Università della Calabria

Dipartimento di Matematica

Dottorato di Ricerca in Matematica ed Informatica

XXIII CICLO

Settore Disciplinare MAT/05 – ANALISI MATEMATICA

Tesi di Dottorato

Uniform distribution of sequences of points and partitions

Maria Infusino

Supervisore	**Coordinatore**
Prof. Aljoša Volčič	Prof. Nicola Leone

A.A. 2009 – 2010

Università della Calabria

Dipartimento di Matematica

Dottorato di Ricerca in Matematica ed Informatica

XXIII CICLO

Settore Disciplinare MAT/05 – ANALISI MATEMATICA

Tesi di Dottorato

Uniform distribution of sequences of points and partitions
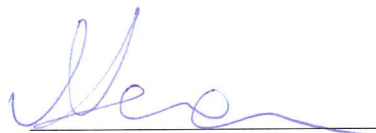
Maria Infusino

Maria Infusino

Supervisore

Prof. Aljoša Volčič

Coordinatore

Prof. Nicola Leone

A.A. 2009 – 2010

To my family

Abstract

The interest for uniformly distributed (u.d.) sequences of points, in particular for low discrepancy sequences, arises from various applications, especially in the field of numerical integration. The basic idea in numerical integration is trying to approximate the integral of a function f by a weighted average of the function evaluated at a set of points $\{x_1, \dots, x_N\}$

$$\int_{I^d} f(x)dx \approx \frac{1}{N} \sum_{i=1}^{N} w_i f(x_i),$$

where I^d is the $d-$dimensional unit hypercube, the x_i's are N points in I^d and $w_i > 0$ are weights such that $\sum_{i=1}^{N} w_i = N$. In some cases it is assumed $w_i = \frac{1}{N}$ for every $1 \leq i \leq N$, as for instance in the classical Monte Carlo method where the points $x_1, \dots, x_N$ are picked from a sequence of random or pseudorandom elements in I^d. Another possibility is to use deterministic sequences with given distribution properties. This procedure is known as Quasi-Monte Carlo method and it is more advantageous than many other approximation techniques. In fact, as the Koksma-Hlawka inequality states, the quality of the approximation provided by the Quasi-Monte Carlo method is linked directly to the discrepancy of the x_i's. The better the nodes are distributed in I^d, the faster the approximation is expected. Hence, a good choice for the integration points is the initial segment of a sequence with small discrepancy.

In this context the construction of u.d. sequences with low discrepancy in various spaces is of crucial importance. The objectives of this thesis are related to this main topic of uniform distribution theory and can be summarized as follows.

(A) the research of explicit techniques for introducing new classes of u.d. sequences of points and of partitions on $[0,1]$ and also on fractal sets;

(B) a quantitative analysis of the distribution behaviour of a class of generalized Kakutani's sequences on $[0,1]$ through the study of their discrepancy.

To achieve these purposes, a fundamental role is played by the concept of u.d. sequences of partitions. In fact when we deal with fractals, and in particular with fractals generated by an Iterated Function System (IFS), partitions turn out to be a convenient tool for introducing a uniform distribution theory. In this thesis we extend to certain fractals the notion of u.d. sequences of partitions, introduced by Kakutani in 1976 for the unit interval and we employ it to construct van der Corput type sequences on a whole class of IFS fractals. More precisely in Chapter 2, where we develop the objective $(\mathbf{A})$, we present a general algorithm to produce u.d. sequences of partitions and of points on the class of fractals generated by a system of similarities on $\mathbb{R}^d$ having the same ratio and verifying the open set condition. We also provide an estimate for the elementary discrepancy of these sequences.

Generalized Kakutani's sequences of partitions of $[0, 1]$ are extremely useful in the extension of these results to a wider class of fractals obtained by eliminating the restriction that the similarities defining the fractal all have the same ratio. According to a remark by Mandelbrot, which allows to see $[0, 1]$ as the attractor of an IFS, the simplest setting for this problem is the unit interval. Perfectly fitting our setting is a recent generalization of Kakutani's splitting procedure on $[0, 1]$, namely the technique of $\rho-$refinements. Consequently, in Chapter 3 we deal with objective $(\mathbf{B})$ and focus on deriving bounds for the discrepancy of the sequences generated by this technique.

Our approach is based on a tree representation of the sequence of partitions constructed by successive $\rho-$refinements, which is exactly the parsing tree generated by Khodak's coding algorithm. This correspondence allows to give bounds of the discrepancy for all the sequences generated by successive $\rho-$refinements, when ρ is a partition of $[0, 1]$ consisting of m subintervals of lenghts $p_1, \ldots, p_m$ such that $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ are rationally related. This result applies also to a countable family of classical Kakutani's sequences and provides estimates of their discrepancy, not known in the existing literature. Moreover, we are also able to cover several irrational cases, which means that at least one of the fractions $\frac{\log p_i}{\log p_j}$ is irrational. More precisely, we discuss some instances in the irrational case when the initial probabilities are p and $q = 1 - p$. In this case we obtain weaker upper bounds for the discrepancy, since they depend heavily on Diophantine approximation properties of the ratio $\frac{\log p}{\log q}$. Finally, we prove bounds for the elementary discrepancy of the sequences of partitions constructed through an adaptation of $\rho-$refinements method to the new class of fractals.

Sommario

L'interesse per le successioni di punti uniformemente distribuite (u.d.) emerge da svariate applicazioni specialmente nell'ambito dell'integrazione numerica. Un approccio tipico di questa disciplina è l'approssimazione dell'integrale di una funzione f con la media pesata dei valori assunti dalla funzione in un insieme di punti $\{x_1, \ldots, x_N\}$

$$\int_{I^d} f(x)dx \approx \frac{1}{N} \sum_{i=1}^{N} w_i f(x_i),$$

dove I^d è l'ipercubo unitario $d-$dimensionale, gli x_i sono N elementi di I^d e i pesi $w_i > 0$ sono tali che $\sum_{i=1}^{N} w_i = N$. In alcuni casi si assume che $w_i = \frac{1}{N}$ per ogni $1 \leq i \leq N$, come ad esempio nel metodo classico di Monte Carlo in cui i punti $x_1, \ldots, x_N$ sono selezionati da una successione casuale o pseudo-casuale di elementi in I^d. Un'altra possibilità è effettuare la scelta degli x_i all'interno di successioni deterministiche con proprietà di distribuzione fissate. Questa procedura è nota come metodo di Quasi-Monte Carlo ed è più vantaggiosa di molte altre tecniche d'approssimazione numerica. Infatti, la disuguaglianza di Koksma-Hlawka stabilisce che la qualità dell'approssimazione fornita dal metodo di Quasi-Monte Carlo è strettamente legata alla discrepanza degli x_i. Pertanto, risulta conveniente scegliere come insieme dei punti di integrazione il segmento iniziale di una successione a bassa discrepanza.

La ricerca di successioni di punti u.d. con bassa discrepanza è dunque di importanza cruciale in ambito applicativo. Gli obiettivi di questo lavoro si collocano all'interno di questo filone di ricerca e interessano due tematiche fondamentali:

(A) la ricerca di tecniche esplicite che consentano di costruire successioni u.d. di punti e di partizioni su $[0, 1]$ e su insiemi frattali;

(B) l'analisi del comportamento asintotico della discrepanza di una classe di successioni di partizioni di Kakutani generalizzate.

Nei risultati proposti uno strumento essenziale è il concetto di successione di partizioni u.d.. Infatti quando si lavora con i frattali, ed in particolare con frattali generati da un Sistema di Funzioni Iterate (IFS), le partizioni risultano essere più convenienti delle successioni di punti in relazione alla teoria della distribuzione uniforme. Pertanto, abbiamo esteso ai frattali la definizione di successione di partizioni u.d., introdotta da Kakutani nel 1976 per partizioni di $[0,1]$, ed abbiamo sfruttato questo concetto per costruire successioni di tipo van der Corput su un'intera classe di frattali IFS. Più precisamente nel Capitolo 2, in cui viene affrontata la tematica $(\mathbf{A})$, presentiamo un algoritmo per generare successioni u.d. di punti e di partizioni sui frattali individuati da un numero finito di similitudini su $\mathbb{R}^d$, aventi tutte lo stesso rapporto di similitudine e che soddifano la condizione dell'insieme aperto. Inoltre abbiamo ricavato una stima della discrepanza elementare delle successioni prodotte.

La seconda problematica studiata è l'estensione dei risultati ottenuti a una classe più ampia di frattali, eliminando la restrizione che le similitudini dell'IFS abbiano tutte lo stesso rapporto. Secondo un'osservazione dovuta a Mandelbrot, che ci consente di vedere $[0,1]$ come attrattore di infiniti IFS, l'ambientazione più semplice per tale problema è proprio l'intervallo unitario. Una tecnica che si adatta perfettamente alle caratteristiche della nuova classe di attrattori è una recente generalizzazione della procedura di Kakutani: la tecnica dei ρ-raffinamenti. Pertanto, nel Capitolo 3 affrontiamo la tematica $(\mathbf{B})$ con l'obiettivo di determinare stime della discrepanza delle successioni di partizioni di $[0,1]$ prodotte tramite tale tecnica.

L'approccio che usiamo è basato su una rappresentazione ad albero di questa classe di successioni che produce lo stesso albero costruito secondo l'algoritmo di Khodak. Questa corrispondenza consente di ricavare stime della discrepanza delle successioni generate dai successivi $\rho-$raffinamenti dell'intervallo unitario, quando ρ è una partizione che consiste di m intervalli di lunghezza $p_1, \ldots, p_m$ tali che $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ siano razionalmente correlati. Questo caso include una classe numerabile di successioni di Kakutani classiche, per le quali otteniamo stime della discrepanza ancora non presenti in letteratura. Per quanto concerne il caso irrazionale, cioè quando almeno uno dei rapporti $\frac{\log p_i}{\log p_j}$ non è razionale, sono state osservate diverse complicazioni. In questo lavoro analizziamo la situazione in cui ρ è costituita da due intervalli di lunghezza p e $q = 1 - p$. Tuttavia, le stime della discrepanza ottenute in questo sottocaso sono più deboli, in quanto dipendono fortemente dalle proprietà di approssimazione diofantea del rapporto $\frac{\log p}{\log q}$. Infine, introduciamo alcuni risultati sulla discrepanza elementare delle successioni di partizioni costruite tramite un adattamento del metodo dei $\rho-$raffinamenti alla nuova classe di frattali.

Contents

List of Figures

Introduction

The theory of uniform distribution was developed extensively within and among several mathematical disciplines and numerous applications. In fact, the main root of this theory is number theory and diophantine approximation, but there are strong connections to various fields of mathematics such as measure theory, probability theory, harmonic analysis, summability theory, discrete mathematics and numerical analysis.

The central goals of this theory are the assessment of uniform distribution and the construction of uniformly distributed (u.d.) sequences in various mathematical spaces. The objectives of this thesis are related to these main topics. In particular, the aim of this work is to introduce new classes of u.d. sequences of points and of partitions on $[0,1]$ and also on fractal sets. Moreover, we intend to present a quantitative analysis of the distribution behaviour of the new sequences produced by studying their discrepancy.

The problem of finding explicit methods for constructing u.d. sequences was originally investigated in the setting of sequence of points. In fact, the starting point of the development of theory was just the study of u.d. sequences of points on the unit interval. The result which marked the beginning of the theory was the discovery that the fractional parts of the multiples of an irrational number are u.d. in the unit interval or, equivalently, on the unit circle. This was a refinement of an approximation theorem due to Kronecker who had already proved the density of this special sequence in the unit interval. So, at the beginning of the last century, many authors independently proposed the theorem about uniform distribution of Kronecker's sequence such as Bohl [5], Sierpiński [62] and Weyl [70]. The latter was the first to estabilish a systematic treatment of uniform distribution theory in his famous paper [72], where the formal definition of u.d. sequences of points in $[0,1]$ was given for the first time. Moreover, in that paper the theory of u.d. sequences of points was generalized to the higher-dimensional unit cube.

1

The uniform distribution of a sequence of points means that the empirical distribution of the sequence is asymptotically equal to the uniform distribution. Therefore in the twenthies and thirties several authors began to study u.d. sequences of points from a quantitative point of view introducing the discrepancy [4, 67, 72]. This quantity is the classical measure of the deviation of a sequence from the ideal uniform distribution. Consequently, having a precise estimate of the discrepancy is very useful for applications but is not a trivial problem. Proving general lower bounds for the discrepancy is a subject still having open questions nowadays.

The interest for u.d. sequences of points, in particular for low discrepancy sequences, arises from various applications in areas like numerical integration, random number generation, stochastic simulation and approximation theory. Indeed, numerical integration was one of the first applications of uniform distribution theory [38]. The basic problem considered by numerical integration is computing an approximate solution to a definite integral. The classical quadrature formulae are less and less efficient the higher the dimension is. To overcome this problem, a typical approach is trying to approximate the integral of a function f by a weighted average of the function evaluated at a set of points $\{x_1, \ldots, x_N\}$

$$\int_{I^d} f(x)dx \approx \frac{1}{N} \sum_{i=1}^{N} w_i f(x_i),$$

where I^d is the $d-$dimensional unit hypercube, the x_i's are N points in I^d and $w_i > 0$ are weights such that $\sum_{i=1}^{N} w_i = N$. In some cases it is assumed $w_i = \frac{1}{N}$ for every $1 \leq i \leq N$, as for instance in the classical Monte Carlo method where the points $x_1, \ldots, x_N$ are picked from a sequence of random or pseudorandom elements in I^d. The advantage of the Monte Carlo method is that it is less sensitive to the increase of the dimension.

Another possibility is to use deterministic sequences with given distribution properties for the choice of the x_i's. This procedure is known as Quasi-Monte Carlo method and it is more advantageous than many other approximation techniques. In fact, the Koksma-Hlawka inequality (1.14) shows that the error of such a method can be bounded by the product of a term only depending on the discrepancy of $\{x_1, \ldots, x_N\}$ and one only depending on the function. Therefore it is convenient to choose the initial segment of a low discrepancy sequence as the set of integration points in the Quasi-Monte Carlo method. These are sequences with a discrepancy of order $\frac{(\log N)^d}{N}$ where d is the dimension of the space in which we take the sequence. Hence, by using low discrepancy sequences, the Quasi-Monte Carlo method

has a faster rate of convergence than a corresponding Monte Carlo method, since in the latter case the point sets necessarily do not have minimal discrepancy. Infact, it behaves, in average, as $\frac{1}{\sqrt{N}}$. Indeed, the Monte Carlo method yields only a probabilistic bound on the integration error. Nevertheless, both Monte Carlo and Quasi-Monte Carlo methods offer the advantage to add further points without recalculating the whole series and this is a big step forward compared to classical methods. Quasi-Monte Carlo methods have an important role in financial and actuary mathematics, where high-dimensional integrals occur. During the last twenty years all these applications have been a rapidly growing area of research, [52, 31].

One of the best known techiniques for generating low discrepancy sequences of points in the unit interval was introduced by van der Corput in 1935, [66]. Successively, van der Corput's procedure was extended to the higher-dimensional case by Halton, [28]. Moreover, a generalization of van der Corput sequences is due to Faure who introduced the permuted or generalized van der Corput sequences. They are also very interesting because there exist formulae for the discrepancy of these sequences which show their good asymptotic behaviour, [13, 24, 25].

The study of van der Corput type sequences has not been limited to the classical setting of the unit interval in one dimension or the unit hypercube in higher dimensions, but interesting extensions have been made to more abstract spaces such as fractals. In fact, the theory of uniform distribution with respect to a given measure has been generalized in several ways: sequences of points in compact and locally compact spaces [45, 51, 32], sequences of probability measures on a separable compact space [60], in particular sequences of discrete measures associated to partitions of a compact interval [41] and to partitions of a separable metric space [14]. In the following we use the basic definitions of uniform distribution theory in compact Hausdorff spaces and in a particular class of fractal compact sets.

Fractals are involved in several applications because they are a powerful tool to describe effectively a variety of phenomena in a large number of fields. To exploit Quasi-Monte Carlo methods on these sets it is essential to study discrepancy bounds for sequences of points on fractals. One of the earlier papers devoted to uniform distribution on fractals is [27], where this theory is developed on the Sierpiński gasket. In this paper the notion of discrepancy on fractals has been introduced for the first time. The authors define several concepts of discrepancy for sequences of points on the Sierpiński gasket by choosing different kinds of partitions on this fractal. Succes-

sively, these notions were generalized also to other fractals, such as the $d-$dimensional Sierpiński carpet in [18, 17]. In particular, in [17] a van der Corput type construction is considered to generate u.d. sequences of points on the $d-$dimensional Sierpiński carpet and the exact order of convergence of various notions of discrepancy is determined for these sequences.

In this work we get a more general result by constructing van der Corput type sequences on a whole class of fractals generated by an Iterated Function System (IFS). More precisely, we are going to study fractals defined by a system of similarities on $\mathbb{R}^d$ having the same ratio and verifying a natural separation condition of their components, namely the open set condition (OSC). This class includes the most popular kind of fractals, but also the unit interval $[0,1]$ that can be seen as the attractor of infinitely many IFS. Starting from this remark, which goes back to Mandelbrot [48], we present a different construction of the classical van der Corput sequences of points on $[0,1]$. By imitating this approach, we introduce an explicit procedure to define u.d. sequences of points on our special class of fractals, (see Subsection 2.2.1). So we call these sequences of *van der Corput type*, just to emphasize the particular order given to the points by our algorithm. It is important to underline that as probability on a fractal F of our class we take the normalized s-dimensional Hausdorff measure, where s is the Hausdorff dimension of F. This is the most natural choice for a probability measure on this kind of fractals, also because the OSC guarantees the existence of an easy formula for evaluating the Hausdorff dimension of these fractals, (see Theorem 2.1.6). A crucial role in the proof of the uniform distribution of the sequences constructed is played by the elementary sets, i.e. the family of all sets generated by applying our algorithm to the whole fractal F. In this way our technique produces also u.d. sequences of partitions of the fractals belonging to the class considered.

The concept of u.d. sequence of partitions on fractals is just one of the most important aspects of this thesis. When we deal with fractals, and in particular with IFS fractals, partitions turn out to be a more convenient tool in relation to the uniform distribution theory. Consequently we extend the notion of u.d. sequences of partitions, introduced by Kakutani in 1976 for the unit interval in [41], to our class of fractals.

The construction ideated by Kakutani, called Kakutani's splitting procedure, allows to construct a whole class of u.d. sequences of partitions of $[0,1]$ and it is based on the concept of $\alpha-$refinement of a partition. For a fixed $\alpha \in]0,1[$, the $\alpha-$refinement of a partition π is obtained by splitting all the intervals of π having maximal lenght

in two parts, proportional to α and $1 - \alpha$ respectively. Kakutani proved that the sequence of partitions generated through successive α−refinements of the trivial partition $\omega = \{[0, 1]\}$ is u.d.. This result received a considerable attention in the late seventies, when other authors provided different proofs of Kakutani's theorem [1] and of its stochastic versions, in which the intervals of maximal lenght are splitted according to certain probability distributions [68, 46, 47, 8, 55]. Recently different generalizations of Kakutani's technique have been introduced. A result in this direction is the extension of Kakutani's splitting procedure to the multidimensional case with a construction which is intrinsically higher-dimensional [12]. Moreover, in a recent paper of Volčič, Kakutani's technique is extended also in the one dimensional case by introducing the concept of ρ−refinement of a partition, which generalizes Kakutani α−refinements. Actually, the ρ−refinement of a partition π is obtained by splitting the longest intervals of π into a finite number of parts homothetically to a given finite partition ρ of $[0, 1]$. The author has proved that the technique of successive ρ−refinements allows to construct new families of u.d. sequences of partitions of $[0, 1]$ in [69]. The last paper also investigates the connections of the theory of u.d. sequences of partitions to the well-estabilished theory of u.d. sequences of points, showing how it is possible to associate u.d. sequences of points to any u.d. sequence of partitions.

Generalized Kakutani's sequences on $[0, 1]$ are a fundamental tool in the extension of the results obtained on our class of fractals. The first attempt of enlarging the class of fractals considered in our previous analysis consists in eliminating the restriction that the similarities defining the fractal all have the same ratio.

The procedure of successive ρ-refinements fits perfectly to the problem of generating u.d. sequences of partitions on this new class of fractals. Indeed, let $\psi = \{\psi_1, \ldots, \psi_m\}$ be a system of m similarities on $\mathbb{R}^d$ having ratio $c_1, \ldots, c_m \in \,]0, 1[$ respectively and such that they verify the OSC. Let F be the attractor of ψ and let s be its Hausdorff dimension. By applying the m similarities successively to the fractal F, we get a first partition consisting of m subsets of F each of probability $p_i = c_i^s$ (where for probability we again mean the normalized s−dimensional Hausdorff measure). At the second step we choose the susbsets with the highest probability and we apply to each of them the m similarities in the same order, and so on. Iterating this procedure, which exploits the same basic idea of ρ−refinements, we obtain a sequence of partitions of F. Now the problem is the assessment of the uniform distribution of these sequences and the estimation of their discrepancy.

