverse class $\mathcal{C}^{-1}$. Let T be a fundamental domain for units as described in Section 2. In view of the correspondence discussed in Section 2 we have,

$$\begin{aligned}&\frac{w}{z}\#\{\mathcal{I} \in \mathcal{C} : N(\mathcal{I}) \leq z \text{ and } |f(\mathcal{I}) - A'(z)| \geq \psi(z)\sqrt{B'(z)}\} \\ &= \frac{1}{z}\#\{d \in \mathcal{L} \cap \rho T(1) : |f(\mathcal{L}^{-1}{<}d{>}) - A'(z)| \geq \psi(z)\sqrt{B'(z)}\}, \qquad (11)\end{aligned}$$

where $\rho = (N(\mathcal{L})z)^{1/s}$.

Therefore, it suffices to show that the right hand side of (11) tends to zero as $z \to \infty$, and then to sum over the finite number of ideal classes, $\mathcal{C}$.

Let $d_1, \cdots, d_s$ be any integral basis of $\mathcal{D}$ and K the corresponding constant (defined by Lemma 3.2). The set $T(1)$ is bounded, so let μ be the positive constant from Theorem 6.1 corresponding to $T(1)$ and $d_1, \cdots, d_s$.

Define the integer valued function,

$$m(z) = [\,\mu\rho\,] + 1 = [\,\mu(N(\mathcal{L})z)^{1/s}\,] + 1. \qquad (12)$$

We have

$$\frac{(m(z))^s}{z} \to \mu^s N(\mathcal{L}) > 0 \text{ as } z \to \infty,$$

and the estimates in iv), v) and vi) above apply. Using these estimates and an argument similar to the proof of Corollary 5.2, it follows that, to show the right hand side of (11) tends to zero it suffices to show that

$$\frac{1}{m^s}\#\{d \in \mathcal{L} \cap \rho T(1) : |f({<}d{>}) - A(m)| \geq 1/2 \cdot \psi(z)\sqrt{B(m)}\,\} \qquad (13)$$

tends to zero as $z \to \infty$, where $A(m) = A'(Km^s)$, $B(m) = B'(Km^s)$ and we have written m for the integer function $m(z)$. (The details are messy but not difficult).

We now need to show (13) tends to zero. We will use Theorem 6.1 with an appropriately chosen $\theta(n)$.

Define

$$\theta(n) = 1/2 \cdot \psi((n-1)^s \mu^{-s} N(\mathcal{L})^{-1}).$$

It then follows, from (12), that $\theta(m(z)) \leq 1/2 \cdot \psi(z)$ because ψ is increasing. Also, from (12), $\rho \leq m(z)\mu^{-1}$ and the properties of T give $\rho T(1) \subseteq m(z)\mu^{-1}T(1)$.

Therefore, we see that (13) is

$$\leq \frac{1}{m^s}\#\{d \in \mathcal{L} \cap m\mu^{-1}T(1) : |f(<d>) - A(m)| \geq \theta(m)\sqrt{B(m)}\}. \quad (14)$$

We now let $z \to \infty$, then, as $m(z) \to \infty$ and $m(z)$ is integral valued, Theorem 6.1 implies that (14) tends to zero. As noted, this means that (13) and (11) tend to zero and the proof is complete.

■

Specializing to the function

$$\omega(\mathcal{I}) = \sum_{\mathcal{P}|\mathcal{I}} 1$$

we obtain, as in Corollary 5.2, the following result.

Corollary 6.3

Let $\psi(z)$ be an increasing function of real z, such that $\psi(z) \to \infty$ as $z \to \infty$. Then, as $z \to \infty$,

$$\frac{1}{z}\#\{\mathcal{I} : N(\mathcal{I}) \leq z \text{ and } |\omega(\mathcal{I}) - \log\log z| \geq \psi(z)\sqrt{\log\log z}\}$$

tends to zero.

■

This result was originally proved by Prachar [1] (in the case of the function $\psi(z) = (\log\log z)^\epsilon$, $\epsilon > 0$). See also Fluch [1].

4. The Distribution of Functions on all Ideals - Some further Results and Speculations.

In Theorem 6.2 above, we have presented a rather round-about proof of Prachar's Hardy-Ramanujan Theorem for ideals, but the proof does show how results for the distribution of functions on principal ideals (of the sort in Chapter 5) could be converted into results for the distribution on all the ideals. However, the method is rather *ad hoc.* An attempt to prove an Erdös-Wintner Theorem for all ideals from Theorem 5.7, by this method, strikes several problems (for example in Theorem 5.7 we need $B'(z)$ bounded as $z \to \infty$). It may be more profitable to start from scratch and to construct a probability space, like that of Chapter 4, with a probability measure that mimics the distribution of ideals. That is, if we have a set $\mathcal{A}$ of ideals of $\mathcal{D}$, the probability should be connected with, for real z, the frequency

$$\frac{1}{z}\#\{\mathcal{I} : N(\mathcal{I}) \leq z \text{ and } \mathcal{I} \in \mathcal{A}\}.$$

As a first step in this general direction, we will now indicate how to prove a Turán-Kubilius inequality for ideals.

Let $C(z)$ denote the number of ideals of $\mathcal{D}$ of norm no larger than z. The following result is well known (see Marcus [1] Chapter 6, Theorems 39 and 40, or Lang [1] Chapter 6, Theorem 3). For an O-constant dependent on the field $\mathbb{K}$ and a positive field constant χ,

$$C(z) = \#\{\mathcal{I} : N(\mathcal{I}) \leq z\} = \chi z + O(z^{1-1/s}). \tag{15}$$

We should note that χ is explicitly given in terms of other field constants.

From (15) the following simple extensions can be deduced. If $\mathcal{L}$ is a fixed ideal,

$$\#\{\mathcal{I} : N(\mathcal{I}) \leq z \text{ and } \mathcal{L}|\mathcal{I}\} = C(z/N(\mathcal{L}))$$

and, consequently,

$$\#\{\mathcal{I} : N(\mathcal{I}) \leq z \text{ and } \mathcal{L}|\mathcal{I}\} = \frac{C(z)}{N(\mathcal{L})} + O\left(\left(\frac{z}{N(\mathcal{L})}\right)^{1-1/s}\right). \tag{16}$$

In (16) the O-constant depends only on the field $\mathbb{K}$.

If we use (16) and (15) and exactly the same argument as we used to establish Theorem 3.4, we obtain the following.

Theorem 6.4 (Ideal Turán-Kubilius Inequality)

Let f be a strongly additive function on the ideals of $\mathcal{D}$, $A'(z)$ and $B'(z)$ be defined by (9) above, and let $C(z)$ be defined by (15) above. Then, for

$z \geq 3$, we have,

$$\sum_{N(\mathcal{I})\leq z} \left(f(\mathcal{I}) - A'(z)\right)^2 = O\left((z + C(z))\, B'(z)\right)$$
$$= O\left(zB'(z)\right),$$

where the constants implied by the O-notation depend on $\mathbb{K}$ but not on f or z.

■

From this we can prove a strengthening of Theorem 6.2 (compare with Fluch [1]).

Theorem 6.5 (Ideal Hardy-Ramanujan Theorem)

Let f be a strongly additive function on the ideals of $\mathcal{D}$, let $A'(z)$ and $B'(z)$ be as defined in (9) above and let $\lambda > 0$. There is a constant c_1, dependent on $\mathbb{K}$ but not on f or λ, such that for all $z \geq 3$,

$$\frac{1}{z}\#\{\mathcal{I} : N(\mathcal{I}) \leq z \text{ and } |f(\mathcal{I}) - A'(z)| \geq \lambda\sqrt{B'(z)}\,\} \leq \frac{c_1}{\lambda^2}.$$

Proof

The argument proceeds as in Theorem 5.1 (first proof). We use the Tchebycheff inequality on the finite set of ideals $\mathcal{I}$, with $N(\mathcal{I}) \leq z$, and then Theorem 6.4. Thus,

$$\#\{\mathcal{I} : N(\mathcal{I}) \leq z \text{ and } |f(\mathcal{I}) - A'(z)| \geq \lambda\sqrt{B'(z)}\,\}$$
$$\leq \frac{1}{\lambda^2 B'(z)} \sum_{N(\mathcal{I})\leq z} \left(f(\mathcal{I}) - A'(z)\right)^2$$
$$= O\left(\frac{zB'(z)}{\lambda^2 B'(z)}\right).$$

■

Encouraged by these results, we could try to construct a probability space, of the type mentioned at the start of this section, for the distribution of functions on all ideals. There are some problems in obtaining, for example, the appropriate analogues of the results in Section 2 of Chapter 5, but this seems to be a fruitful area for ongoing research.

Finally, it would be desirable to obtain a generalized Erdös-Kac Theorem (see Chapter 1, Section 1) in two senses. Firstly, to obtain a version for the distribution of additive functions among principal ideals (compare with Rieger [2], discussed in Chapter 1, Section 3 above). We have set up the appropriate probability space in this thesis and with the correct sieve results (like those used in Rieger [1]) this should be possible. Sieve results, however, are beyond the scope of this thesis. Secondly, we could hope to deduce an Erdös-Kac Theorem for distribution among all ideals, given that we could construct a new, appropriate, probability space (as mentioned above). This is another area for future work.

BIBLIOGRAPHY

Babu, G.J.