According to the Mandelbrot's remark the simplest setting for this problem is the unit interval. In fact, if we consider $[0, 1]$ as the attractor of m similarities $\varphi_1, \dots, \varphi_m$ having different ratios and satisfying the OSC and we apply the procedure described above, then we get exactly the sequence of $\rho-$refinements $(\rho^n \omega)$, where $\rho = \{\varphi_1([0, 1]), \dots, \varphi_m([0, 1])\}$ and $\omega = \{[0, 1]\}$.

In the second part of this work we focus on deriving bounds for the discrepancy of the generalized Kakutani's sequences of partitions of $[0, 1]$ generated by using the techinique of successive $\rho-$refinements. The problem of estimating the asymptotic behaviour of the discrepancy of this kind of sequences has been posed for the first time in [69]. At the moment the only known discrepancy bounds for a class of such sequences have been given by Carbone [10]. In this paper the author considered the so-called LS-sequences that are generated by successive $\rho-$refinements where ρ is a partition with L subintervals of $[0, 1]$ of length α and S subintervals of length α^2 (where α is given by the equation $L\alpha + S\alpha^2 = 1$).

To study this problem in more generality we use a correspondence between the procedure of successive $\rho-$refinements and Khodak's algorithm [43]. This new approach is based on a parsing tree related to the Khodak coding algorithm, which represents the successive ρ-refinements. We introduce improvements of the results obtained in [20] to provide significative bounds of the discrepancy for all the sequences generated by successive $\rho-$refinements, when ρ is a partition of $[0, 1]$ consisting of m subintervals of lenghts $p_1, \dots, p_m$ such that $\log\left(\frac{1}{p_1}\right), \dots, \log\left(\frac{1}{p_m}\right)$ are rationally related. This result applies also to classical Kakutani's sequences and provides, for the first time after thirty years, quantitative estimates of their discrepancy. Moreover, the class of generalized Kakutani's sequences belonging to this rational case also includes the $LS-$sequences.

In the following we are also able to cover several irrational cases, when at least one of the fractions $\frac{\log p_i}{\log p_j}$ is irrational. This case is much more involved than the rational one. In this work we discuss some instances in the irrational case when the initial probabilities are two, namely p and q. The upper bounds for the discrepancy that we obtain in this subcase are weaker, since they depend heavily on Diophantine approximation properties of the ratio $\frac{\log p}{\log q}$. Furthermore, if the initial partition is composed of more than two intervals, then the analysis of the behaviour of the discrepancy is even more complicated, as evident by comparing with [26].

The approach applied for achieving these bounds of the discrepancy of generalized Kakutani's sequences on $[0, 1]$ can be also used for the sequences of partitions

constructed on fractals defined by similarities which do not have the same ratio and satisfing the OSC. In fact, we have described above an analogue of the method of successive $\rho-$refinements which allows to produce sequences of partitions on this new class of fractals. We actually introduce a new correspondence between nodes of the tree associated to Khodak's algorithm and the subsets belonging to the partitions generated on the fractal. Consequently, with a technique similar to the one used on $[0, 1]$, we can prove bounds for the elementary discrepancy of these sequences of partitions, too.

Let us give a brief outline of the thesis.

Chapter 1 provides the basic background knowledge on the areas of uniform distribution theory that are investigated in this thesis. The first part of the chapter deals with the classical part of the theory. Basic definitions and properties of u.d. sequences of points on the unit interval are introduced and specific examples of u.d. sequences of points are described throughout. Then a whole section is devoted to the more recent theory of u.d. sequences of partitions, which plays an essential role in this work. Some extensions of the theory of uniform distribution are also touched on in this chapter, such as the multidimensional case and the theory in Hausdorff compact spaces.

Chapter 2 regards the uniform distribution on a special class of fractals. More precisely, we are concerned with fractals generated by an Iterated Function System (IFS) of similarities having the same ratio and satisfying the Open Set Condition (OSC). We propose an algorithm for generating u.d. sequences of partitions and of points on this class of fractals. Furthermore, in the last part of this chapter we study the order of convergence of the elementary discrepancy of van der Corput type sequences constructed on these fractals. The results presented in this chapter have been first published in [40].

In **Chapter 3** we extend the results given in the second chapter to a wider class of fractals by using a new approach, which allows us to derive bounds for the discrepancy of a class of generalized Kakutani's sequences of partitions of $[0, 1]$, constructed through successive $\rho-$refinements. We present the recent technique of $\rho-$refinements and the generalization of Kakutani's theorem to the class of sequences of partitions generated by this procedure. Then, we analyze the behaviour of the discrepancy of these sequences from a new point of view. The crucial idea is a tree representation of the sequence of partitions constructed by successive $\rho-$refinements, which is precisely the parsing tree generated by Khodak's coding algorithm. The

correspondence between the two techniques allows not only to give optimal upper bounds in the so-called rational case on $[0, 1]$ but also to extend the results obtained in the second chapter to a wider class of fractals. Moreover, we study the irrational case which is more involved then the rational one. Finally, we give some examples and applications of the results achieved so far. The new contributions presented in this chapter are collected in [19].

The thesis concludes by reviewing, in **Chapter 4**, the main results we have obtained and indicates open problems and directions of future research.

Chapter 1

Preliminary topics

This chapter is meant to give a short overview of known results about uniform distribution theory not only in the classical setting of $[0,1]$ but also in more general spaces. First we intend to mention necessary definitions and basic results concerning the u.d. sequences of points in $[0,1]$. Then we will introduce the more recent theory of u.d. sequences of partitions which is fundamental in the development of this work. Finally, we will point out the main aspects of uniform distribution theory on the unite hypercube and on compact spaces.

1.1 Uniformly distributed sequences of points in $[0,1]$

In this section we develop the classical part of uniform distribution theory. The standard references for this topic are [45] and [21]. We start introducing the basic concepts related to u.d. sequences of points and then we proceed to consider the quantitative aspect of the theory. Moreover, a whole subsection is devoted to a special class of sequences with certain advantageous distribution properties, namely the van der Corput sequences.

1.1.1 Definitions and basic properties

First of all, let us state the main definition of the theory.

Definition 1.1.1.
A sequence (x_n) of points in $[0,1]$ is said to be uniformly distributed (u.d.) *if for any real number a such that $0 < a \leq 1$ we have*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \chi_{[0,a[}(x_n) = a \tag{1.1}$$

where $\chi_{[0,a[}$ is the characteristic function of the interval $[0, a[$.

Let us introduce some concepts which are very useful to characterize u.d. sequences of points.

Definition 1.1.2.
A class $\mathcal{F}$ of Riemann-integrable functions on $[0, 1]$ is said to be determining *for the uniform distribution of sequences of points, if for any sequence (x_n) in $[0, 1]$ the validity of the relation*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) = \int_0^1 f(x) \ dx \qquad (1.2)$$

for all $f \in \mathcal{F}$ already implies that (x_n) is u.d.. In particular, a system of subsets of $[0, 1]$ such that the family of their characteristic functions is determining is called discrepancy system*.*

Hence, we can restate the Definition 1.1.1 saying that the family of all characteristic functions $\chi_{[0,a[}$ for $0 < a \leq 1$ is determining or that the system of all sets $[0, a[$ for $0 < a \leq 1$ is a discrepancy system.

An important determining class is the family of all continuous (real or complex-valued) functions on $[0, 1]$. This result is due to Weyl and it is very useful to extend the theory to more general spaces, [71, 72].

Theorem 1.1.3 (Weyl's Theorem).
A sequence (x_n) of points in $[0, 1]$ is u.d. if and only if for any (real or complex-valued) continuous function f defined on $[0, 1]$ the equation (1.2) holds.

Proof.
Let (x_n) be u.d. and let f be a step function

$$f(x) = \sum_{i=0}^{k-1} c_i \chi_{[a_i, a_{i+1}[}(x_n) \qquad (1.3)$$

where $0 = a_0 < a_1 < \ldots < a_k = 1$ and $c_i \in \mathbb{R}$ for $i = 0, \ldots, k - 1$. Then it follows

by (1.1) and (1.3) that

$$
\begin{aligned}
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) &= \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \sum_{i=0}^{k-1} c_i \chi_{[a_i, a_{i+1}[}(x_n) \\
&= \sum_{i=0}^{k-1} c_i \lim_{N \to \infty} \frac{1}{N} \left(\sum_{n=1}^{N} \chi_{[0, a_{i+1}[}(x_n) - \sum_{n=1}^{N} \chi_{[0, a_i[}(x_n) \right) \\
&= \sum_{i=0}^{k-1} c_i \left(\int_0^1 \chi_{[0, a_{i+1}[}(x) \, dx - \int_0^1 \chi_{[0, a_i[}(x) \, dx \right) \\
&= \sum_{i=0}^{k-1} \int_0^1 c_i \chi_{[a_i, a_{i+1}[}(x) \, dx \\
&= \int_0^1 f(x) \, dx
\end{aligned}
$$

Now, assume that f is a real-valued function defined on $[0, 1]$. Fixed $\varepsilon > 0$, by the definition of the Riemann integral, there exist two step functions f_1 and f_2 such that

$$
f_1(x) \leq f(x) \leq f_2(x) \ , \ \forall x \in [0, 1]
$$

and

$$
\int_0^1 (f_2(x) - f_1(x)) \, dx \leq \varepsilon.
$$

Then we have the following chain of inequalities

$$
\begin{aligned}
\int_0^1 f(x) \, dx - \varepsilon &\leq \int_0^1 f_2(x) \, dx - \varepsilon \leq \int_0^1 f_1(x) \, dx = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f_1(x_n) \\
&\leq \liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) \leq \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) \\
&\leq \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f_2(x_n) = \int_0^1 f_2(x) \, dx \\
&\leq \int_0^1 f_1(x) \, dx + \varepsilon \leq \int_0^1 f(x) \, dx + \varepsilon
\end{aligned}
$$

So the relation (1.2) holds for all continuous functions.

Conversely, let (x_n) a sequence of points in $[0, 1]$ such that the (1.2) holds for every real-valued continuous function f defined on $[0, 1]$. Let $a \in]0, 1[$, then given any $\varepsilon > 0$ there exist two continuous functions g_1 and g_2 such that

$$
g_1(x) \leq \chi_{[0, a[}(x) \leq g_2(x) \ , \ \forall x \in [0, 1]
$$

and

$$\int_0^1 \left(g_2(x) - g_1(x) \right) \, dx \leq \varepsilon.$$

Then we have

$$
\begin{aligned}
a - \varepsilon \quad &\leq \quad \int_0^1 g_2(x) \, dx - \varepsilon \leq \int_0^1 g_1(x) \, dx = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} g_1(x_n) \\
&\leq \quad \liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \chi_{[0,a[}(x_n) \leq \limsup_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \chi_{[0,a[}(x_n) \\
&\leq \quad \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} g_2(x_n) = \int_0^1 g_2(x) \, dx \\
&\leq \quad \int_0^1 g_1(x) \, dx + \varepsilon \leq \ a + \varepsilon
\end{aligned}
$$

Since ε is arbitrarily small, we have (1.1).

$\square$

Moreover, we can state a more general result.

Theorem 1.1.4.

A sequence (x_n) of points in $[0,1]$ is u.d. if and only if for any Riemann-integrable function f defined on $[0,1]$ the equation (1.2) holds.

Proof.
The sufficiency directly follows by the previous theorem, because every continuous function is Riemann-integrable. The other implication was shown by De Bruijn and Post [9], who proved that if f is defined on $[0,1]$ and if the averages $\frac{1}{N} \sum_{n=1}^{N} f(x_n)$ admit limit for any (x_n) u.d., then f is Reimann-integrable.

$\square$

The problem of finding the largest reasonable determining classes has been addressed also in [14] and [57].

Other examples of determing classes are the following ones.

Examples

- The class of all characteristic functions of open (closed or half-open) subintervals of $[0, 1]$ is determining.

- The class of all characteristic functions of all intervals of the type $[0, q]$ with $q \in \mathbb{Q}$ is determining.

- The class of all step functions, i.e. functions given by finite linear combinations of characteristic functions of half-open subintervals of $[0, 1]$ is determining.

- The class of all continuous (real or complex-valued) functions g on $[0, 1]$ such that $g(0) = g(1)$ is determining.

- The class of all polynomials with rational coefficients is determining.

Now, consider all functions of the type $f(x) = e^{2\pi i h x}$ where h is a non-zero integer. One of the most important facts of uniform distribution theory is that these functions give a criterion to determine if a sequence of points is u.d..

Theorem 1.1.5 (Weyl's Criterion).
The sequence (x_n) is u.d. if and only if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n} = 0$$

for all integers $h \neq 0$.

This important result was proved for the first time by Weyl in [72], but a lot of proofs can be find in literature. Moreover, this criterion has a variety of applications in the uniform distribution theory and also in the estimation of exponential sums. In particular, Weyl applied this theorem to the special sequence $(\{n\theta\})$, with θ irrational, to give a new proof of the following theorem.

Let us recall that for any $x \in \mathbb{R}$, we denote by $\{x\}$ the fractional part of x, which satisfies $\{x\} = x - [x]$, where $[x]$ is the integral part of x (i.e the greatest integer less or equal to x).

Theorem 1.1.6.
Let θ an irrational number. Then the sequence $(\{n\theta\})$ is u.d..

This result was independently estabilished by Weyl [70], Bohl [5] and Sierpiński [62] in 1909-1910. The problem of the distribution of this special sequence has its origin in the theory of secular perturbations in astronomy and signs the beginning of the theory of u.d. sequences of points. Theorem 1.1.6 improves a previous theorem due to Kronecker, who proved that the points $e^{in\theta}$ are dense in the unit circle, whenever θ is an irrational multiple of π (Kronecker's approximation theorem). For this reason the sequence $(\{n\theta\})$ with θ irrational is called *Kronecker sequence*.

Finally, it is important to underline that uniform distribution has also a measure-theoretic aspect. In fact, if we look at Definition 1.1.1, we realize that a sequence (x_n) of points in $[0,1]$ is u.d. if and only if the sequence of discrete measures $\left(\frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \right)$ converges weakly to the Lebesgue measure λ on $[0,1]$, where δ_t is the Dirac measure concentrated in t.

The notion of weak convergence of measures represents the link between u.d. sequence of points and u.d. sequence of partitions.

1.1.2 Discrepancy of sequences

As a quantitative measure of the distribution behaviour of a u.d. sequence we consider the so-called discrepancy, that is the maximal deviation between the empirical distribution of the sequence and the uniform distribution. This notion was studied for the first time in a paper of Bergström, who used the term "Intensitätdispersion", [4]. The term discrepancy was probably coined by van der Corput. Moreover, the first intensive study of discrepancy is due to van der Corput and Pisot in [67].

Definition 1.1.7 (Discrepancy)**.**
Let $\omega_N = \{x_1, \ldots, x_N\}$ be a finite set of real numbers in $[0,1]$. The number

$$D_N(\omega_N) = \sup_{0 \leq a < b \leq 1} \left| \frac{1}{N} \sum_{i=1}^{N} \chi_{[a,b[}(x_i) - (b-a) \right|$$

is called the discrepancy *of the given set ω_N.*

If (x_n) is an infinite sequence of points, we associate to it the sequence of positive real numbers $D_N(\{x_1, x_2, \ldots x_N\})$. So, the symbol $D_N(x_n)$ denotes the discrepancy of the initial segment $\{x_1, x_2, \ldots x_N\}$ of the infinite sequence.

The importance of the concept of discrepancy in uniform distribution theory is revealed by the following fact, [72].

Theorem 1.1.8.
A sequence (x_n) of points in $[0,1]$ is u.d. if and only if

$$\lim_{n \to \infty} D_N(x_n) = 0.$$

Sometimes it is useful to restrict the family of intervals considered in the definition of discrepancy. The most important type of restriction is to consider only intervals of the form $[0, a[$ with $0 < a \leq 1$.

Definition 1.1.9 (Star discrepancy).

Let $\omega_N = \{x_1, \ldots, x_N\}$ be a finite set of real numbers in $[0,1]$, we define star discrepancy of ω_N the quantity

$$D_N^*(\omega_N) = \sup_{0 < a \leq 1} \left| \frac{1}{N} \sum_{i=1}^N \chi_{[0,a[}(x_i) - a \right|.$$

The definition D_N^* is extended to the infinite sequence in the same way as we did for D_N. Moreover, the discrepancy and star discrepancy are related by the following inequality.

Theorem 1.1.10.

For any sequence (x_n) of points in $[0,1]$ we have

$$D_N^*(x_n) \leq D_N(x_n) \leq 2D_N^*(x_n).$$

The most prominent open problem in theory of irregularities of distribution is to determine the optimal lower bound for the discrepancy. A first trivial lower bound is given by

Proposition 1.1.11.

For any finite set $\omega = \{x_1, \ldots, x_N\}$ in $[0,1]$ we have that

$$\frac{1}{N} \leq D_N(\omega) \leq 1.$$

The finite set $x_n = \frac{n}{N}$, $n = 1, \ldots, N$ satisfies $D_N(x_1, \ldots, x_N) = \frac{1}{N}$. But sequences of this kind can only exist in the one-dimensional case, by a theorem due to Roth [56] and this shows that the lower bound is optimal. Moreover, in this example it is easy to see that for every N a new set $x_1, \ldots, x_N$ is constructed. So the natural question is if there exists an infinite sequence (x_n) in $[0,1]$ such that $D_N(x_n) = \mathcal{O}\left(\frac{1}{N}\right)$ as $N \to \infty$.

Van der Corput made the conjecture that there are no sequences of this kind in the unit interval and this was proved by van Aardenne-Ehrenfest in [64, 65]. But the van der Corput conjecture was completely solved also from a quantitative point of view with the following important result due to Schmidt, [61].

Theorem 1.1.12 (Schmidt's Theorem).

For any sequence (x_n) in $[0,1]$ we have that

$$N D_N(x_n) > c \log N$$

for infinitely many positive integers N, where $c > 0$ is an absolute constant.

This lower bound is the best possible in the one-dimensional case.

Usually, sequences having discrepancy of the order $\mathcal{O}\left(\frac{\log N}{N}\right)$ are called *low discrepancy sequences* and they are very important for several applications. An interesting example of this kind of sequences are the van der Corput sequences.

1.1.3   The van der Corput sequence

In 1935 van der Corput introduced a procedure to generate low discrepancy sequences on $[0,1]$, (see [66]). These sequences are considered the best distributed on $[0,1]$, because no infinite sequence has yet been found with discrepancy of smaller order of magnitude than the van der Corput sequences. The technique of van der Corput is based on a very simple idea. First of all we have to define the radical inverse function which is at the basis of this construction.

Definition 1.1.13 (Radical-inverse function)**.**
Let $b \geq 2$ an integer and let $n = \sum\limits_{k=0}^{r} a_k b^k$ be the digital expansion of the integer $n \geq 1$ in base b, $a_k \in \{0, \ldots, b-1\}$. The function

$$\gamma_b(n) = \sum_{k=0}^{r} a_k b^{-k-1}$$

is called radical inverse function in base b.

The radical inverse function $\gamma_b(n)$ represents the fraction lying between 0 and 1 constructed by reversing the order of the digits in the $b-$adic expansion of n.

Definition 1.1.14 (van der Corput sequences)**.**
Let $b \geq 2$ a fixed prime integer. The sequence $(x_n)_{n \geq 1}$, where

$$x_n = \gamma_b(n-1),$$

is called van der Corput sequence in base b.

For example, the van der Corput sequence in base $b = 2$ is given by

$$0, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \cdots$$

The construction of these points is explicitely showed in the following table.

$\mathbf{n-1}$	$\mathbf{(n-1)}$ in base 2	$\mathbf{x_n}$ in base 2	$\mathbf{x_n}$
0	0.0	0.0	0
1	1.0	0.1	$\frac{1}{2}$
2	10.0	0.01	$\frac{1}{4}$
3	11.0	0.11	$\frac{3}{4}$
4	100.0	0.001	$\frac{1}{8}$
5	101.0	0.101	$\frac{5}{8}$
6	110.0	0.011	$\frac{3}{8}$
7	111.0	0.111	$\frac{7}{8}$

and so on. Let us introduce the estimate of the convergence order of the discrepancy of van der Corput sequence in base 2. But before proving this result, we need some further notions (see [45]).

Definition 1.1.15.
For $0 \leq \delta < 1$ and $\varepsilon > 0$, a finite set $x_1 < x_2 < \cdots < x_N$ in $[0, 1]$ is said to be an almost-arithmetic progression *if there exists an η with $0 < \eta \leq \varepsilon$ such that the following conditions are satisfied:*

- $0 \leq x_1 \leq \eta + \delta\eta$;

- $\eta - \eta\delta \leq x_{n+1} - x_n \leq \eta + \eta\delta$ for $1 \leq n \leq N - 1$;

- $1 - \eta - \delta\eta \leq x_N < 1$.

It is clear that if $\delta = 0$, then we have a genuine arithmetic progression with difference η.

Theorem 1.1.16.
Let $x_1 < x_2 < \cdots < x_N$ be an almost-arithmetic progression in $[0, 1]$ and let η be the parameter in Definition 1.1.15. Then

$$D_N^*(x_1, \ldots, x_N) \leq \frac{1}{N} + \frac{\delta}{1 + \sqrt{1 - \delta^2}} \quad for \; \delta > 0$$

$$D_N^*(x_1, \ldots, x_N) \leq \min\left\{\eta, \frac{1}{N}\right\} \quad for \; \delta = 0.$$

Theorem 1.1.17.
Let $\omega = \{x_1, \ldots, x_N\}$ be a finite set of N points in $[0, 1]$. For $1 \leq j \leq r$ let ω_j be a subset of ω consisting of N_j elements such that its discrepancy is $D_{N_j}(\omega_j)$, its star

discrepancy is $D_{N_j}^*(\omega_j)$, $\omega_j \cap \omega_i = \emptyset$ for all $j \neq i$ and $N = N_1 + \ldots + N_r$. Then

$$D_N(\omega) \leq \sum_{j=1}^{r} \frac{N_j}{N} D_{N_j}(\omega_j)$$

and also

$$D_N^*(\omega) \leq \sum_{j=1}^{r} \frac{N_j}{N} D_{N_j}^*(\omega_j).$$

Now, we are ready to prove the following result.

Theorem 1.1.18.
The discrepancy $D_N(x_n)$ of the van der Corput sequence in base 2 satisfies

$$D_N(x_n) \leq c \, \frac{\log(N+1)}{N}$$

where $c > 0$ is an absolute constant.

Proof.
Let $N \geq 1$. We represent N by its dyadic expansion

$$N = 2^{h_1} + \ldots + 2^{h_r} \qquad \text{with} \quad h_1 > h_2 > \ldots > h_r \geq 0.$$

Partition the interval $[1, N] \cap \mathbb{N}$ of integers in r subsets $M_1, \ldots, M_r$ defined as follows

$$M_j = [2^{h_1} + \ldots + 2^{h_{j-1}} + 1, 2^{h_1} + \ldots + 2^{h_{j-1}} + 2^{h_j}] \cap \mathbb{N} \qquad \text{for} \quad 1 < j \leq r$$

and put $M_1 = [0, 2^{h_1}] \cap \mathbb{N}$.

An integer $n \in M_j$ can be written in the form

$$n = 1 + 2^{h_1} + \ldots + 2^{h_{j-1}} + \sum_{i=0}^{h_j - 1} a_i 2^i, \quad \text{with} \quad a_i \in \{0, 1\}.$$

In fact, we get all 2^{h_j} integers in M_j if we let the a_j run through all the possible combinations of 0 and 1. It follows that the point x_n of the van der Corput sequence is given by

$$x_n = 2^{-h_1 - 1} + \ldots + 2^{-h_{j-1} - 1} + \sum_{i=0}^{h_j - 1} a_i 2^{-i-1} = y_j + \sum_{i=0}^{h_j - 1} a_i 2^{-i-1}$$

where y_j only depends on j and not on n.

If n runs through M_j, then the sum $\sum_{i=0}^{h_j - 1} a_i 2^{-i-1}$ runs through all fractions $0, 2^{-h_j}, \ldots, (2^{h_j} - 1) \cdot 2^{-h_j}$. Moreover, we can note that $0 \leq y_j < 2^{-h_j}$.

We conclude that if the elements x_n with $n \in M_j$ are ordered according to their magnitude, then we obtain a sequence ω_j consisting of $N_j = 2^{h_j}$ elements that is an arithmetic progression with parameters $\delta = 0$ and $\eta = 2^{-h_j}$, (see Definition 1.1.15). By Theorem 1.1.16, we have that

$$D_{N_j}^*(\omega_j) \leq \min\left\{\eta, \frac{1}{N_j}\right\} = 2^{-h_j}.$$

The set of the first N terms of the van der Corput sequence, i.e. $\omega = \{x_1, \ldots, x_N\}$, can be decomposed in the r subset ω_j defined above, since $N = N_1 + \cdots + N_r = 2^{h_1} + \ldots + 2^{h_r}$. Hence, by Theorem 1.1.17 we have

$$D_N^*(\omega) \leq \sum_{j=1}^r \frac{N_j}{N} D_{N_j}^*(\omega_j) \leq \sum_{j=1}^r \frac{1}{N} = \frac{r}{N}. \tag{1.4}$$

It remains to estimate r in terms of N. Since $h_1 > h_2 > \ldots > h_r \geq 0$ then we have that $h_r \geq 0$, $h_{r-1} \geq 1$, $h_{r-2} \geq 2, \ldots, h_1 \geq r - 1$. So we have that

$$N = 2^{h_1} + \ldots + 2^{h_r} \geq 2^{r-1} + \ldots + 2^0 = 2^r - 1,$$

and so

$$r \leq \frac{\log(N+1)}{\log 2}. \tag{1.5}$$

Finally, by combinig (1.4) and (1.5) we have

$$D_N^*(\omega) \leq \frac{\log(N+1)}{N \log 2}$$

and since Theorem 1.1.10 holds, we have

$$D_N(\omega) \leq \left(\frac{2}{\log 2}\right) \frac{\log(N+1)}{N}.$$

$\square$

1.2 Uniformly distributed sequences of partitions on $[0,1]$

In this section, we will consider u.d. sequences of partitions of $[0,1]$, a concept which has been introduced in 1976 by Kakutani, [41]. In particular, we will sketch the theory of u.d. sequences of partitions introducing the significant example constructed by Kakutani. In the second part of this section, we will investigate the relation between u.d. sequences of partitions and u.d. sequences of points. This topic is analyzed more thoroughly in [69].

Firstly, let us give the basic definitions.

Definition 1.2.1.

Let (π_n) be a sequence of partitions of $[0, 1]$, where $\pi_n = \{[t^n_{i-1}, t^n_i] : 1 \le i \le k(n)\}$. The sequence (π_n) is said to be uniformly distributed (u.d.) if for any continuous function f on $[0, 1]$ we have

$$\lim_{n \to \infty} \frac{1}{k(n)} \sum_{i=1}^{k(n)} f(t^n_i) = \int_0^1 f(t) \, dt. \tag{1.6}$$

Equivalently, (π_n) is u.d. if the sequence of discrepancies

$$D_n = \sup_{0 \le a < b \le 1} \left| \frac{1}{k(n)} \sum_{i=1}^{k(n)} \chi_{[a,b[}(t^{(n)}_i) - (b - a) \right| \tag{1.7}$$

tends to 0 as $n \to \infty$.

Similarly to the sequences of points, we can note that the uniform distribution of the sequence of partitions (π_n) is equivalent to the weak convergence to λ of the associated sequences of measures (ν_n), with

$$\nu_n = \frac{1}{k(n)} \sum_{i=1}^{k(n)} \delta_{t^n_i}. \tag{1.8}$$

Moreover, it is easy to see that the uniform distribution of the sequence of partitions (π_n) is equivalent to each of the following two conditions.

1. For any choice of the points $\tau^n_i \in [t^n_{i-1}, t^n_i]$ we have

$$\lim_{n \to \infty} \frac{1}{k(n)} \sum_{i=1}^{k(n)} f(\tau^n_i) = \int_0^1 f(t) \, dt$$

   for any continuous function f on $[0, 1]$.

2. For any choice of the points $\tau^n_i \in [t^n_{i-1}, t^n_i]$ we have that the sequence of measures

$$\frac{1}{k(n)} \sum_{i=1}^{k(n)} \delta_{\tau^n_i}$$

   converges weakly to the Lebesgue measure λ on $[0, 1]$.

1.2.1 Kakutani's splitting procedure

Let us describe a particular technique which allows to construct a whole class of u.d. sequences of partitions of $[0, 1]$. This procedure was introduced by Kakutani in 1976 and works through successive $\alpha-$refinements of the unit interval, [41].

Definition 1.2.2.

If $\alpha \in]0, 1[$ and $\pi = \{[t_{i-1}, t_i] : 1 \leq i \leq k\}$ is any partition of $[0, 1]$, then Kakutani's α-refinement of π (which will be denoted by $\alpha\pi$) is obtained by splitting only the intervals of π having maximal lenght in two parts, proportional to α and $\beta = 1 - \alpha$ respectively.

We will denote by $\alpha^2\pi$ the α-refinement of $\alpha\pi$ and, in general, by $\alpha^n\pi$ the $\alpha-$refinement of $\alpha^{n-1}\pi$. Starting with the trivial partition ω of $[0, 1]$, i.e. $\omega = \{[0, 1]\}$, we get Kakutani's sequence of partitions $\kappa_n = \alpha^n\omega$.

For example, if $\alpha < \beta$ we have that
$\kappa_1 = \{[0, \alpha], [\alpha, 1]\}$
$\kappa_2 = \{[0, \alpha], [\alpha, \alpha + \alpha\beta], [\alpha + \alpha\beta, 1]\}$
and so on.

About this splitting procedure Kakutani proved the following result.

Theorem 1.2.3.

For every $\alpha \in]0, 1[$ the sequence of partitions (κ_n) of $[0, 1]$ is u.d..

The most transparent proof of this theorem is due to Adler and Flatto and follows from a combination of classical results from ergodic theory, [1]. Indeed, Kakutani's procedure caught the attention of several authors in the late seventies also from a stochastic point of view. In fact, Kakutani's theorem was a partial answer to the following question posed by the physicist H. Araki, which regarded random splitting of the interval $[0, 1]$. Let X_1 be choosen randomly with respect to uniform distribution on $[0, 1]$. Once $X_1, \ldots, X_n$ have been choosen, let X_{n+1} be a point choosen at random and accordingly to the uniform distribution in the largest of the $n + 1$ intervals determined by the previous n points. Kakutani had been originally asked whether the associated sequence of empirical distribution functions converges uniformly, with probability 1, to the distribution function of the uniform random variable on $[0, 1]$.

This question has been studied in [68, 46, 47, 8] and later in [55]. It is important to note that in the probabilistic setting the possibility that the partition obtained at the $n-$th step has more than one interval of maximal lenght can be neglected, since it is an event which has probability equal to zero. On the other hand, in Kakutani's splitting procedure for every α the partition $\alpha^n\omega$ has more than one interval of maximal lenght for infinitely many values of n.

Recently, some new results and ideas revived the interest for this subject. In fact, Kakutani's technique has been generalized in several directions. In [12] the splitting procedure has been generalized to higher dimensions, providing a sequence

of nodes in the hypercube $[0,1]^d$ which is proved to be u.d.. In [11], a von Neumann type theorem is presented for sequences of partitions of $[0,1]$. More precisely, u.d. sequences of partitions of the unit interval are constructed starting from sequences of partitions π_n whose diameter tends to zero for $n \to \infty$. In [69], the concept of $\alpha-$refinement is generalized and it is introduced a new splitting procedure for constructing a larger class of u.d. sequences of partitions on $[0,1]$. Moreover, in this paper it is analyzed the deep relation between the theory of u.d. sequences of partitions and the theory of u.d. sequences of points. This strong connection between the two theories makes more interesting the study of u.d. sequences of partitions in view of possible applications to Quasi-Monte Carlo methods.

1.2.2 Associated uniformly distributed sequences of points

In the following, we intend to study the problem of associating to a u.d. sequence of partitions a u.d. sequence of points. Before investigating this problem, let us note that the converse problem results to be easier in many cases.

Theorem 1.2.4.
If (x_n) is a u.d. sequence of points in $[0,1]$ such that $x_n \neq x_m$ when $n \neq m$ and $x_n \notin \{0,1\}$ for any $n \in \mathbb{N}$, then the sequence of partitions (π_n), where each π_n is determined by the points $\{0, 1, x_k \text{ with } k \leq n\}$ ordered by magnitude, is u.d. .

Proof.
By using the assumption that (x_n) is u.d. and the Theorem 1.1.3, it follows that the relation (1.6) holds for any continuous function f defined on $[0,1]$.

$\square$

The requirement that $x_n \neq x_m$ when $n \neq m$ is important and it is not possible to avoid this assumption in the theorem as it is shown in the following example.

Example
Consider the sequence (x_n) defined by consecutive blocks of $4m$ points for $m \in \mathbb{N}$. Each block is defined as follows

$$\left\{ \frac{1}{2m+1}, \frac{1}{2m+1}, \ldots, \frac{m}{2m+1}, \frac{m}{2m+1}, \ldots, \frac{1}{2}, \frac{2m+1}{4m}, \frac{2m+2}{4m}, \ldots, \frac{4m-1}{4m} \right\}.$$

In each block the first m points are repeated twice, while the others are all distinct. In this way, the points of the sequence have double density in the right half of $[0,1]$, but they have however a good distribution because of the repetition in the left

half of $[0,1]$. So the sequence (x_n) is u.d.. But when we take in consideration the sequence of partitions (π_n) associated to (x_n), according to the procedure described in the previous theorem, the repetitions are cancelled. Hence, we get a sequence (π_n) having twice as many subintervals in $\left[\frac{1}{2}, 1\right]$ than in $\left[0, \frac{1}{2}\right[$ and so (π_n) is not u.d..

Now, consider our starting problem of associating a u.d. sequence of points to a fixed u.d. sequence of partitions. Let us introduce an important result proved by Volčič in [69], where a probabilistic answer to this problem is given.

Suppose (π_n) is a u.d. sequence of partitions in $[0,1]$ with $\pi_n = \{[t_{i-1}^n, t_i^n] : 1 \leq i \leq k(n)\}$. The natural question is if it is possible to rearrange the points t_i^n determining the partitions π_n, for $1 \leq i \leq k(n)$, in order to get a u.d. sequence of points. Clearly, there exist many ways of reordering the points t_i^n. A natural restriction is that we first reorder all the points determining π_1 then those defining π_2, and so on. This kind of reorderings are called *sequential reorderings*.

Before presenting the result of Volčič, we need some preliminaries. In particular, we introduce a version of the strong law of large numbers for negatively correlated random variables, which is attributed to Aleksander Rajchman and can be proved following the lines of Theorem 5.1.2 in [16].

Lemma 1.2.5.
Let (φ_n) be a sequence of real, negatively correlated random variables with variances uniformly bounded by V on the probability space (W, P). Moreover, suppose that

$$\lim_{i \to \infty} E(\varphi_i) = M.$$

Then

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \varphi_i = M \qquad \text{almost surely.}$$

Proof.
We may assume $E(\varphi_i) = 0$ and remove afterwards this restriction by applying the conclusions to the sequence of random variables $\varphi_i - E(\varphi_i)$. Put $S_n = \sum_{i=1}^{n} \varphi_i$. For

any $\varepsilon > 0$, by using the Čebišev inequality we have

$$
\begin{aligned}
P\left(\frac{1}{n^2}S_{n^2} \geq \varepsilon\right) &\leq \frac{1}{\varepsilon^2}Var\left(\frac{1}{n^2}S_{n^2}\right) = \frac{1}{n^4\varepsilon^2}E\left(S_{n^2}^2\right) \\
&= \frac{1}{n^4\varepsilon^2}E\left(\left(\sum_{i=1}^{n^2}\varphi_i\right)^2\right) \\
&= \frac{1}{n^4\varepsilon^2}E\left(\sum_{i=1}^{n^2}\varphi_i^2 + \sum_{i=1}^{n^2}\sum_{\substack{i\neq j \\ j=1}}^{n^2}\varphi_i\varphi_j\right) \\
&= \frac{1}{n^4\varepsilon^2}\left(\sum_{i=1}^{n^2}E\left(\varphi_i^2\right) + \sum_{i=1}^{n^2}\sum_{\substack{i\neq j \\ j=1}}^{n^2}E\left(\varphi_i\varphi_j\right)\right).
\end{aligned}
$$

Now, because of the negative correlation of the φ_i's we have that the terms $E\left(\varphi_i\varphi_j\right)$ for $i \neq j$ are not positive. So by using this fact and the bound for the variance, we get the estimate

$$
P\left(\frac{1}{n^2}S_{n^2} \geq \varepsilon\right) \leq \frac{1}{n^4\varepsilon^2}\left(\sum_{i=1}^{n^2}E\left(\varphi_i^2\right)\right) = \frac{1}{n^4\varepsilon^2}\left(\sum_{i=1}^{n^2}Var\left(\varphi_i\right)\right) \leq \frac{V}{n^2\varepsilon^2}.
$$

Since the series of the upper bounds is convergent, the series $\sum_{n=1}^{\infty}P\left(\frac{1}{n^2}S_{n^2} \geq \varepsilon\right)$ is convergent, too. Therefore by the Borel-Cantelli lemma, we have that

$$
\lim_{n\to\infty}\frac{1}{n^2}S_{n^2} = 0 \quad \text{a.s..} \tag{1.9}
$$

Define now

$$
L_n = \max_{n^2 \leq j < (n+1)^2}\left|S_j - S_{n^2}\right|.
$$

For the same ε, the Čebišev inequality implies that

$$
\begin{aligned}
P\left(\frac{L_n}{n^2} \geq \varepsilon\right) \;\; &\leq \;\; \frac{1}{n^4\varepsilon^2}E\left(L_n^2\right) \leq \frac{1}{n^4\varepsilon^2}E\left(\sum_{j=n^2+1}^{(n+1)^2-1}|S_j - S_{n^2}|^2\right) \\[2mm]
&= \;\; \frac{1}{n^4\varepsilon^2}E\left(\sum_{j=n^2+1}^{(n+1)^2-1}\left(\sum_{i=n^2+1}^{j}\varphi_i\right)^2\right) \\[2mm]
&= \;\; \frac{1}{n^4\varepsilon^2}E\left(\sum_{j=n^2+1}^{(n+1)^2-1}\left(\sum_{i=n^2+1}^{j}\varphi_i^2 + \sum_{i=n^2+1}^{j}\sum_{\substack{h=n^2+1 \\ i\neq h}}^{j}\varphi_i\varphi_h\right)\right) \\[2mm]
&= \;\; \frac{1}{n^4\varepsilon^2}\sum_{j=n^2+1}^{(n+1)^2-1}\left(\sum_{i=n^2+1}^{j}E(\varphi_i^2) + \sum_{i=n^2+1}^{j}\sum_{\substack{h=n^2+1 \\ i\neq h}}^{j}E(\varphi_i\varphi_h)\right) \\[2mm]
&\leq \;\; \frac{1}{n^4\varepsilon^2}\sum_{j=n^2+1}^{(n+1)^2-1}\sum_{i=n^2+1}^{j}Var(\varphi_i) \\[2mm]
&\leq \;\; \frac{1}{n^4\varepsilon^2}\sum_{j=n^2+1}^{(n+1)^2-1}\sum_{i=n^2+1}^{(n+1)^2-1}Var(\varphi_i) \\[2mm]
&\leq \;\; \frac{V(2n-1)^2}{n^4\varepsilon^2}.
\end{aligned}
$$

Since the series of the upper bounds is convergent, the series $\sum\limits_{n=1}^{\infty} P\left(\frac{1}{n^2}L_n \geq \varepsilon\right)$ is convergent and therefore, again by the Borel-Cantelli lemma, we have

$$
\lim_{n\to\infty}\frac{1}{n^2}L_n = 0 \quad \text{a.s..} \tag{1.10}
$$

Since for any m with $n^2 \leq m < (n+1)^2$ we have

$$
\frac{|S_m|}{m} \leq \frac{1}{n^2}\left(|S_{n^2}| + L_n\right)
$$

the conclusion follows by (1.9) and (1.10).

$\hfill\square$

Let φ be the random variable taking with probability $\frac{1}{k}$ values in the sample space $W = \{w_i \in [0,1], 1 \leq i \leq k\}$ with $k \geq 2$. We assume that $w_{i-1} < w_i$ for $1 \leq i \leq k$. Denote by φ_i the value assumed by φ in the $i-$th draw from W without replacement. Fix $c \in\,]0,1[$ and let $\psi_i = \chi_{[0,c[}(\varphi_i)$. Then the following property holds.

Proposition 1.2.6.

The variances of the random variables ψ_i, $1 \leq i \leq k$, are bounded by $\frac{1}{4}$ and the ψ_i's are negatively correlated.

Proof.

The expectation of ψ_i is given by

$$E(\psi_i) = \frac{1}{k} \sum_{i=1}^{k} \chi_{[0,c[}(\omega_i),$$

so $E(\psi_i^2) = E(\psi_i)$. Then

$$Var(\psi_i) = E(\psi^2) - (E(\psi_i))^2 = E(\psi_i)(1 - E(\psi_i)).$$

Now, it is easy to see that $\frac{1}{4}$ is an upper bound for the right-hand side and so we have that

$$Var(\psi_i) \leq \frac{1}{4}.$$

Since all pairs of distinct ψ_i's have the same joint distribution, we may evaluate just the covariance of ψ_1 and ψ_2. Suppose that $w_i \in [0,c[$ if and only if $i \leq h$, with $0 < h < k$. Then

$$
\begin{aligned}
Cov(\psi_1, \psi_2) &= E(\psi_1\psi_2) - E(\psi_1)E(\psi_2) \\
&= \frac{1}{k(k-1)} \sum_{i=1}^{h} \sum_{\substack{j=1 \\ i \neq j}}^{h} \chi_{[0,c[}(w_i)\chi_{[0,c[}(w_j) - \left(\frac{1}{k} \sum_{i=1}^{h} \chi_{[0,c[}(w_i) \right)^2 \\
&= \frac{h(h-1)}{k(k-1)} - \frac{h^2}{k^2} = \frac{h(h-k)}{k^2(k-1)} < 0
\end{aligned}
$$

$\square$

Now, we are ready to introduce the result of Volčič, [69]. In the following, we consider the *sequential random reordering* of the points (t_i^n), defined as follows.