[1] *Probabilistic methods in the theory of arithmetic functions.* The Macmillan Company of India, Delhi, 1978.

[2] *Some results on the distribution of additive arithmetic functions II.* Acta Arith. 23(1973), pp.315-328.

de Kroon, J.P.M.

[1] *The asymptotic behaviour of additive functions in algebraic number theory.* Compos. Math. 17(1965), pp.207-261.

Dugundji, J.

[1] *Topology.* Allyn and Bacon, Boston, 1966.

Elliott, P.D.T.A.

[1] *Probabilistic number theory Vols.I and II.* Springer-Verlag, New York, 1979-1980.

Erdös, P.

[1] *On the density of some sequences of numbers III.* Journ. London Math. Soc. 13(1938), pp.119-127.

Erdös, P. and Kac, M.

[1] *The Gaussian law of errors in the theory of additive number theoretic functions.* Amer. Journ. Math. 62(1940), pp.738-742.

Erdös, P. and Wintner, A.

[1] *Additive arithmetical functions and statistical independence.* Amer. Journ. Math. 61(1939), pp.713-721.

Fluch, W.

[1] *Über einen Satz von Hardy-Ramanujan.* Monatsh. Math. 73(1969), pp.31-35.

Goldstein, L.J.
[1] *Analytic number theory.* Prentice-Hall, Englewood Cliffs, New Jersey, 1971.

Halmos, P.R.
[1] *Measure theory.* Springer-Verlag, New York, 1974.

Hardy, G.H. and Ramanujan, S.
[1] *The normal number of prime factors of a number n.* Quart. Journ. Math.(Oxford) 48(1917), pp.76-92.

Hasse, H.
[1] *Zahlentheorie.* Akademie-Verlag, Berlin, 1949.

Husain, T.
[1] *Introduction to topological groups.* W.B.Saunders Company, Philadelphia, 1966.

Kubilius, J.
[1] *Probabilistic methods in the theory of numbers.* Amer. Math. Soc. Translations of Math. Monographs No.11, Providence, 1964.

Lang, S.
[1] *Algebraic number theory.* Addison-Wesley, Reading, Mass., 1970.

Lévy, P.
[1] *Sur les séries dont les termes sont des variables éventuelles indépendants.* Studia Math. 3(1931), pp.119-155.

Luthar, I.S.
[1] *A note on a result of Mahler's.* Journ. Aust. Math. Soc. 6(1966), pp.399-401.

Mackey, G.W.
[1] *Unitary group representations in physics, probability and number theory.* Benjamin/Cummings, Reading, Mass., 1978.

Mahler, K.

[1] *Inequalities for ideal bases in algebraic number fields.* Journ. Aust. Math. Soc. 4(1964), pp.425-448.

Marcus, D.A.

[1] *Number fields.* Springer-Verlag, New York, 1977.

McFeat, R.B.

[1] *Geometry of numbers in adele spaces.* Dissertationes Mathematicae 88(1971).

Narkiewicz, W.

[1] *Elementary and analytic theory of algebraic numbers.* Polish Scientific Publishers, Warsaw, 1973.

Novoselov, E.V.

[1] *A new method in probabilistic number theory.* Amer. Math. Soc. Translations(2) 52(1966), pp.217-275.

Philipp, W.

[1] *Mixing sequences of random variables and probabilistic number theory.* Amer. Math. Soc. Memoirs No.114, 1971.

Prachar, K.

[1] *Verallgemeinerung eines Satzes von Hardy und Ramanujan auf algebraische Zahlkörper.* Monatsh. Math. 56(1952), pp.229-232.

[2] *Primzahlverteilung.* Springer-Verlag, Berlin, 1957.

Rényi, A.

[1] *Foundations of probability.* Holden-Day, San Francisco, 1970.

Rieger, G.J.

[1] *Verallgemeinerung der Siebmethode von A.Selberg auf algebraische Zahlkörper III.* Journ. Reine Angew. Math. 208(1961), pp.79-90.

[2] *Über die Anzahl der Primfaktoren algebraischer Zahlen und das Gausssche Fehlergesetz.* Math. Nachr. 24(1962), pp.77-89.

Schaal, W.

[1] *Obere und untere Abschätzungen in algebraischen Zahlkörpern mit Hilfe des linearen Selbergschen Siebes.* Acta Arith. 13(1968), pp.267-313.

Stewart, I.N. and Tall, D.O.

[1] *Algebraic number theory.* Chapman and Hall, London, 1979.

Taylor, S.J.

[1] *Introduction to measure and integration.* Cambridge University Press, Cambridge, 1973.

Wilson, R.J.

[1] *The large sieve in algebraic number fields.* Mathematika 16(1969), pp.189-204.

Weiss, E.

[1] *Algebraic number theory.* McGraw-Hill, New York, 1963.