Definition 1.2.7.

If (π_n) is a u.d. sequence of partitions of $[0,1]$ with $\pi_n = \{[t_{i-1}^n, t_i^n] : 1 \leq i \leq k(n)\}$, the sequential random reordering *of the points t_i^n is a sequence (φ_m) made up of consecutive blocks of random variables. The n-th block consists of $k(n)$ random variables which have the same law and represent the drawing, without replacement, from the sample space $W_n = \left\{ t_1^n, \ldots, t_{k(n)}^n \right\}$ where each singleton has probability $\frac{1}{k(n)}$.*

Denote by T_n the set of all permutations on W_n, endowed with the natural probability P compatible with the uniform probability on W_n, i.e. $P(\tau_n) = \frac{1}{k(n)!}$ with $\tau_n \in T_n$.

Any sequential random reordering of (π_n) corresponds to a random selection of $\tau_n \in T_n$ for each $n \in \mathbb{N}$. The permutation $\tau_n \in T_n$ identifies the reordered $k(n)$-tuple of random variables φ_i with $K(n-1) \leq i \leq K(n)$, where $K(n) = \sum_{i=1}^{n} k(i)$. Therefore, the set of all sequential random reorderings can be endowed with the natural product probability on the space $T = \prod_{n=1}^{\infty} T_n$.

Theorem 1.2.8.

If (π_n) is a u.d. sequence of partitions of $[0,1]$, then the sequential random reordering of the points t_i^n defining them is almost surely a u.d. sequence of points in $[0,1]$.

Proof.

Let (φ_m) be the sequential random reordering of (π_n). First of all, note that if $0 < c < 1$ and φ_m belongs to the $n-$th block of $k(n)$ random variables, then

$$E(\chi_{[0,c[}(\varphi_m)) = \frac{1}{k(n)} \sum_{i=1}^{k(n)} \chi_{[0,c[}(t_i^n)$$

and this quantity tends to c, when m and hence n tends to infinity, since (π_n) is u.d. by assumption.

If we consider $\psi_m = \chi_{[0,c[}(\varphi_m)$ for $K(n-1) \leq m \leq K(n)$, then Proposition 1.2.6 holds and so the ψ_m's are negatively correlated for $K(n-1) \leq m \leq K(n)$, i.e. when the φ_m belong to the same block. On the other hand, the correlation is zero when the φ_m belong to different blocks, since they are independent.

Let $\{c_h, h \in \mathbb{N}\}$ be a dense subset of $[0,1]$. Fix $h \in \mathbb{N}$ and consider the sequence $\left(\chi_{[0,c_h[}(\varphi_m)\right)$. Hence, we may apply the Lemma 1.2.5 and get that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \chi_{[0,c_h[}(\varphi_i) = c_h \qquad \text{a.s.}$$

for any c_h. But this is a sufficient condition for the uniform distribution and so we have our conclusion.

$\square$

1.3 Uniform distribution theory on $[0,1]^d$

In this section we deal with the extension of uniform distribution theory to the unit hypercube. We will introduce the basic definitions and results of the theory with

a particular attention to the study of discrepancy and to some special u.d. sequences of points.

1.3.1   Definitions and basic properties

Let d an integer with $d \geq 2$. Let $J = [a_1, b_1[\times \cdots \times [a_d, b_d[\subset \mathbb{R}^d$ be a rectangle with sides parallel to the axes in the $d-$dimensional space $\mathbb{R}^d$. If we denote by λ_d the $d-$dimensional Lebesgue measure, then the volume of J is given by

$$\lambda_d(J) = \prod_{i=1}^{d}(b_i - a_i)$$

. Let us denote by I^d the $d-$dimensional unit hypercube, i.e. $I^d = [0,1]^d$.

Definition 1.3.1.

A sequence (x_n) of points in I^d is said to be uniformly distributed (u.d.) *if for any rectangle R of the form $R = [0, a_1[\times \cdots \times [0, a_d[\subset I^d$ we have*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \chi_R(x_n) = \lambda_d(R) \tag{1.11}$$

where χ_R is the characteristic function of the rectangle R.

As in the one-dimensional case we can introduce the concept of *determining class* of functions.

Definition 1.3.2.

A class $\mathcal{F}$ of Riemann-integrable functions on I^d is said to be determining *for the uniform distribution of sequences of points, if for any sequence (x_n) in I^d the validity of the relation*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) = \int_{I^d} f \, d\lambda_d \tag{1.12}$$

for all $f \in \mathcal{F}$ already implies that (x_n) is u.d. .

Weyl was the first to extend to the multidimensional case the uniform distribution theory. So, we can give also in this case his classical results, [71, 72].

Theorem 1.3.3 (Weyl's Theorem).

A sequence (x_n) of points in I^d is u.d. if and only if for any (real or complex-valued) continuous function f defined on I^d the equation (1.12) holds.

Moreover, let $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ be in $\mathbb{R}^d$ and let us denote by $x \cdot y$ the usual inner product in $\mathbb{R}^d$, i.e. $x \cdot y = \sum_{i=1}^{d} x_i y_i$. Then we can give the generalization of the Weyl's Criterion.

Theorem 1.3.4 (Weyl's Criterion).
The sequence (x_n) in I^d is u.d. if and only if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h \cdot x_n} = 0$$

for all non zero integer lattice points $h \in \mathbb{Z}^d - \{(0, \ldots, 0)\}$.

Weyl applied this theorem to the Kronecker sequence also in the multidimensional case for giving a new proof of the Kronecker's approximation theorem in $\mathbb{R}^d$, [72].

Theorem 1.3.5 (Kronecker's Approximation Theorem).
Let $\theta = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$ such that $1, \theta_1, \ldots, \theta_d$ are linearly independent over the rationals. Then the sequence of fractionals parts $(\{n\theta\})$, where $\{n\theta\} = (\{n\theta_1\}, \ldots, \{n\theta_d\})$, is dense in I^d.

Furthermore, Weyl's criterion implies that a sequence of the form $(n\theta)$ is u.d. if and only if $1, \theta_1, \ldots, \theta_d$ are linearly independent over $\mathbb{Q}$. Hence it follows that $(n\theta)$ is u.d. if and only if $(\{n\theta\})$ is dense in I^d.

1.3.2   Estimation of discrepancy

Definitions 1.1.7 and 1.1.9 may be extended to sequences of points in I^d as follows.

Definition 1.3.6.
Let $\omega_N = \{x_1, \ldots, x_N\}$ be a finite set of points in I^d.

- *The discrepancy of ω_N is defined by*

$$D_N(\omega_N) = \sup_J \left| \frac{1}{N} \sum_{i=1}^{N} \chi_J(x_i) - \lambda_d(J) \right|,$$

  where J runs through all rectangles in I^d of the form $J = [a_1, b_1[\times \cdots \times [a_d, b_d[$ with $0 \le a_i < b_i \le 1$.

- *The star discrepancy of ω_N is defined by*

$$D_N^*(\omega_N) = \sup_R \left| \frac{1}{N} \sum_{i=1}^{N} \chi_R(x_i) - \lambda_d(R) \right|,$$

where R runs through all rectangles in I^d of the form $R = [0, a_1[\times \cdots \times [0, a_d[$ with $0 < a_i \leq 1$.

Moreover, the discrepancy and star discrepancy are related by the following inequality.

Theorem 1.3.7.
For any sequence (x_n) of points in I^d we have

$$D_N^*(x_n) \leq D_N(x_n) \leq 2^d D_N^*(x_n).$$

In the same way as in the one dimensional case if (x_n) is an infinite sequence of points, we associate to it the sequence of positive real numbers $D_N(\{x_1, x_2, \ldots x_N\})$. So, the symbol $D_N(x_n)$ denotes the discrepancy of the initial segment $\{x_1, x_2, \ldots x_N\}$ of the infinite sequence. It is easy to see that

Theorem 1.3.8.
A sequence (x_n) of points in I^d is u.d. if and only if

$$\lim_{n \to \infty} D_N(x_n) = 0.$$

Equivalently a sequence (x_n) of points in I^d is u.d. if and only if

$$\lim_{n \to \infty} D_N^*(x_n) = 0.$$

The immediate lower bound given in Proposition 1.1.11 holds also in the higher-dimensional case. In fact, we have

Proposition 1.3.9.
For any finite set $\omega = \{x_1, \ldots, x_N\}$ of points in I^d we have that

$$\frac{1}{N} \leq D_N(\omega) \leq 1.$$

Proof.
The right-hand side inequality is evident from the definition of discrepancy. Now, choose $\varepsilon > 0$ and consider the first point of ω, namely $x_1 = \left(x_1^{(1)}, \ldots, x_1^{(d)} \right) \in I^d$. Let $J = [x_1^{(1)}, x_1^{(1)} + \varepsilon[\times \cdots \times [x_1^{(d)}, x_1^{(d)} + \varepsilon[$. Since $x_1 \in J$ then we have

$$D_N(\omega) \geq \frac{1}{N} - \lambda_d(J) = \frac{1}{N} - \varepsilon^d$$

and so the conclusion follows.

$\square$

As we have already said, only in the one-dimensional case we have examples of sequences such that $D_N(x_n) = \frac{1}{N}$. In fact, in the higher-dimensional case such examples cannot exist by Roths's theorem, [56]. So far this is the best known result for $d > 3$.

Theorem 1.3.10 (Roth's Theorem).
Let $d \geq 2$. Then the discrepancy $D_N(x_n)$ of the finite set $\omega = \{x_1, \ldots, x_N\} \subset I^d$ is bounded from below by

$$D_N(\omega) \geq c_d \frac{(\log N)^{\frac{d-1}{2}}}{N},$$

where $c_d > 0$ is an absolute constant given by $c_d = \frac{1}{2^{4d}((d-1)\log 2)^{\frac{d-1}{2}}}$.

For further information on bounds for the dimensions 2 and 3 and refinements of Roth's theorem we refer to [21].

A well known conjecture states that for every dimension d there exists a constant c_d such that for any infinite sequence $(x_n) \in \mathbb{R}^d$ with $d \geq 1$ we have

$$D_N(x_n) \geq c_d \frac{(\log N)^d}{N}$$

for infinitely many N. This conjecture has been proved by Schmidt only for $d = 1$ (see Theorem 1.1.12), while it is still open for $d \geq 2$.

Usually, sequences of points in $\mathbb{R}^d$ having discrepancy bounded from above by $\mathcal{O}\left(\frac{(\log N)^d}{N}\right)$ are called *low discrepancy sequences*. We have already described an important class of low discrepancy sequences in the one-dimensional case, that is the van der Corput sequences. In the following, we will introduce their higher-dimensional generalization. Before defining these special u.d. sequences, let us give a result that proves the important role played by low discrepancy sequences in numerical integration.

1.3.3 The Koksma-Hlawka inequality

The concept of discrepancy gives a quantitative measure of the order of convergence in the relation (1.11) defining the uniform distribution of a given sequence. Consequently, it is also very interesting to get information on the order of convergence in (1.12). Referring to this problem, a very useful estimate is provided by the Koksma-Hlawka inequality. In fact, it states that the order of convergence between the actual value of the integral in (1.12) and its approximation can be estimated in terms of the variation of the function and the star discrepancy. Before we can write down this result, we need to define the variation of a function $f : I^d \to \mathbb{R}$.

By a partition P of I^d we mean a set of d finite sequences $(\eta_i^{(0)}, \ldots, \eta_i^{(m_i)})$ for $i = 1, \ldots, d$ with $0 = \eta_i^{(0)} \leq \eta_i^{(1)} \leq \cdots \leq \eta_i^{(m_i)} = 1$. In connection with such a partition we define for each $i = 1, \ldots, d$ an operator Δ_i by

$$
\begin{aligned}
\Delta_i f(x_1, \ldots, x_{i-1}, \eta_i^{(j)}, x_{i+1}, \ldots, x_d) &= f(x_1, \ldots, x_{i-1}, \eta_i^{(j+1)}, x_{i+1}, \ldots, x_d) \\
&- f(x_1, \ldots, x_{i-1}, \eta_i^{(j)}, x_{i+1}, \ldots, x_d)
\end{aligned}
$$

for $0 \leq j < m_i$. Operators with different subscrites obviously commute and $\Delta_{i_1, \ldots, i_k}$ stands for $\Delta_{i_1} \cdots \Delta_{i_k}$. Such an operator commutes with summation over variables on which it does not act.

Definition 1.3.11 (Function of bounded variation in the sense of Vitali).
For a function $f : I^d \to \mathbb{R}$ we set

$$
V^{(d)}(f) = \sup_P \sum_{j_1=0}^{m_1-1} \cdots \sum_{j_d=0}^{m_d-1} \left| \Delta_{1,\ldots,d} f(\eta_1^{(j_1)}, \ldots, \eta_d^{(j_d)}) \right|,
$$

where the supremum is extended over all partitions P of I^d.
If $V^{(d)}(f)$ is finite the f is said to be of bounded variation on I^d in the sense of Vitali.

Definition 1.3.12 (Function of bounded variation in the sense of Hardy and Krause.).
Let $f : I^d \to \mathbb{R}$ and assume that f is of bounded variation in the sense of Vitali. If the restriction $f^{(F)}$ of f to each face F of I^d of dimension $1, 2, \ldots, d-1$ is of bounded variation on F in the sense of Vitali, then f is said to be of bounded variation on I^d in the sense of Hardy and Krause.

So we can state

Theorem 1.3.13 (Koksma-Hlawka's Inequality).
Let f be a function of bounded variation on I^d in the sense of Hardy and Krause. Let $\omega = (x_1, \ldots, x_N)$ be a finite set of points in I^d. Let us denote by ω_l the projection of ω on the $(d-l)-$dimensional face F_l of I^d defined by $F_l = \{(u_1, \ldots, u_d) \in I^d : u_{i_1} = \cdots \cdots = u_{i_l} = 1$. Then we have

$$
\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_{I^d} f(x)dx \right| \leq \sum_{l=0}^{d-1} \sum_{F_l} D_N^*(\omega_l) V^{(d-l)}(f^{(F_l)}), \qquad (1.13)
$$

where the second sum is extended over all $(d-l)-$dimensional faces F_l of the form $u_{i_1} = \cdots = u_{i_l} = 1$. The discrepancy $D_N^*(\omega_l)$ is clearly computed in the face of I^d in which ω_l is contained.

Remark 1.3.14.

Trivially $D_N^*(\omega_l)$ can be bounded by $D_N^*(\omega)$. Hence we get from (1.13) that

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_{I^d} f(x)dx \right| \leq V(f) D_N^*(x_n) \qquad (1.14)$$

where

$$V(f) = \sum_{l=0}^{d-1} \sum_{F_l} V^{(d-l)}(f^{(F_l)})$$

is called the variation of Hardy and Krause.

A proof can be found in [45], but the original proof is given in [37]. This relation provides a strong motivation for the choice of low discrepancy sequences in Quasi-Monte Carlo integration.

1.3.4 The Halton and Hammersley sequences

A very important application of u.d. sequences is numerical integration. In fact, given a function f on I^d, the basic idea of classical Monte Carlo integration is to approximate the integral

$$I(f) = \int_{I^d} f d\lambda_d$$

with the mean

$$I_N = \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

where $x_1, \ldots, x_N$ are N points choosen randomly or pseudorandomly in I^d.

For a large class of functions, Quasi-Monte Carlo methods have a faster rate of convergence than the Monte Carlo method. Indeed, the Quasi-Monte Carlo methods works by choosing deterministically the N integration points instead of actual random points. Therefore, it is essential that the nodes are well distributed on I^d. This means that it is convenient if their distribution is close to the uniform distribution. So a good choice for the integration points is the initial segment of a sequence with small discrepancy, since the Koksma-Hlawka inequality holds, i.e

$$|I_N(f) - I(F)| \leq V(f) D_N^*(f),$$

where $V(f)$ is the variation of f in the sense of Hardy-Krause (see Subsection 1.3.3).

Finally, the deterministic nature of Quasi-Monte Carlo methods provides many advantages with respect to Monte Carlo methods. First of all, the Quasi-Monte Carlo

method allows to work with deterministic points rather than random samples and then it offers the availability of deterministic error bounds instead of the probabilistic Monte Carlo rate of convergence. Moreover, with the same computational effort, the Quasi-monte Carlo method achieves a significantly higher accuracy than the Monte Carlo method just thanks to the choice of the integration points with small discrepancy.

In this subsection, we want to introduce some important classes of sequences of points in I^d with small discrepancy: the Halton sequences and the Hammersley sequences. Both constructions are based on the radical inverse function (see Definition 1.1.13).

Definition 1.3.15 (Halton sequence).
For a given dimension $d \geq 2$ the $d-$dimensional Halton sequence (x_n) in I^d is defined by

$$x_n = (\gamma_{b_1}(n), \ldots, \gamma_{b_d}(n))$$

where $b_1, \ldots, b_d$ are given coprime integers.

As it was shown in [28], the Halton sequence is a low discrepancy sequence. In fact, it has a discrepancy of order $\mathcal{O}\left(\frac{(\log N)^d}{N}\right)$.

For $d = 1$ we just get the van der Corput sequence (see Subsection 1.1.3). So, Halton construction is a generalization of van der Corput one to the higher-dimensional case.

For example, let us consider $b_1 = 2$ and $b_2 = 3$. By applying the Halton's construction we first have to generate the van der Corput sequence in base 2 that is $(\gamma_2(n))$, i.e.

$$\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \cdots$$

and then we have to generate the van der Corput sequence in base 3 that is $(\gamma_3(n))$, i.e.

$$\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \cdots$$

Finally, the Halton sequence $(\gamma_2(n), \gamma_3(n))$ in the unit square I^2 is obtained by pairing up these two sequences

$$\left(\frac{1}{2}, \frac{1}{3}\right), \left(\frac{1}{4}, \frac{2}{3}\right), \left(\frac{3}{4}, \frac{1}{9}\right), \left(\frac{1}{8}, \frac{4}{9}\right), \left(\frac{5}{8}, \frac{7}{9}\right), \left(\frac{3}{8}, \frac{2}{9}\right), \left(\frac{7}{8}, \frac{5}{9}\right), \cdots$$

While the performance of standard Halton sequences is very good in low dimensions, problems with correlation have been observed among sequences generated from higher primes. This can cause serious problems in the estimation of models

with high-dimensional integrals. In order to deal with this problem, various other methods have been proposed; one of the most prominent solutions is the technique of *scrambled Halton sequence*, which uses permutations of the coefficients employed in the construction of the standard sequences, [59, 6, 50].

Definition 1.3.16 (Hammersley sequence).
For given integers $d \geq 2$ and N, the $d-$dimensional Hammersley sequence (x_n) of size N in I^d is defined by

$$x_n = \left(\frac{n}{N}, \gamma_{b_1}(n), \ldots, \gamma_{b_{d-1}}(n) \right)$$

where $b_1, \ldots, b_{d-1}$ are given coprime integers.

As it was shown in [29], the Hammersley sequence has a discrepancy of order $\mathcal{O}\left(\frac{(\log N)^{d-1}}{N} \right)$.

Note that the Hammersley sequence is a finite set of size N which cannot be extended to an infinite sequence. So in the approximation of the integral $I(f)$, one should have to decide in advance the value of N in order to perform the calculation, since the first coordinate depends on N. In the computational practice of Quasi-Monte Carlo integration it is often convenient to be able to increase the value of N without losing the previously calculated function values. For this purpose, it is preferable to work with a whole low discrepancy sequence of nodes and then take its first N terms whenever a value of N has been selected. In this way, N can be increased while all data from the earlier computations can be still used. Therefore, in several cases the Halton sequences are more convenient in Quasi-Monte Carlo integration than the Hammersley point sets.

1.4 Uniform distribution theory in compact spaces

A theory of uniform distribution can be developed in settings more abstract than the unit interval and the unite hypercube. In this section, we present its generalization to compact Hausdorff spaces. The study of this theory was intiated by Hlawka in [35, 36]. The notion of u.d. sequences in such spaces is related to a given nonnegative regular normalized Borel measure, but for convenience we will consider a regular probability.

Let X be a compact Hausdorff space and let us denote by $\mathcal{B}$ the σ-algebra of Borel subsets of X. Suppose μ is a regular probability on $\mathcal{B}$.

Definition 1.4.1 (Regular Borel measure)**.**
A positive measure μ defined on $\mathcal{B}$ is said to be regular *if*

$$\mu(E) = \sup\{\mu(C)\colon C \subseteq E, \ C \ closed\} = \inf\{\mu(D)\colon E \subseteq D, \ D \ open\}$$

for all $E \in \mathcal{B}$.

Notation 1.4.2. *Let us denote by*

- $\mathcal{B}(X)$ *the set of all bounded real valued measurable functions defined on X;*

- $\mathcal{C}(X)$ *the subset of $\mathcal{B}(X)$ consisting of all continuous real valued functions defined on X.*

The space $\mathcal{B}(X)$, endowed with the norm $\|f\|_\infty = \sup\limits_{x \in X} |f(x)|$, is a Banach space.
Among the various characterizations of the concept of u.d. sequences of points in $[0,1]$, the most easily adaptable to this more general situation is the Weyl's Theorem (that is Theorem 1.1.3), which allows to give the following definition.

Definition 1.4.3 (U.d. sequences of points)**.**
A sequence (x_i) of elements in X is said to be uniformly distributed (u.d.) *with respect to μ, if*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} f(x_i) = \int_X f(t) \, d\mu(t)$$

for all $f \in \mathcal{C}(X)$.

In order to generalize to compact spaces the concepts of u.d. sequences of partitions, we need to introduce the notion of $\mu-$continuity set.

Definition 1.4.4 (μ-continuity set)**.**
A Borel set $M \subset X$ is called a μ-continuity set if $\mu(\partial M) = 0$, where ∂M denotes the boundary of M with respect to the relative topology on X.

Definition 1.4.5 (U.d. sequences of partitions)**.**
Let (π_n) be a sequence of partitions of X, where $\pi_n = \left\{ A_1^n, A_2^n, \ldots, A_{k(n)}^n \right\}$ and the A_i^n's are μ-continuity sets. The sequence (π_n) is said to be μ-uniformly distributed if for any $f \in \mathcal{C}(X)$, and any choice $t_i^n \in A_i^n$ we have

$$\lim_{n \to \infty} \frac{1}{k(n)} \sum_{i=1}^{k(n)} f(t_i^n) = \int_X f(t) \, d\mu(t).$$

The existence of u.d. sequences of partitions in separable metric spaces has been addressed, but not completely solved, in [14]. On the other hand, the existence problem for u.d. sequences of points can be easily settled in compact Hausdorff spaces satisfying the second axiom of countability. In fact, if X is a compact Hausdorff with countable basis, then almost all sequences in X are u.d. with respect to μ, [35, 45]. Nevertheless, the existence problem is still open in the general setting of compact Hausdorff spaces. The strongest constructive result is due to Hedrlín who showed that u.d. sequences of points exist in every compact metric space, by using an explicit construction, [30]. Interesting results on this topic are proposed in [51].

Definition 1.4.6 (Determining functions for sequences of points).
A class $\mathcal{F}$ of Riemann-integrable functions is said to be determining for the uniform distribution of sequences of points if for any sequence (x_i) in X the validity of the relation

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} f(x_i) = \int_X f(t) \, d\mu(t) \, , \tag{1.15}$$

for all $f \in \mathcal{F}$ already implies that (x_i) is u.d..

Similarly, we can give the analogous definition for sequences of partitions.

Definition 1.4.7 (Determining functions for sequences of partitions).
A class $\mathcal{F}$ of Riemann-integrable functions is said to be determining for the uniform distribution of sequences of partitions if for any sequence (π_n), where $\pi_n = \left\{ A_1^n, A_2^n, \ldots, A_{k(n)}^n \right\}$, the validity of the relation

$$\lim_{n \to \infty} \frac{1}{k(n)} \sum_{i=1}^{k(n)} f(t_i^n) = \int_X f(t) \, dt$$

for all $f \in \mathcal{F}$ and for any choice $t_i^n \in A_i^n$ already implies that (π_n) is u.d..

Observe that the determining classes for the sequences of points play the same role for the sequences of partitions and viceversa.

As in uniform distribution theory on $[0,1]$, a family of $\mu-$continuity set $\mathcal{G}$ such that the class $\mathcal{F} = \{\chi_M, M \in \mathcal{G}\}$ is determining is called *discrepancy system*. Obviously, it is possible to define this notion independently of the concept of the determining class (see [21]).

Definition 1.4.8 (Discrepancy system)**.**

A system $\mathcal{G}$ of $\mu-$continuity sets of X is called discrepancy system *if*

$$\lim_{N \to \infty} \sup_{M \in \mathcal{G}} \left| \frac{1}{N} \sum_{n=1}^{N} \chi_M(x_n) - \mu(M) \right| = 0$$

holds if and only if the sequence (x_n) is u.d. with respect to μ.

For a family of real valued functions $\mathcal{F}$, we will denote by $span(\mathcal{F})$ the linear space generated by $\mathcal{F}$ and by $\overline{span(\mathcal{F})}$ its closure. The construction of many important determining classes is based on the following theorem.

Theorem 1.4.9.

Let (x_n) a sequence of points in X. If $\mathcal{F}$ is a class of functions from $\mathcal{B}(X)$ such that (1.15) holds for all $f \in \mathcal{F}$ and $\overline{span(\mathcal{F})} \supset \mathcal{C}(X)$, then $\mathcal{F}$ is a determining class for (x_n).

Proof.

Let us first show that (1.15) holds for all $g \in \text{span}(\mathcal{F})$. In fact any $g \in \text{span}(\mathcal{F})$ is of the form $g = \alpha_1 f_1 + \ldots + \alpha_k f_k$ with $f_i \in \mathcal{F}$ and $\alpha_i \in \mathbb{R}$, $1 \leq i \leq k$. Since (1.15) holds for all $f \in \mathcal{F}$, in particular holds for all f_i. Therefore by linearity, the function g satisfies the relation (1.15).

Now, let us consider $f \in \mathcal{C}(X)$. Fixed $\varepsilon > 0$, by the assumption of density there exists $h \in \text{span}(\mathcal{F})$ such that $\|f - h\|_\infty < \varepsilon$. Then we have

$$
\begin{aligned}
\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_X f \, d\mu \right| &\leq \left| \frac{1}{N} \sum_{n=1}^{N} (f - h)(x_n) - \int_X (f - h) \, d\mu \right| \\
&+ \left| \frac{1}{N} \sum_{n=1}^{N} h(x_n) - \int_X h \, d\mu \right| \\
&\leq \frac{1}{N} \sum_{n=1}^{N} |(f - h)(x_n)| + \int_X |f - h| \, d\mu \\
&+ \left| \frac{1}{N} \sum_{n=1}^{N} h(x_n) - \int_X h \, d\mu \right| \\
&\leq 2\|f - h\|_\infty + \left| \frac{1}{N} \sum_{n=1}^{N} h(x_n) - \int_X h \, d\mu \right| \\
&< 3\varepsilon
\end{aligned}
$$

for sufficiently large N.

$\square$

Now, we can generalize to the compact spaces the concept of discrepancy.

Definition 1.4.10 ($\mathcal{G}-$discrepancy)**.**
Let $\mathcal{G}$ be a discrepancy system in X and $\omega_N = \{x_1, \ldots, x_N\}$ a finite sequence of points in X. Then the discrepancy with respect to $\mathcal{G}$ (or $\mathcal{G}-$discrepancy) is defined by

$$D_N^{\mathcal{G}}(\omega_N) = \sup_{M \in \mathcal{G}} \left| \frac{1}{N} \sum_{n=1}^{N} \chi_M(x_n) - \mu(M) \right|.$$

If (x_n) is an infinite sequence of points in X, we associate to it the sequence of positive real numbers $D_N^{\mathcal{G}}(\{x_1, x_2, \ldots x_N\})$. Often it is used the symbol $D_N^{\mathcal{G}}(x_n)$ to denote the quantity $D_N^{\mathcal{G}}(\{x_1, x_2, \ldots x_N\})$.

It follows from the definition that (x_n) is u.d. if and only if $D_N^{\mathcal{G}}(x_n)$ tends to zero when N tends to infinity.

Chapter 2

Uniform distribution on fractals

In this chapter, we will be concerned with uniform distribution theory on a special class of fractals, namely those which are defined by an Iterated Function System (IFS) of similarities having the same ratio and satisfying the Open Set Condition (OSC). More precisely, we will give an explict procedure to generate u.d. sequences of partitions and of points on this class of fractals and we will present some results obtained about the elementary discrepancy of these sequences, [40].

2.1   Fractals defined by Iterated Function Systems

Let us introduce a general method of construction for some fractals which is based on their self-similarity, that is the property of many fractals to be made up of parts similar to the whole. For instance, the Cantor set is given by the union of two similar copies of itself and the von Koch curve consists of four similar copies. This property may actually be used to define these fractals, which are called IFS fractals because they are generated by an Iterated Function System. Before introducing this kind of construction, let us recall some basic notions (see [23, 22]).

Let $\| \cdot \|$ be the usual norm on the $d-$dimensional Euclidean space $\mathbb{R}^d$. By the diameter of a set $U \subset \mathbb{R}^d$ we mean the quantity $\mathrm{diam}(U) = \sup_{x,y \in U} \|x - y\|$.

Definition 2.1.1 ($\delta-$covering).
Let $E \subset \mathbb{R}^d$. Fixed $\delta > 0$, a countable family (U_i) of sets of $\mathbb{R}^d$ is said to be a $\delta-$covering of E if the union of all U_i's covers the set E and for each i we have $0 < \mathrm{diam}(U_i) \leq \delta$.

Definition 2.1.2 ($s-$dimensional Hausdorff measure).
For $E \subset \mathbb{R}^d$ and $s \geq 0$, we define for each $\delta > 0$

$$\mathcal{H}^s_\delta(E) = \inf \left\{ \sum_{i=1}^{\infty} \mathrm{diam}(U_i)^s : \ (U_i) \ is \ a \ \delta - covering \ of \ E \right\}$$

and subsequently the $s-$dimensional Hausdorff measure of E is given by

$$\mathcal{H}^s(E) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(E).$$

$\mathcal{H}^s$ is not a measure in the usual sense, but it is an outer measure. In fact, in general the countable additivity does not hold but it is possible to prove that $\mathcal{H}^s$ is a measure only when it is defined over the Borel sets of $\mathbb{R}^d$ (see [23]). The Haudorff measure generalizes the concept of the Lebesgue measure on $\mathbb{R}^d$. Indeed, $\mathcal{H}^d$ is equal to the $d-$dimensional Lebesgue measure λ_d up to a constant, i.e.

$$\mathcal{H}^d(E) = c_d \lambda_d(E)$$

where $c_d = \frac{\pi^{\frac{1}{2}d}}{2^d \left(\frac{1}{2}d\right)!}$. Moreover, the Hausdorff measure has a very useful scaling property.

Proposition 2.1.3.
Let $E \subset \mathbb{R}^d$, $k > 0$ and $s \geq 0$ then

$$\mathcal{H}^s(kE) = k^s \mathcal{H}^s(E)$$

where $kF = \{kx \colon x \in E\}$, i.e. the set kE is the set E scaled of a factor k.

$\mathcal{H}^s(E)$ is non-increasing with s and there exists a unique value of s where $\mathcal{H}^s(E)$ jumps from ∞ to 0. This value is called *Hausdorff dimension* of E and it is given by

$$\dim_H(E) = \inf\{s \colon \mathcal{H}^s(E) = 0\} = \sup\{s \colon \mathcal{H}^s(E) = \infty\}.$$

Let us denote by $\mathcal{K}(\mathbb{R}^d)$ the space of all the non-empty compact subsets of $\mathbb{R}^d$ endowed with the Hausdorff distance, which makes it a complete metric space. Let us recall the definition of *Hausdorff distance*

$$d_H(K, L) = \max \left\{ \sup_{x \in K} \inf_{y \in L} \|x - y\|, \sup_{x \in L} \inf_{y \in K} \|x - y\| \right\}$$

or equivalently

$$d_H(K, L) = \min \left\{ \lambda \geq 0 \colon K \subset L_\lambda \text{ and } L \subset K_\lambda \right\}$$

where

$$K_\lambda = \{x \in \mathbb{R}^n \colon \|x - y\| \le \lambda \text{ for some } y \in K\}$$

and

$$L_\lambda = \{x \in \mathbb{R}^n \colon \|x - y\| \le \lambda \text{ for some } y \in L\}.$$

Let us give the following results, due to Hutchinson which show how an IFS defines a unique non-empty self-similar compact set, [39].

Theorem 2.1.4.
Let $\psi_1, \ldots, \psi_m$ be m contractions defined on $\mathbb{R}^d$ so that $|\psi_i(x) - \psi_i(y)| \le c_i \|x - y\|$, for all $x, y \in \mathbb{R}^d$, with $0 < c_i < 1$ for each i. Then the mapping $\psi(E) \mapsto \bigcup_{i=1}^{m} \psi_i(E)$ is a contraction on $\mathcal{K}(\mathbb{R}^d)$ and its unique fixed point is a non-empty compact set F, called the attractor of the IFS. The set F is called a self-similar set and we have

$$F = \bigcup_{i=1}^{m} \psi_i(F).$$

Moreover, if $F_0 \in \mathcal{K}(\mathbb{R}^d)$ is such that $\psi_i(F_0) \subset F_0$ for $1 \le i \le m$, then the sequence of iterates $(\psi^n(F_0))$ is decreasing and convergent to F in the Hausdorff metric as $n \to \infty$, with

$$F = \bigcap_{n=0}^{\infty} \psi^n(F_0)$$

(where $\psi^0(F_0) = F_0$ and $\psi^{n+1}(F_0) = \psi(\psi^n(F_0))$ for $n \ge 0$).

The set $F_0 = \psi^0(F_0)$ is called *initial set* and the iterates $\psi^n(F_0)$ are called *pre-fractals* for F.

One advantages of dealing with fractals generated by an IFS is that their Hausdorff dimension is often easy to calculate. In particular, the evaluation of the Hausdorff dimension is very simple when we consider m similarities $\psi_1, \ldots, \psi_m$ on $\mathbb{R}^d$ with ratios $0 < c_i < 1$ for each i, i.e $|\psi_i(x) - \psi_i(y)| = c_i \|x - y\|$, for all $x, y \in \mathbb{R}^d$ and assume that the following condition holds.

Definition 2.1.5 (OSC).
A class of similarities $\psi_1, \ldots, \psi_m \colon \mathbb{R}^d \to \mathbb{R}^d$ satisfies the open set condition *if there exists a non-empty bounded open set V such that*

$$V \supset \bigcup_{i=1}^{m} \psi_i(V)$$

where $\psi_i(V)$ are pairwise disjoint.

Then for this special class of IFS fractals we have the following theorem due to Moran, (see [49], [22]).

Theorem 2.1.6.
Assume that m similarities $\psi_1, \ldots, \psi_m$ defined on $\mathbb{R}^d$ with ratios $0 < c_i < 1$ (for $1 \leq i \leq m$) satisfy the OSC. Let F be the attractor of the ψ_i's then the Hasudorff dimension s of F is the solution of the equation

$$\sum_{i=1}^{m} c_i^s = 1 \tag{2.1}$$

Moreover, we have that the $s-$dimensional Hausdorff measure $\mathcal{H}^s(F)$ is positive and finite, i.e. $0 < \mathcal{H}^s(F) < \infty$.

Note that, the OSC ensures that the components $\psi_i(F)$ of the invariant set F cannot overlap too much and this property is expressed by the following corollary of the previous theorem.

Corollary 2.1.7.
Let $\psi_1, \ldots, \psi_m$ be similarities on $\mathbb{R}^d$ with ratio $0 < c_i < 1$ for each i and let F be the attractor. If the OSC holds, then $\mathcal{H}^s(\psi_i(F) \cap \psi_j(F)) = 0$ for $i \neq j$.

Proof.
Using the assumption that the ψ_i's are similarities and Property 2.1.3 of the Hausdorff measure, we have

$$\sum_{i=1}^{m} \mathcal{H}^s\left(\psi_i(F)\right) = \sum_{i=1}^{m} c_i^s \mathcal{H}^s(F)$$

Since the OSC holds, we can use the relation (2.1) and so we have

$$\sum_{i=1}^{m} \mathcal{H}^s\left(\psi_i(F)\right) = \mathcal{H}^s(F) \sum_{i=1}^{m} c_i^s = \mathcal{H}^s(F) = \mathcal{H}^s\left(\bigcup_{i=1}^{m} \psi_i(F)\right).$$

By Theorem 2.1.6 we have $0 < \mathcal{H}^s(F) < \infty$, so the previous relation can only happen if $\mathcal{H}^s\left(\psi_i(F) \cap \psi_j(F)\right) = 0$ for $i \neq j$.

$\square$

Let us just cite some of the most popular examples of fractals which are included in the class considered by Theorem 2.1.6.

Examples 2.1.8.

Cantor set

The Cantor set is constructed starting from the unit interval through a sequence of deletion operations. Put $C_0 = [0, 1]$. At the first step we remove the open middle third of C_0 and so we obtain the set $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. Deleting the open middle third of the intervals $\left[0, \frac{1}{3}\right]$ and $\left[\frac{2}{3}, 1\right]$, we obtain four intervals of lenght $\frac{1}{9}$. So at the end of the second step we have constructed the set $C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$. By repeating this procedure, at the $k-$th step we have that C_k consists of 2^k intervals of lenght 3^{-k}, generated by removing the open middle third of each interval in C_{k-1}. In Figure 2.1 the first three steps of this construction are illustrated. The Cantor set C is given by
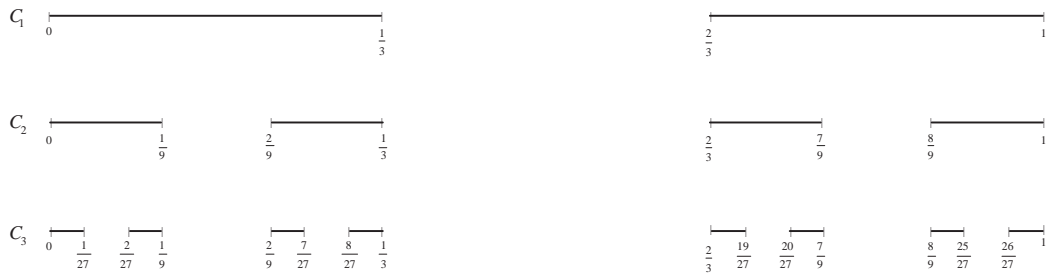
$$C = \bigcap_{k=0}^{\infty} C_k.$$

Figure 2.1: Construction of the Cantor set

So the Cantor set C is the attractor of the two following similarities

$$\begin{cases} S_1(x) = \frac{1}{3}x \\ \\ S_2(x) = \frac{1}{3}x + \frac{2}{3} \end{cases}.$$

The set C satisfies the OSC taking $V =]0, 1[$. Then by Theorem 2.1.6 we have that the Hausdorff dimension s of C is given by

$$\sum_{i=1}^{2} \left(\frac{1}{3}\right)^s = 1 \; \Rightarrow \; 2 = 3^s \; \Rightarrow \; s = \frac{\log 2}{\log 3}.$$

Sierpiński Triangle

The Sierpiński triangle T is constructed starting from an equilateral triangle by repeatedly removing inverted equilateral triangles. In fact, let T_0 be an equilateral triangle in $\mathbb{R}^2$ and take the three middle points of its sides. These three points and the vertices of T_0 define four equilateral congruent triangles and we remove the central open one. At the end of the first step, we have obtained three congruent closed triangles and we denote their union by T_1. At the second step, we repeat this procedure on each triangle of T_1, so we get nine triangles whose union is T_2. In figure 2.2 the first four steps of this construction are illustrated. The Sierpiński triangle T is given by
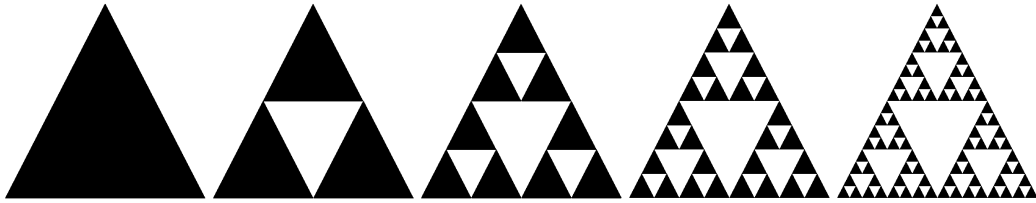
$$T = \bigcap_{k=0}^{\infty} T_k.$$

Figure 2.2: Construction of the Sierpiński triangle

So the Sierpiński triangle T is the attractor of the following similarities

$$\begin{cases} S_1(x,y) = \left(\frac{1}{2}x, \frac{1}{2}y \right) \\[2ex] S_2(x,y) = \left(\frac{1}{2}x + \frac{1}{4}, \frac{1}{2}y + \frac{\sqrt{3}}{4} \right) \\[2ex] S_3(x,y) = \left(\frac{1}{2}x + \frac{1}{2}, \frac{1}{2}y \right) \end{cases}$$

where the origin is taken in the left down vertex of the initial triangle. The set T satisfies the OSC taking V as the interior of the initial triangle T_0. Consequently, by Theorem 2.1.6, we have that the Hausdorff dimension s of T is given by

$$\sum_{i=1}^{3} \left(\frac{1}{2} \right)^s = 1 \;\Rightarrow\; 3 = 2^s \;\Rightarrow\; s = \frac{\log 3}{\log 2}.$$

von Koch Curve

The von Koch curve K is constructed starting from the unit interval $K_0 = [0, 1]$. At the first step we remove the open middle third of K_0 and replace it by the other two sides of the equilateral triangle based on the removed segment. The union of these four segments is denoted by K_1. We construct K_2 applying this procedure to each segment in K_1, and so on. The sequence of polygonal curves K_j tends to a limiting curve K, called von Koch curve. In figure 2.3 the first four steps of this construction are illustrated.
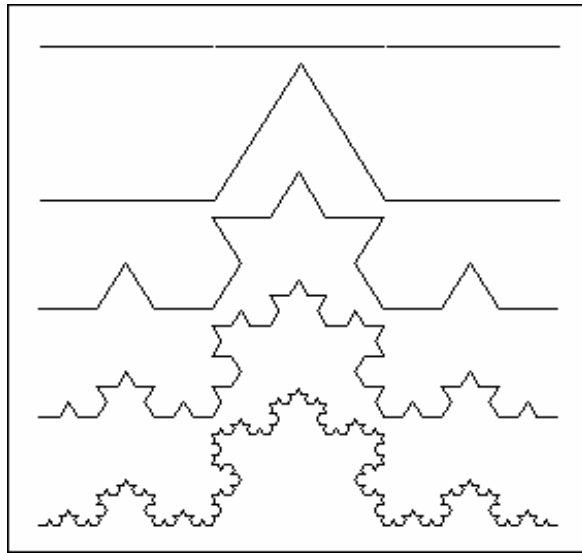

Figure 2.3: Construction of the von Koch curve

The Von Koch curve K is the attractor of the following similarities.

$$
\begin{cases}
S_1(x, y) = \left(\frac{1}{3}x, \frac{1}{3}y \right) \\[2ex]
S_2(x, y) = \left(\frac{1}{6}x - \frac{\sqrt{3}}{6}y + \frac{1}{3}, \frac{\sqrt{3}}{6}x + \frac{1}{6}y \right) \\[2ex]
S_3(x, y) = \left(-\frac{1}{6}x + \frac{\sqrt{3}}{6}y + \frac{2}{3}, \frac{\sqrt{3}}{6}x + \frac{1}{6}y \right) \\[2ex]
S_4(x, y) = \left(\frac{1}{3}x + \frac{2}{3}, \frac{1}{3}y \right)
\end{cases}
.
$$

The curve K satisfies the OSC taking V as the interior of the isosceles triangle of height equal to $\frac{\sqrt{3}}{6}$ with basis the segment K_0. So by Theorem 2.1.6 we have that the Hausdorff dimension s of K is given by

$$\sum_{i=1}^{4} \left(\frac{1}{3}\right)^s = 1 \ \Rightarrow \ 4 = 3^s \ \Rightarrow \ s = \frac{\log 4}{\log 3}$$

2.2 Van der Corput sequences on fractals

In this section, we extend to certain fractals the concept of u.d. sequences of partitions defined for the interval $[0, 1]$ in Subsection 1.2. In particular, we introduce our recent results in this setting, [40]. We present a general algorithm to produce u.d. sequences of partitions and of points on the class of fractals generated by a system of similarities on $\mathbb{R}^d$ having the same ratio and verifying the open set condition. Moreover, we provide an estimate for the elementary discrepancy of van der Corput type sequences constructed on this class of fractals.

2.2.1 Algorithm to construct u.d. sequence of points and of partition on a class of fractals

The classical concept of u.d. sequences of points is more natural when we deal with the interval $[0, 1]$ and with manifolds. On the other hand, when we work on fractals, in particular with fractals generated by Iterated Function Systems, partitions become a convenient tool for introducing a uniform distribution theory.

The advantage of considering partitions was implicitely used by Grabner and Tichy in [27] and by Cristea and Tichy in [18] even if they treated u.d. sequences of points. In these papers various concepts of discrepancy were introduced on the planar Sierpiński gasket and on the multidimensional Sierpiński carpet respectively, by using different kinds of partitions on these two fractals. In [27] an analogue of the classical van der Corput sequence has been constructed on the planar Sierpiński gasket. Similarly, in a succesive paper of Cristea, Pillichshammer, Pirsic and Scheicher [17] a sequence of van der Corput type has been defined on the s-dimensional Sierpiński carpet by exploiting the IFS-addresses of the carpet points. In all these papers the order of convergence of the several notions of discrepancy is determined for the van der Corput type sequences constructed on these fractals.

The idea to study this special kind of sequences in relation to uniform distribution on IFS fractals is also our starting point. In fact, the algorithm we are going to introduce generalizes the results cited above and allows to construct van der Corput

sequences on a whole class of fractals including the ones considered in [27], [18] and [17].

From now on, we consider m similarities $\psi_1, \ldots, \psi_m$ defined on $\mathbb{R}^d$ having all the same ratio, i.e. $|\psi_i(x) - \psi_i(y)| = c\|x - y\|$, for all $x, y \in \mathbb{R}^d$, with $0 < c < 1$. Moreover, we assume that our system of similarities satisfies the OSC. According to Theorem 2.1.6, the Hausdorff dimension of the attractor F of this IFS is $s = -\frac{\log m}{\log c}$ and its s-dimensional Hausdorff measure $\mathcal{H}^s$ is positive and finite.

Our class of fractals includes the most popular kind of fractals as for instance the ones cited in Examples 2.1.8. But also $[0, 1]$ can be seen as the attractor of an IFS, in fact of infinitely many IFS's. Indeed, fix a positive integer m and consider the mappings $\varphi_1, \ldots, \varphi_m$, from $\mathbb{R}$ to $\mathbb{R}$, where

$$\varphi_k(x) = \frac{k-1}{m} + \frac{x}{m} \text{ , for } 1 \le k \le m. \tag{2.2}$$

Then $[0, 1]$ is the attractor of this IFS. This observation goes back to Mandelbrot (see [48]) and suggests how to define on the kind of fractals we are considering (and also on $[0, 1]$) the van der Corput sequences of base $m > 1$.

Let $\psi = \{\psi_1, \ldots, \psi_m\}$ be our IFS and F its attractor. Assume that F_0 is the initial set such that

$$\psi_i(F_0) \subset F_0 \text{ for } i = 1, \ldots, m.$$

Fix a point $x_0 \in F$ and apply $\psi_1, \ldots, \psi_m$ in this order to x_0 getting so the points $x_1, \ldots, x_m$. At the second step, we apply the m mappings first to x_1, then to x_2 and so on, getting finally m^2 points ordered in a precise manner. Now we keep going, applying the functions of the IFS first to x_1, then to x_2 and continue so until we reach the point x_{m^2}, getting so m^3 points in the order determined by the construction. Iterating this procedure we get a sequence (x_n) of points in F which will be called the van der Corput sequence.

Example
Consider the triangle $T_0 \subset \mathbb{R}^2$ of vertices $(0,0)$, $(0,1)$ e $(1,1)$ and the similarities S_1, S_2 e S_3, defined as follows

$$\begin{cases} S_1(x,y) = \left(\frac{1}{2}x, \frac{1}{2}y\right) \\\\ S_2(x,y) = \left(\frac{1}{2}x, \frac{1}{2}y + \frac{1}{2}\right) \\\\ S_3(x,y) = \left(\frac{1}{2}x + \frac{1}{2}, \frac{1}{2}y + \frac{1}{2}\right) \end{cases} \quad .$$

Let T be the Sierpiński triangle generated by this IFS starting from the initial set T_0. Fixed $x_0 = (0,0)$ by applying the algortithm we have

I step
$$x_1 = (0,0) \quad x_2 = \left(0, \tfrac{1}{2}\right) \quad x_3 = \left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

II step
$$x_1 = (0,0) \quad x_2 = \left(0, \tfrac{1}{2}\right) \quad x_3 = \left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

$$x_4 = \left(0, \tfrac{1}{4}\right) \quad x_5 = \left(0, \tfrac{3}{4}\right) \quad x_6 = \left(\tfrac{1}{2}, \tfrac{3}{4}\right)$$

$$x_7 = \left(\tfrac{1}{4}, \tfrac{1}{4}\right) \quad x_8 = \left(\tfrac{1}{4}, \tfrac{3}{4}\right) \quad x_9 = \left(\tfrac{3}{4}, \tfrac{3}{4}\right)$$

III step
$$x_1 = (0,0) \quad x_2 = \left(0, \tfrac{1}{2}\right) \quad x_3 = \left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

$$x_4 = \left(0, \tfrac{1}{4}\right) \quad x_5 = \left(0, \tfrac{3}{4}\right) \quad x_6 = \left(\tfrac{1}{2}, \tfrac{3}{4}\right)$$

$$x_7 = \left(\tfrac{1}{4}, \tfrac{1}{4}\right) \quad x_8 = \left(\tfrac{1}{4}, \tfrac{3}{4}\right) \quad x_9 = \left(\tfrac{3}{4}, \tfrac{3}{4}\right)$$

$$x_{10} = \left(0, \tfrac{1}{8}\right) \quad x_{11} = \left(0, \tfrac{5}{8}\right) \quad x_{12} = \left(\tfrac{1}{2}, \tfrac{5}{8}\right)$$

$$x_{13} = \left(0, \tfrac{3}{8}\right) \quad x_{14} = \left(0, \tfrac{7}{8}\right) \quad x_{15} = \left(\tfrac{1}{2}, \tfrac{7}{8}\right)$$

$$x_{16} = \left(\tfrac{1}{4}, \tfrac{3}{8}\right) \quad x_{17} = \left(\tfrac{1}{4}, \tfrac{7}{8}\right) \quad x_{18} = \left(\tfrac{3}{4}, \tfrac{7}{8}\right)$$

$$x_{19} = \left(\tfrac{1}{8}, \tfrac{1}{8}\right) \quad x_{20} = \left(\tfrac{1}{8}, \tfrac{5}{8}\right) \quad x_{21} = \left(\tfrac{5}{8}, \tfrac{5}{8}\right)$$

$$x_{22} = \left(\tfrac{1}{8}, \tfrac{3}{8}\right) \quad x_{23} = \left(\tfrac{1}{8}, \tfrac{7}{8}\right) \quad x_{24} = \left(\tfrac{5}{8}, \tfrac{7}{8}\right)$$

$$x_{25} = \left(\tfrac{3}{8}, \tfrac{3}{8}\right) \quad x_{26} = \left(\tfrac{3}{8}, \tfrac{7}{8}\right) \quad x_{27} = \left(\tfrac{7}{8}, \tfrac{7}{8}\right)$$

and so on. In Figures 2.4, 2.5 and 2.6 show the first three steps of the algorithm.

Observe that if $[0,1]$ is seen as the attractor of the IFS described in (2.2) and if $x_0 = 0$, the sequence (x_n) generated by this procedure is exactly the classical van der Corput sequence (see Subsection 1.1.3).
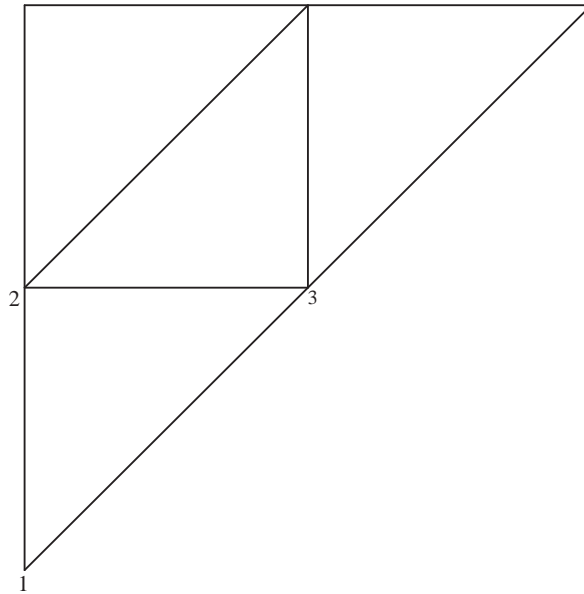
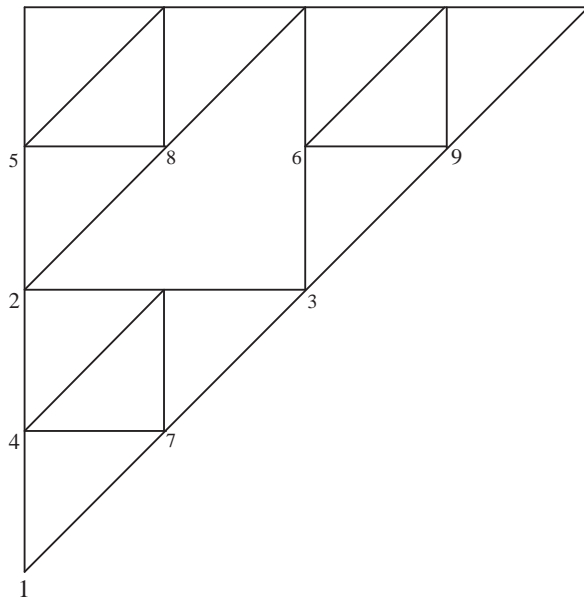Figure 2.4: Construction of a u.d. sequence of points on T (I step)

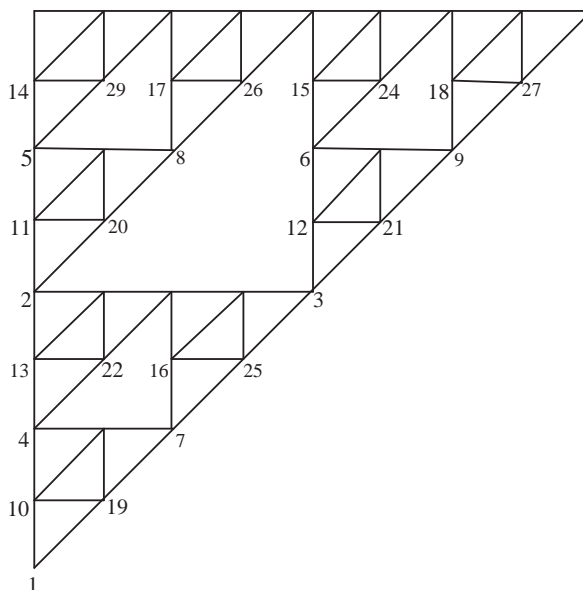Figure 2.5: Construction of a u.d. sequence of points on T (II step)

Figure 2.6: Construction of a u.d. sequence of points on T (III step)

Let us come back to the general situation, showing now how a similar construction produces sequences of u.d. partitions.

Note that if we apply the ψ_i's to F, in the same order as before, we construct a sequence (π_k) of partitions of F

$$\pi_k = \left\{ \psi_{j_k} \psi_{j_{k-1}} \cdots \psi_{j_1}(F) : j_1, \ldots, j_k \in \{1, \ldots, m\} \right\}.$$

Each of the m^k sets E_j^k of the partition π_k contains exactly one point of the van der Corput sequence (x_n) constructed above for $n = m^k$. We order the sets E_j^k accordingly.

Let us denote by $\mathscr{E}_k$ the collection of the m^k sets in π_k and by $\mathscr{E}$ the union of the families $\mathscr{E}_k$, for $k \in \mathbb{N}$. The sets of the class $\mathscr{E}$ are called *elementary sets*.

As probability on F we consider the normalized s-dimensional Hausdorff measure P, i.e.

$$P(A) = \frac{\mathcal{H}^s(A)}{\mathcal{H}^s(F)} \text{ for any Borel set } A \subset F \tag{2.3}$$

which is a regular probability.

The sequences of points and of partitions generated on F by the algorithm are both u.d. with respect to the probability P. But before proving these results, we need to introduce some preliminary lemmas about properties of the family of the elementary sets.

Lemma 2.2.1.

The elementary sets are P-continuity sets.

Proof.

Consider an elementary set $E_i = \psi_i(F) \in \mathscr{E}_1$. Let $x \in \partial E_i$. By definition, every neighbourhood U of x in the relative topology is such that $U \cap E_j \neq \emptyset$ for some $j \in \{1, 2, \ldots, m\}$ and $j \neq i$. But each E_j is closed, therefore $x \in E_j$. Hence ∂E_i is contained in $\bigcup\limits_{\substack{j=1 \\ j\neq i}}^{m} (E_i \cap E_j)$. By Theorem 2.1.7, we have

$$
\begin{aligned}
0 &\leq \mathcal{H}^s(\partial E_i) \leq \mathcal{H}^s\left(\bigcup_{\substack{j=1 \\ j\neq i}}^{m} \big(E_i \cap E_j\big)\right) = \mathcal{H}^s\left(\bigcup_{\substack{j=1 \\ j\neq i}}^{m} \big(\psi_i(F) \cap \psi_j(F)\big)\right) \\
&\leq \sum_{\substack{j=1 \\ j\neq i}}^{m} \mathcal{H}^s\big(\psi_i(F) \cap \psi_j(F)\big) = 0.
\end{aligned}
$$

Now, a generic elementary set $A \in \mathscr{E}_k$ with $k \geq 2$ is a homothetic image of an elementary set in $\mathscr{E}_1$ and therefore $\mathcal{H}^s(\partial A) = 0$, too.

$\square$

Lemma 2.2.2.

The class $\mathscr{E}$ consisting of all elementary sets is determining.

Proof.

Let $\mathcal{M}$ be the class consisting of all characteristic functions of sets $E \in \mathscr{E}$ and $f \in \mathcal{C}(F)$. By uniform continuity, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x') - f(x'')| < \varepsilon$ whenever $\|x' - x''\| < \delta$. Chose $n \in \mathbb{N}$ such that every $E_k^n \in \mathscr{E}_n$ has diameter smaller than δ. Take for any $E_k^n \in \mathscr{E}_n$ a point t_k and consider the function

$$
g(y) = \sum_{k=1}^{m^n} f(t_k)\chi_{E_k^n}(y), \quad y \in F.
$$

For some k we have $y \in E_k^n$ and so $|g(y) - f(y)| = |f(t_k) - f(y)| < \varepsilon$.

Hence, span$(\mathcal{M})$ is uniformly dense in $\mathcal{C}(F)$ and the conclusion follows by Theorem 1.4.9.

$\square$

Now, we are ready to introduce the main results of this section.

Theorem 2.2.3.

The sequence (π_n) of partitions of F generated by the algorithm is u.d. with respect to the probability P.

Proof.

By Lemma 2.2.2, we have to show that

$$\lim_{n\to\infty} \frac{1}{m^n} \sum_{j=1}^{m^n} \chi_{E_h^k}(t_j^n) = \int_F \chi_{E_h^k}(t)\, dP(t)$$

for every $E_h^k \in \pi_k$ and for every choice of $t_j^n \in E_j^n$. Let us fix $E_h^k \in \pi_k$. The previous relation is equivalent to

$$\lim_{n\to\infty} \frac{1}{m^n} \sum_{j=1}^{m^n} \chi_{E_h^k}(t_j^n) = \frac{1}{m^k}$$

because

$$\int_F \chi_{E_h^k}(t)\, dP(t) = P(E_h^k) = c^{sk}P(F) = c^{sk} = \frac{1}{m^k}.$$

Now, observe that for $n > k$, among the m^n sets generated by the algorithm, exactly one set of π_n is contained in the fixed set E_h^k. Since there are m^{n-k} sets of π_n which are contained in E_h^k, then

$$\lim_{n\to\infty} \frac{1}{m^n} \sum_{j=1}^{m^n} \chi_{E_h^k}(t_j^n) = \frac{m^{n-k}}{m^n} = \frac{1}{m^k}.$$

$\square$

Theorem 2.2.4.

The sequence (x_i) of points of F generated by the algorithm is u.d. with respect to P.

Proof.

By Lemma 2.2.2, the class $\mathscr{E}$ is determining. Hence, for a fixed set $E \in \mathscr{E}_k$, we have to prove that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \chi_E(x_i) = \int_F \chi_E\, dP = \frac{1}{m^k}. \tag{2.4}$$

Let $m^t \le N < m^{t+1}$, then

$$
\begin{aligned}
\frac{1}{N} \sum_{i=1}^{N} \chi_E(x_i) &= \frac{1}{N} \sum_{i=1}^{m+m^2+\ldots+m^{t-1}} \chi_E(x_i) + \frac{1}{N} \sum_{i=\frac{m^t-m}{m-1}}^{N} \chi_E(x_i) \\
&= \frac{\left(\frac{m^t-m}{m-1}\right)}{N} \cdot \frac{1}{\left(\frac{m^t-m}{m-1}\right)} \sum_{i=1}^{m+m^2+\ldots+m^{t-1}} \chi_E(x_i) \\
&+ \frac{N - \left(\frac{m^t-m}{m-1}\right)}{N} \cdot \frac{1}{N - \left(\frac{m^t-m}{m-1}\right)} \sum_{i=\frac{m^t-m}{m-1}}^{N} \chi_E(x_i) \quad (2.5)
\end{aligned}
$$

since $1 + m + m^2 + \ldots + m^{t-1} = \frac{m^t-1}{m-1}$.

Observe that for $i > \frac{m^t-1}{m-1}$, because of the order of the points x_i, among the first m^k points exactly one point of the sequence (x_i) is contained in the fixed set E. Hence, for $t \to \infty$ we have

$$
\frac{1}{\left(\frac{m^t-m}{m-1}\right)} \sum_{i=1}^{m+m^2+\ldots+m^{t-1}} \chi_E(x_i) \to \frac{1}{m^k}. \quad (2.6)
$$

Writing N as $N = \left(\frac{m^t-m}{m-1}\right) + Mm^k + r$ with $0 \le r < m^k$, we have

$$
\begin{aligned}
\frac{1}{N - \left(\frac{m^t-m}{m-1}\right)} \sum_{i=\frac{m^t-m}{m-1}}^{N} \chi_E(x_i) &= \frac{Mm^k}{N - \left(\frac{m^t-m}{m-1}\right)} \cdot \frac{1}{Mm^k} \sum_{i=\frac{m^t-m}{m-1}}^{N-r} \chi_E(x_i) \\
&+ \frac{r}{N - \left(\frac{m^t-m}{m-1}\right)} \cdot \frac{1}{r} \sum_{i=N-r+1}^{N} \chi_E(x_i). \quad (2.7)
\end{aligned}
$$

By the previous remarks we get that

$$
\frac{1}{Mm^k} \sum_{i=\frac{m^t-m}{m-1}}^{N-r} \chi_E(x_i) = \frac{1}{m^k},
$$

while for $N \to \infty$ and hence for $t \to \infty$ we have

$$
\frac{r}{N - \left(\frac{m^t-m}{m-1}\right)} \sum_{i=N-r+1}^{N} \chi_E(x_i) \to 0
$$

because $0 \le \frac{r}{N} < \frac{m^k}{m^t}$.

Using the last two relations in (2.7) and taking the limit for $N \to \infty$ (and hence for $t \to \infty$) we have

$$
\frac{1}{N - \left(\frac{m^t-m}{m-1}\right)} \sum_{i=1}^{N} \chi_E(x_i) \to \frac{1}{m^k}. \quad (2.8)
$$

Finally, (2.5) is a convex combination of two terms which both tend to $\frac{1}{m^k}$ for $N \to \infty$ because of (2.6) and (2.8). Therefore, the conclusion (2.4) holds.

$\square$

2.2.2 Order of convergence of the elementary discrepancy

In the following, we will provide an estimate for the elementary discrepancy of the sequences of van der Corput type generated by our explicit algorithm.

Note that Lemma 2.2.2 implies the family $\mathscr{E}$ of all elementary sets of F is a discrepancy system (see Definition 1.4.8). So, according to Definition 1.4.10, the *elementary discrepancy* of a sequence ω of points in F is given by

$$D_N^{\mathscr{E}}(\omega) = \sup_{E \in \mathscr{E}} \left| \frac{1}{N} \sum_{i=1}^{N} \chi_E(x_i) - P(E) \right|.$$

The choice of the elementary discrepancy is convenient because the family of elementary sets is obtained in the most natural way by the construction of the fractal F and because the elementary sets can be constructed for every IFS fractal regardless of the complexity of its geometric structure.

In the papers [27], [18] and [17] the authors also gave estimates for the elementary discrepancy of the van der Corput type sequences produced on the particular fractals considered, finding that is of the order $\mathcal{O}\left(\frac{1}{N}\right)$. Our results include theirs, giving a more trasparent proof and taking in consideration the whole class of fractals described in the previous subsection.

The next theorem evaluates the elementary discrepancy for the sequences of points generated by the algorithm.

Theorem 2.2.5.
Let (x_i) be the sequence of points generated on F by the algorithm described in the previous subsection and let $N \geq 1$. Then for the elementary discrepancy we have

$$\lim_{N \to \infty} N D_N^{\mathscr{E}}(\omega_N) = 1.$$

where $\omega_N = (x_1, \ldots, x_N)$.

Proof.
The lower bound is trivial. In fact, if for any $k \in \mathbb{N}$, we have

$$D_N^{\mathscr{E}}(\omega_N) \geq \frac{1}{N} - \frac{1}{m^k}.$$

In order to find an upper bound for $D_N^{\mathscr{E}}(\omega_N)$, let us consider $D_N^{\mathscr{E}_k}(\omega_N)$ for any $k \in \mathbb{N}$. Fix $k \in \mathbb{N}$ and let $E \in \mathscr{E}_k$.

$$D_N^{\mathscr{E}_k}(\omega_N) = \sup_{E \in \mathscr{E}_k} \left| \frac{1}{N} \sum_{i=1}^{N} \chi_E(x_i) - \frac{1}{m^k} \right|.$$

Among the first m^k points of the sequence (x_i) exactly one point is contained in the fixed set E because of the special order induced by the algorithm.

Let us distinguish two different cases:

1. For $N \leq m^k$, the set E contains at most one point of ω_N. Hence

$$\left| \frac{1}{N} \sum_{i=1}^{N} \chi_E(x_i) - \frac{1}{m^k} \right| = \max \left\{ \left| \frac{1}{N} - \frac{1}{m^k} \right|, \left| 0 - \frac{1}{m^k} \right| \right\} \leq \frac{1}{N}.$$

2. If $N > m^k$, we can write N as follows

$$N = Q \cdot m^k + r \text{ with } 0 \leq r < m^k \text{ and } Q \geq 1.$$

Therefore, every $E \in \mathscr{E}_k$ contains either Q points or $Q + 1$ points and hence

$$\left| \frac{1}{N} \sum_{i=1}^{N} \chi_E(x_i) - \frac{1}{m^k} \right| \leq \max \left\{ \left| \frac{Q}{N} - \frac{1}{m^k} \right|, \left| \frac{Q+1}{N} - \frac{1}{m^k} \right| \right\}.$$

Note that

$$\left| \frac{Q}{N} - \frac{1}{m^k} \right| = \left| \frac{Qm^k - N}{Nm^k} \right| = \left| \frac{-r}{Nm^k} \right| < \frac{m^k}{Nm^k} = \frac{1}{N},$$

while

$$\left| \frac{Q+1}{N} - \frac{1}{m^k} \right| = \left| \frac{Qm^k + m^k - N}{Nm^k} \right| = \left| \frac{m^k - r}{Nm^k} \right| = \left| \frac{1}{N} - \frac{r}{Nm^k} \right| < \frac{1}{N}.$$

So we have that

$$\left| \frac{1}{N} \sum_{i=1}^{N} \chi_E(x_i) - \frac{1}{m^k} \right| < \frac{1}{N}.$$

It follows that for any $k \in \mathbb{N}$ we have $D_N^{\mathscr{E}_k}(\omega_N) < \frac{1}{N}$. This implies that $D_N^{\mathscr{E}}(\omega_N) \leq \frac{1}{N}$, as we wanted to prove.

$\square$

Note that $D_N^{\mathscr{E}}(\omega_N)$ is equal to zero for infinitely many N and precisely when $N = \sum_{i=1}^{n} m^i$ for any $n \in \mathbb{N}$. This is due to the fact that the elementary discrepancy of the sequence of partitions (π_n) generated by the algorithm is exactly zero.

Chapter 3

Generalized Kakutani's sequences of partitions

A first extension of the results introduced in Section 2.2 can be obtained by eliminating the restriction that the similarities defining the fractal have all the same ratio. The study of this problem on the simplest case of attractor of an IFS, namely $[0, 1]$, has taken us to consider Kakutani's sequences of partitions and their recent generalization that is the $\rho-$refinements, [69]. In this chapter, we firstly introduce the technique of successive $\rho-$refinements which generalizes Kakutani's splitting procedure and allows to construct new families of u.d. sequences of partitions. Successively, we analyze the problem of deriving bounds for the discrepancy of these sequences. The approach that we use is based on a tree representation of the sequence of partitions constructed by successive $\rho-$refinements, which is precisely the parsing tree generated by Khodak's coding algorithm. Finally, with the help of this technique, we present an application of these results to a class of fractals which includes the one considered in Section 2.2.

3.1   A generalization of Kakutani's splitting procedure

In Subsection 1.2.1 we introduced Kakutani's splitting procedure, which works through successive $\alpha-$refinements of the unit interval. In a recent paper [69], this concept has been generalized through the new notion of $\rho-$refinement and it has been proved that the sequence of partitions generated by successive $\rho-$refinements of the trivial partiton is u.d..

3.1.1 $\rho-$refinements

Consider any non trivial finite partition ρ of $[0,1]$ and from now on we keep it fixed.

Definition 3.1.1.
Let π be any partition of $[0,1]$. The ρ-refinement of π (which will be denoted by $\rho\pi$) is obtained by splitting all the intervals of π having maximal lenght into a finite number of parts positively homothetically to ρ.

Note that, if $\rho = \{[0,\alpha],[\alpha,1]\}$ then the $\rho-$refinement is just Kakutani's $\alpha-$refinement. As in Kakutani's case, we can iterate the splitting procedure. We will denote by $\rho^2\pi$ the ρ-refinement of $\rho\pi$ and, in general, by $\rho^n\pi$ the ρ-refinement of $\rho^{n-1}\pi$.

In the following we will consider the sequence $(\rho^n\omega)$, where ω is the trivial partition of $[0,1]$, and we will prove that $(\rho^n\omega)$ is u.d..

Remark 3.1.2.
It is important to note that in general $(\rho^n\pi)$ is not u.d. for any partition π. For instance, let $\pi = \left\{\left[0,\frac{2}{5}\right],\left[\frac{2}{5},1\right]\right\}$ and $\rho = \left\{\left[0,\frac{1}{2}\right],\left[\frac{1}{2},1\right]\right\}$. It is clear that the $\rho-$refinement operates alternatively on $\left[\frac{2}{5},1\right]$ and $\left[0,\frac{2}{5}\right]$. So, if we consider the sequence of measures (ν_n) associated to $(\rho^n\pi)$ (see (1.8) for the definition), then the subsequence (ν_{2n}) converges to μ_1 while the subsequence (ν_{2n+1}) converges to μ_2 where

$$\mu_1(E) = \frac{5}{4} \cdot \lambda\left(E \cap \left[0,\frac{2}{5}\right]\right) + \frac{5}{6} \cdot \lambda\left(E \cap \left[\frac{2}{5},1\right]\right)$$

$$\mu_2(E) = \frac{5}{6} \cdot \lambda\left(E \cap \left[0,\frac{2}{5}\right]\right) + \frac{10}{9} \cdot \lambda\left(E \cap \left[\frac{2}{5},1\right]\right)$$

for any measurable set $E \subset [0,1]$. Hence, (ν_n) does not converge and consequently $(\rho^n\pi)$ is not u.d..

We can find the problem showed by this example also in the simplest case of Kakutani's splitting procedure. So, it could be interesting to find significant sufficient conditions on π in order to guarantee the uniform distribution of $(\alpha^n\pi)$ or more in general of $(\rho^n\pi)$.

Before introducing the analogous of Kakutani's theorem for these new sequences of partitions (Theorem 3.1.12), let us fix some notations and recall some preliminary results.

Firstly, we need some classical definitions from ergodic theory (see [53, Chapter 29] due to F.Blume).

Definition 3.1.3.

A measurable function $\varphi : [0,1] \to [0,1]$ is said to be measure preserving *if*

- *φ is bijective,*

- *$\varphi(A), \varphi^{-1}(A)$ are measurable when A is measurable,*

- *$\lambda(\varphi^{-1}(A)) = \lambda(A)$.*

A countable family $\mathscr{F}$ of measurable functions is said to be measure preserving if any $\varphi \in \mathscr{F}$ is measure preserving.

Definition 3.1.4.

Given a measurable function $\varphi : [0,1] \to [0,1]$, a measurable set A is called $\varphi-$invariant if

$$\lambda(\varphi^{-1}(A) \; \Delta \; A) = 0,$$

where Δ is the symmetric difference.

    If $\mathscr{F}$ is a countable family of measurable functions, a measurable set A is said to be $\mathscr{F}-$invariant if it is $\varphi-$invariant for any $\varphi \in \mathscr{F}$.

Definition 3.1.5.

A measurable function $\varphi : [0,1] \to [0,1]$ is called ergodic *if it is measure preserving and if for each set A φ-invariant we have*

$$\lambda(A) = 0 \;\; or \;\; \lambda(A) = 1.$$

A countable family $\mathscr{F}$ of measurable functions is said to be ergodic if any $\varphi \in \mathscr{F}$ is ergodic.

Definition 3.1.6.

A real-valued function f on $[0,1]$ is said to be φ-invariant if

$$f(\varphi(x)) = f(x)$$

holds $\lambda-$almost everywhere.

    If $\mathscr{F}$ is a countable family of measurable functions, f is called $\mathscr{F}-$invariant if it is φ-invariant for any $\varphi \in \mathscr{F}$.

    It is easy to prove that

Theorem 3.1.7.

Let $\mathscr{F}$ be a countable ergodic family and suppose that f is a measurable real-valued function on $[0,1]$. If f is $\mathscr{F}-$invariant, then f is constant almost everywhere.

Proof.

Consider for every $\alpha \in \mathbb{R}$ the set

$$A_\alpha = \{x \in [0,1] : f(x) > \alpha\}.$$

A_α is measurable, since the function f is measurable. Moreover, A_α is $\mathscr{F}-$invariant, because f is $\mathscr{F}-$invariant. Therefore we have

$$\lambda(A_\alpha) = 0 \ \ \text{or} \ \ \lambda(A_\alpha) = 1 \ \ \text{for all} \ \alpha \in \mathbb{R}$$

since $\mathscr{F}$ is an ergodic family. It follows that f is constant almost everywhere. In fact, if not, there would exist an $\alpha \in \mathbb{R}$ such that $0 < \lambda(A_\alpha) < 1$ which would contradict the ergodicity of $\mathscr{F}$.

$\square$

Let us recall an important theorem due to Hewitt and Savage, [34, Theorem 11.3].

Theorem 3.1.8.

A product measure on an infinite product of measure spaces can assume only the values 0 and 1 for sets which are invariant under all finite permutations of the coordinates.

Now, we can come back to the $\rho-$refinements and introduce some concepts and properties necessary to the proof of Theorem 3.1.12.

Let $\rho = \{[u_{i-1}, u_i] : 1 \leq i \leq k\}$ be the fixed partition of $[0,1]$ and let us denote by $\alpha_i = u_i - u_{i-1}$, for $1 \leq i \leq k$, the lenghts of the k intervals of ρ. Let $[\rho]^n$ be the $n-$th $\rho-adic\ partition$ of $[0,1]$, obtained from $[\rho]^{n-1}$ (where $[\rho]^1 = \rho$) by subdividing all its k^{n-1} intervals positively homothetically to ρ. If an interval belongs to $[\rho]^n$, we will say that it has rank n.

The k intervals of ρ will be denoted by

$$I(\alpha_i) = \left[\sum_{h=1}^{i-1} \alpha_h, \sum_{h=1}^{i} \alpha_h\right] = [u_{i-1}, u_i].$$

If $[y_{j-1}, y_j] = I(\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{n-1}})$ with $1 \leq j \leq k^{n-1}$ is a generic interval of rank $n-1$, then its subintervals belonging to $[\rho]^n$ are

$$I(\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{n-1}} \alpha_{i_n}) = \left[y_{j-1} + (y_j - y_{j-1})\sum_{h=1}^{i_n-1} \alpha_h, \ y_{j-1} + (y_j - y_{j-1})\sum_{h=1}^{i_n} \alpha_h\right]$$

for $\alpha_{i_n} = \alpha_1, \dots, \alpha_k$. Moreover, by varying $[y_{j-1}, y_j] \in [\rho]^{n-1}$, $1 \leq j \leq k^{n-1}$, we obtain all the k^n intervals of $[\rho]^n$. Note that

$$\lambda(I(\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{n-1}} \alpha_{i_n})) = \prod_{m=1}^{n} \alpha_{i_m}. \tag{3.1}$$

Example

Let $\rho = \left\{ \left[0, \frac{1}{4} \right], \left[\frac{1}{4}, \frac{1}{2} \right], \left[\frac{1}{2}, 1 \right] \right\}$. In this case we have $\alpha_1 = \alpha_2 = \frac{1}{4}$ and $\alpha_3 = \frac{1}{2}$. We only want to construct $[\rho]^n$ for $n = 1, 2$. Then

$$[\rho]^1 = \rho = \left\{ \underbrace{\left[0, \frac{1}{4} \right]}_{I(\alpha_1)}, \underbrace{\left[\frac{1}{4}, \frac{1}{2} \right]}_{I(\alpha_2)}, \underbrace{\left[\frac{1}{2}, 1 \right]}_{I(\alpha_3)} \right\}.$$

Now, $[\rho]^2$ can be obtained from $[\rho]^1$ by splitting all its intervals homothetically to ρ. Practically, we have to subdivide each interval in two equal parts, then take the first of these two halves and split it again in two equal parts. So we have

$$[\rho]^2 = \left\{ \underbrace{\left[0, \frac{1}{16} \right]}_{I(\alpha_1\alpha_1)}, \underbrace{\left[\frac{1}{16}, \frac{1}{8} \right]}_{I(\alpha_1\alpha_2)}, \underbrace{\left[\frac{1}{8}, \frac{1}{4} \right]}_{I(\alpha_1\alpha_3)}, \underbrace{\left[\frac{1}{4}, \frac{5}{16} \right]}_{I(\alpha_2\alpha_1)}, \underbrace{\left[\frac{5}{16}, \frac{3}{8} \right]}_{I(\alpha_2\alpha_2)}, \underbrace{\left[\frac{3}{8}, \frac{1}{2} \right]}_{I(\alpha_2\alpha_3)}, \underbrace{\left[\frac{1}{2}, \frac{5}{8} \right]}_{I(\alpha_3\alpha_1)}, \underbrace{\left[\frac{5}{8}, \frac{3}{4} \right]}_{I(\alpha_3\alpha_2)}, \underbrace{\left[\frac{3}{4}, 1 \right]}_{I(\alpha_3\alpha_3)} \right\}.$$

Let $X = \{\alpha_1, \ldots, \alpha_k\}$ and let σ be the probability on X such that $\sigma(\{\alpha_i\}) = \alpha_i$ for $1 \le i \le k$. Put $X_m = X$ and $\sigma_m = \sigma$ for any $m \in \mathbb{N}$. Denote by

$$Y = \prod_{m=1}^{\infty} X_m$$

and consider on Y the usual product probability μ.

If $C = C(\alpha_{i_1} \alpha_{i_2} \ldots \alpha_{i_n}) = \prod_{m=1}^{\infty} X'_m$, where $X'_m = \{\alpha_{i_m}\}$ for $m \le n$ and $X'_m = X$ for $m > n$, is a cylinder set then

$$
\begin{aligned}
\mu(C) &= \mu\left(\{\alpha_{i_1}\} \times \{\alpha_{i_2}\} \times \cdots \times \{\alpha_{i_n}\} \times X'_{n+1} \times X'_{n+2} \times \cdots \right) \\
&= \sigma(\{\alpha_{i_1}\}) \cdot \sigma(\{\alpha_{i_2}\}) \cdots \sigma(\{\alpha_{i_n}\}) \cdot \sigma(X) \cdot \sigma(X) \cdots \\
&= \prod_{m=1}^{n} \alpha_{i_m}.
\end{aligned}
\tag{3.2}
$$

To every point $t \in [0, 1]$ we can associate a sequence (α_{i_m}) such that

$$t \in \bigcap_{m=1}^{\infty} I(\alpha_{i_1} \ldots \alpha_{i_m}),$$

that is

$$[t]_\rho = \alpha_{i_1} \ldots \alpha_{i_m} \ldots.$$

We called $[t]_\rho$ the $\rho-$adic representation of t.

It is important to take care of an expected ambiguity of this representation. In fact, there are two such sequences (α_{i_m}) associated to a t in the countable set of points belonging to the endpoints of some $[\rho]^n$. In this case to solve the problem, we associate to t the sequence for which definitively $\alpha_{i_m} = \alpha_1$.

This defines a $1-1$ mapping $\phi : [0,1] \to Y'$, where Y' is a subset of Y obtained by removing from Y the countable set of sequences (α_{i_m}) such that definitively $\alpha_{i_m} = \alpha_k$, i.e.

$$Y' = Y \setminus \left\{ (\alpha_{i_m}) : \alpha_{i_m} = \alpha_k \quad \forall m \geq m_0 \right\}.$$

Note that $\mu(Y \setminus Y') = 0$. In fact, if $y \in Y \setminus Y'$ then y is of the form

$$y = (\alpha_{i_1} \dots \alpha_{i_{m_0-1}} \alpha_k \alpha_k \alpha_k \dots)$$

so

$$Y \setminus Y' \subset C(\alpha_{i_1} \dots \alpha_{i_{m_0-1}} \underbrace{\alpha_k \dots \alpha_k}_{n \text{ times}}).$$

Therefore we have that

$$0 \leq \mu(Y \setminus Y') \leq \mu\left(C(\alpha_{i_1} \dots \alpha_{i_{m_0-1}} \underbrace{\alpha_k \dots \alpha_k}_{n \text{ times}}) \right) = \alpha_{i_1} \cdots \alpha_{i_{m_0-1}} \alpha_k^n \to 0$$

as $n \to \infty$, since $\alpha_k < 1$.

Moreover, ϕ is a measure preserving mapping if we take on $[0,1]$ the Lebesgue measure λ and on Y' the restriction of μ. This follows immediately noting that $\rho-$adic intervals and cylinder sets with the same indices have the same measure (see (3.1) and (3.2)). Hence, ϕ is a measure isomorphism between $\big([0,1], \lambda|_{[0,1]}\big)$ and $(Y', \mu|_{Y'})$.

Let I and J be two disjoint subintervals of $[0,1]$ having the same lenght and let $J = I + c$ with $0 < c < 1$. Let us define the following function

$$f_{I,J} = \begin{cases} x + c & \text{if } x \in I \\ x - c & \text{if } x \in J \\ x & \text{otherwise} \end{cases}. \tag{3.3}$$

It is important to observe that $f_{I,J}$ is measure preserving. Let us denote by $\mathcal{F}$ the family of all functions $f_{I,J}$ such that I and J are two $\rho-$adic intervals having the same lenght. The intervals I and J do not need to have necessarily the same rank.

Lemma 3.1.9.
The family $\mathcal{F}$ is ergodic.

Proof.

Let us denote by $\mathcal{F}'$ the family of transformations on Y' correspondent to $\mathcal{F}$ by using the isomorphism described above, i.e

$$\mathcal{F}' = \{f' : Y' \to Y' \text{ s.t. } f' = \phi^{-1} f \phi \text{ for some } f \in \mathcal{F}\}.$$

When $f = f_{I,J}$ and I and J have the same rank, the correspondent function f' on Y' is a permutation of a finite number of coordinates and it preserves the product measure μ. Let us denote by $\mathcal{G}'$ the family of such functions. By Theorem 3.1.8 the family $\mathcal{G}'$ is ergodic and consequently $\mathcal{F}'$ is also ergodic, since $\mathcal{G}' \subset \mathcal{F}'$. In conclusion, since the isomorphism ϕ is measure preserving the family $\mathcal{F}$ results to be ergodic, too.

$\square$

For the partition $\rho^n \omega$, let A_n be the lenght of the longest interval and a_n the lenght of the shortest interval. Moreover, let us denote by $k(n)$ is the number of intervals of the partition $\rho^n \omega$. We have the following results.

Lemma 3.1.10.

1. *For any $n \in \mathbb{N}$ we have $a_1 A_n \leq a_n$.*

2. *If $\pi_n = \rho^n \omega$ then $\lim_{n\to\infty} diam(\pi_n) = 0$.*

Proof.

1. Since $A_n < 1$ for any $n \in \mathbb{N}$, the strict inequality holds for $n = 1$. Now, proceed by induction. Suppose that $a_1 A_{n-1} \leq a_{n-1}$ holds. There are two possibilities either $a_n = a_{n-1}$ or $a_n < a_{n-1}$. In the first case, since $A_n < A_{n-1}$ we have

$$a_1 A_n < a_1 A_{n-1} \leq a_{n-1} = a_n.$$

In the second case, the shortest interval of the partition $\rho^n \omega$ is obtained by splitting the longest interval of $\rho^{n-1} \omega$, so $a_n = a_1 A_{n-1}$. Hence, we have

$$a_1 A_n < a_1 A_{n-1} = a_n.$$

2. According to the notation introduced above, obviously $a_n < \frac{1}{n}$ for any $n \in \mathbb{N}$. By applying the relation just proved we have

$$diam(\pi_n) = A_n \leq \frac{a_n}{a_1} < \frac{1}{a_1 n}$$

and so the conclusion follows.

$\square$

Lemma 3.1.11.

The family of the characteristic functions of all intervals belonging to the partitions $\rho^n \omega$ for $n \in \mathbb{N}$ is determining.

Proof.

Let $\mathcal{M}$ be the class consisting of the characteristic functions of all intervals belonging to the partitions $\rho^n \omega$ for $n \in \mathbb{N}$ and let $f \in \mathcal{C}([0,1])$.

By uniform continuity we have that for any $\varepsilon > 0$ there exists $\overline{\delta} > 0$ such that $|f(x') - f(x'')| < \varepsilon$ whenever $|x' - x''| < \overline{\delta}$. Moreover, by the second part of Lemma 3.1.10, we have that for any $\delta > 0$ there exist $\overline{n} \in \mathbb{N}$ such that for every $n > \overline{n}$ we have $diam(\rho^n \omega) < \delta$. Hence for every $n > \overline{n}$ we have that each $E_j^n \in \rho^n \omega$ has diameter smaller than δ.

So we can choose $n \in \mathbb{N}$ such that each $E_j^n \in \rho^n \omega$ has diameter smaller than $\overline{\delta}$. Fixed a such n, take a point t_j in any $E_j^n \in \rho^n \omega$ and consider the function

$$g(y) = \sum_{j=1}^{k(n)} f(t_j) \chi_{E_j^n}(y), \quad y \in [0,1].$$

Now, $y \in E_j^n$ for some $j \in \{1, \ldots, k(n)\}$ then $|g(y) - f(y)| = |f(t_j) - f(y)| < \varepsilon$. Hence, $span(\mathcal{M})$ is uniformly dense in $\mathcal{C}([0,1])$ with respect to the $||\cdot||_\infty$ and so the conclusion follows by Theorem 1.4.9.

$\square$

3.1.2 A generalization of Kakutani's Theorem

Theorem 3.1.12.

The sequence $(\rho^n \omega)$ is u.d..

Proof.

Let us denote by (ν_n) the sequence of measures associated to the sequence of partitions $(\rho^n \omega)$ defined accordingly to (1.8). We have to prove that (ν_n) is weakly convergent to λ.

It is well known that the set of all Borel probability measures on $[0,1]$, with the topology associated to the weak convergence, is metrizable and compact (see [54, Theorem 6.4]). Then (ν_n) admits weakly convergent subsequences. So all we need to prove is that any such subsequence converges to λ.

First of all, let us prove that the family of the characteristic functions of all $\rho-$adic intervals is determining. In fact, every interval $J \in \rho^n \omega$ belongs to some $[\rho]^m \omega$. The

viceversa is also true, namely every interval $I \in [\rho]^m \omega$ sooner or later belongs to some $\rho^n \omega$. This is due to the fact that

$$s = \sup \{ r : J_r \in \rho^r \omega, \ I \subset J_r \}$$

is well defined, since the ρ−adic intervals are either disjoint or contained one in the other. Moreover, by the second part of Lemma 3.1.10, the diameter of $\rho^n \omega$ tends to zero and so $I = J_s$. Therefore, by Lemma 3.1.11 we can conclude that the family of the characteristic functions of all ρ−adic intervals is determining, too.

Consequently, it is sufficient to prove that for each weakly convergent subsequence (ν_{n_k}) we have

$$\nu_{n_k}(J) \rightharpoonup \lambda(J), \ \ \forall J \in [\rho]^m \omega, \ \forall m \in \mathbb{N}$$

where we denote by "$\rightharpoonup$" the weak convergence.

Let J be any ρ−adic interval and suppose that $m \in \mathbb{N}$ is such that for any $n \geq m$ every $\rho^n \omega$ subdivides J. If J is splitted in k intervals by $\rho^n \omega$, then

$$k a_n \leq \lambda(J) \leq k A_n$$

where a_n and A_n are the quantities considered in Lemma 3.1.10.

If $k(n)$ is the number of intervals in $\rho^n \omega$, we have that for all $n \geq m$

$$\frac{\lambda(J)}{k(n) A_n} \leq \nu_n(J) \leq \frac{\lambda(J)}{k(n) a_n}. \tag{3.4}$$

In fact, let us denote by $K(J)$ the number of intervals in which J is subdivided by $\rho^n \omega$, i.e

$$K(J) := \sum_{i=1}^{k(n)} \delta_{t_i^n}(J),$$

where the points t_i^n are the points determining the partition $\rho^n \omega$.
So $\nu_n(J) = \frac{1}{k(n)} \sum_{i=1}^{k(n)} \delta_{t_i^n}(J) = \frac{K(J)}{k(n)}$ and therefore

- $\lambda(J) \geq K(J) a_n = \nu_n(J) k(n) a_n \Rightarrow \nu_n(J) \leq \frac{\lambda(J)}{k(n) a_n}$,

- $\lambda(J) \leq K(J) A_n = \nu_n(J) k(n) A_n \Rightarrow \nu_n(J) \geq \frac{\lambda(J)}{k(n) A_n}$.

By (3.4) and by the first part of Lemma 3.1.10 we have

$$a_1 \lambda(J) \leq \frac{a_1 \lambda(J)}{k(n) a_n} \leq \nu_n(J) \leq \frac{\lambda(J)}{k(n) a_1 A_n} \leq \frac{\lambda(J)}{a_1}. \tag{3.5}$$

Now, suppose that (ν_{n_k}) is a subsequence weakly convergent to ν. Then, by (3.5) we have that for any $\rho-$adic interval J the following holds

$$a_1 \lambda(J) \leq \nu(J) \leq \frac{\lambda(J)}{a_1}. \tag{3.6}$$

Since the family of the characteristic functions of all $\rho-$adic intervals is determining, the relation (3.6) holds for any Borel set B in $[0,1]$. Therefore $\lambda \ll \nu \ll \lambda$ and if we denote by $\frac{d\nu}{d\lambda}$ the Radon-Nikodym derivative of ν with respect to λ, then we have that

$$a_1 \leq \frac{d\nu}{d\lambda} \leq \frac{1}{a_1}.$$

Note that if I and J are two intervals having the same lenght and belonging to some $\rho^n \omega$ (not necessarily having the same rank), then the splitting procedure behaves on them in the same way. This implies that $\nu(I) = \nu(J)$. Then by applying Radon-Nikodym's theorem we have

$$\int_I \frac{d\nu}{d\lambda} d\lambda = \int_J \frac{d\nu}{d\lambda} d\lambda$$

and so

$$\frac{d\nu}{d\lambda}(I) = \frac{d\nu}{d\lambda}(J) \ \text{ a.e.}.$$

But $J = f_{I,J}(I)$ for (3.3), so the previous relation becomes

$$\frac{d\nu}{d\lambda}(I) = \frac{d\nu}{d\lambda}(f_{I,J}(I)) \ \text{ a.e.}.$$

It follows that $\frac{d\nu}{d\lambda}$ is $\mathcal{F}-$invariant. Hence, since by Lemma 3.1.9 the family $\mathcal{F}$ is ergodic, Theorem 3.1.7 implies that $\frac{d\nu}{d\lambda}$ is costant a.e.. In particular, we have

$$\frac{d\nu}{d\lambda} = 1 \ \text{ a.e.}$$

because

$$1 = \nu([0,1]) = \int_{[0,1]} \frac{d\nu}{d\lambda} d\lambda = \frac{d\nu}{d\lambda} \cdot \lambda([0,1]) = \frac{d\nu}{d\lambda}.$$

Then $\nu = \lambda$.

$\square$

3.2 Discrepancy of some generalized Kakutani's sequences

A natural problem, posed in [69], is to estimate the behaviour of the discrepancy (1.7) of the sequence of partitions generated by successive $\rho-$refinements as n tends

to infinity. In particular, it is interesting to find partitions ρ such that the speed of convergence of the discrepancy to zero is as high as possible. The only known discrepancy bounds for sequences of this kind have been obtained by Carbone [10], who considered the so-called LS-sequences that evolve from partitions ρ with L subintervals of $[0, 1]$ of length α and S subintervals of length α^2 (where α is given by the equation $L\alpha + S\alpha^2 = 1$).

In this section, we analyze this problem with a new approach based on a parsing tree (related to the Khodak coding algorithm [43]) which represents the successive ρ-refinements. In particular, we use refinements of the results proved in [20] about Khodak's algorithm to give bounds of the discrepancy for a class of sequences of partitions constructed by successive $\rho-$refinements. Finally, we present some examples and applications including LS-sequences and u.d. sequences of partitions on a class of fractals. These results are collected in [19].

3.2.1 Correspondence between $\rho-$refinements and Khodak's algorithm

From now on, consider a partition ρ of $[0, 1]$ consisting of m intervals of lengths $p_1, \ldots, p_m$ and the sequence of ρ-refinements of the trivial partition $\omega = \{[0, 1]\}$. Our goal is to construct recursively an m-ary tree which represents the process of successive $\rho-$refinements of ω.

Definition 3.2.1.
An m-ary tree is an ordered rooted tree, where each node has either m ordered successors or it is a leaf with no successors. A node with m successors is called internal node, while a leaf that has no successors is called also external node.

The numbers $p_1, \ldots, p_m$ induce a natural labelling on the nodes. Suppose that the unique path from the root to a node x at level l is encoded by the sequence $(j_1, j_2, \ldots, j_l)$, $j_i \in \{1, 2, \ldots, m\}$, then we set $P(x) = p_{j_1} p_{j_2} \cdots p_{j_l}$. This can be also considered as the probability of reaching the node x with a random walk that starts at the root and moves away from it according to the probabilities $p_1, \ldots, p_m$. For completeness the root a is labelled with $P(a) = 1$. If T is a finite m-ary tree then the labels of the external nodes sum up to 1. Hence, the shape of an m-ary tree (together with $p_1, \ldots, p_m$) gives rise of a probability distribution.

The start of our iteration is a tree that only consists of the root which is then an external node with probability 1. In the first step, the root is replaced by an internal node together with m ordered successing leaves that are given the probability

distribution $p_1, \ldots, p_m$. At each further iteration we select all leaves y with largest label $P(y)$ and grow m children out of each of them. Actually, this construction corresponds to the procedure of successive ρ-refinements. The leaves of the tree correspond to the intervals of $\rho^n \omega$ and the labels of the leaves to the lengths of these intervals.

This procedure exactly leads to the same parsing tree of the *Tunstall code* [20]. In fact, the m initial leaves correspond to the symbols of an $m-$ary alphabet $\mathcal{A}$ and so the words $(j_1, j_2, \ldots, j_l)$ that encode the paths from the root to the leaves are the phrases of the dictionary. It is important to note that at each iteration we can have different leaves of the same probability, but Tunstall's algorithm selects (randomly) only one of these leaves and grow m children out of it.

There is a second way to describe this tree evolution process, namely by *Khodak's algorithm* [43]. Fix a real number $r \in]0, p_{min}[$, where $p_{min} = \min\{p_1, \ldots, p_m\}$, and consider all nodes x among in an infinte m-ary tree with $P(x) \geq r$. Let us denote these nodes by $\mathcal{I}(r)$. Of course, if $P(x) \geq r$ then all nodes x' on the path from the root to x satisfy $P(x') \geq r$, too. Hence, these nodes of $\mathcal{I}(r)$ constitute a finite subtree. These nodes will be the *internal nodes* of Khodak's construction. Finally, we append to these internal nodes all successor nodes y. By construction all these nodes satisfy $p_{min}r \leq P(y) < r$ and we denote them by $\mathcal{E}(r)$. These nodes are the *external nodes* of Khodak's construction. We denote by $M_r = |\mathcal{E}(r)|$ the number of external nodes. Obviously, we have got a finite m-ary tree $\mathcal{T}(r) = \mathcal{I}(r) \cup \mathcal{E}(r)$ and it is clear that these trees grow when r decreases. For certain values r, precisely the external nodes y of largest value $P(y) = r$ turn into internal nodes and all their successors become new external nodes. Actually, the tree $\mathcal{T}(r)$ grows in correspondence to a decreasing sequence of values (r_j). Indeed, when $r \in]r_j, r_{j-1}]$ the tree remains the same, i.e $\mathcal{T}(r_{j-1}) = \mathcal{T}(r)$.

The parsing tree resulting from Khodak's algorithm is exactly the same as the tree constructed by Tunstall's algorithm. However, we have to observe that in Khodak's construction all leaves with the same highest probability are selected to generate the children at once, while in Tunstall'algortithm they are selected one by one in an arbitrary order. Now, in the procedure of successive $\rho-$refinements, at each step we select the intervals having maximal lenght at once and we split them at the same time. So, Khodak's algorithm and $\rho-$refinements procedure not only are exactly represented by the same tree but they also have a common structure which allows to create a useful correspondence between them.

In fact, if we fix a step j in $\rho-$refinements procedure, then the tree associated

to the partition $\rho^j \omega$ is exactly $\mathcal{T}(r_j)$. Therefore, we will only consider the values of the sequence (r_j) for which the tree constructed by Khodak's algorithm actually grow. Note that we have to start with the value $r_1 = 1$, because we intend to consider the whole procedure since the first step which corresponds to $\rho\omega$. Hence, in our correspondence between Khodak's algorithm and the procedure of successive $\rho-$refinements the value $r_j \in]0, 1]$ corresponds to the partition $\rho^j \omega$. Consequently, the number of external nodes in $\mathcal{E}(r_j)$ equals the number of points defining the partition $\rho^j \omega$, i.e. $M_{r_j} = k(j)$. Moreover, if $r \in]r_j, r_{j-1}]$ then $M_r = M_{r_{j-1}} = k(j-1)$.

From here on we denote by $\mathscr{E}_{r_j}$ the family of all intervals of the partition $\rho^j \omega$ corresponding to the leaves belonging to $\mathcal{E}(r_j)$ and the order of the intervals in $\mathscr{E}_{r_j}$ corresponds to the left-to-right order of the external nodes in $\mathcal{E}(r_j)$. We will call *elementary intervals* all the intervals belonging to each $\mathscr{E}_r$ for $r \in]0, 1]$.

In the following we denote by H the entropy of the probability distribution $p_1, \ldots, p_m$, which is defined as

$$H = p_1 \log \left(\frac{1}{p_1} \right) + \cdots + p_m \log \left(\frac{1}{p_m} \right).$$

Example
Let $\rho = \left\{ \left[0, \frac{1}{4} \right], \left[\frac{1}{4}, \frac{1}{2} \right], \left[\frac{1}{2}, 1 \right] \right\}$ and $\omega = \{ [0, 1] \}$. So in this case we have that $p_1 = p_2 = \frac{1}{4}$ and $p_3 = \frac{1}{2}$. In Figure 3.1 the correspondence between the tree constructed by Khodak's algorithm and the $\rho-$refinements of ω is illustrated in the first three steps of these procedures. In particular, the internal nodes are coloured in black and the external ones in grey. Moreover, it is easy to note that the label of each node corresponds to the lenght of the corresponding interval.

The start of the procedure of $\rho-$refinements is the trivial partition ω which corresponds to the root node of probability 1. At the first step we take $r_1 = 1$, so the root becomes an internal node and we grow three leaves out of it. The three children nodes have probability p_1, p_2 and p_3 respectively and each of them corresponds to an interval of $\rho\omega$ (see Figure 3.1a). The next value of r for which we have an actual grewth of the tree is $r_2 = \frac{1}{2}$. Consequently, at the second step, we select all the nodes x having $P(x) \geq r_2$ and grow three children out of each of them. The external nodes generated at the end of this step correspond exactly to the intervals of $\rho^2 \omega$ (see Figure 3.1b). At the third step we choose $r_3 = \frac{1}{4}$ and we proceed at the same way of the previous steps and so we get 11 leaves which are associated to the intervals of $\rho^3 \omega$ (see Figure 3.1c). By iterating this procedure for all the values of (r_j), we will get the whole infinity tree corresponding to the sequence of partitions $(\rho^j \omega)$.
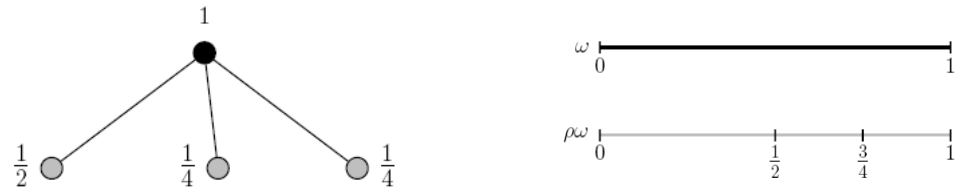
(a) I step
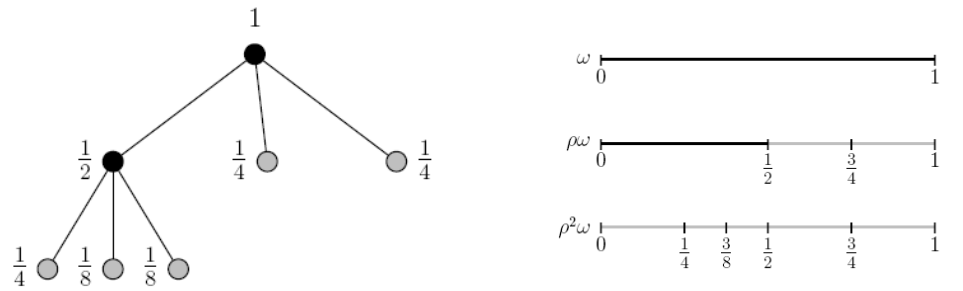
(b) II step
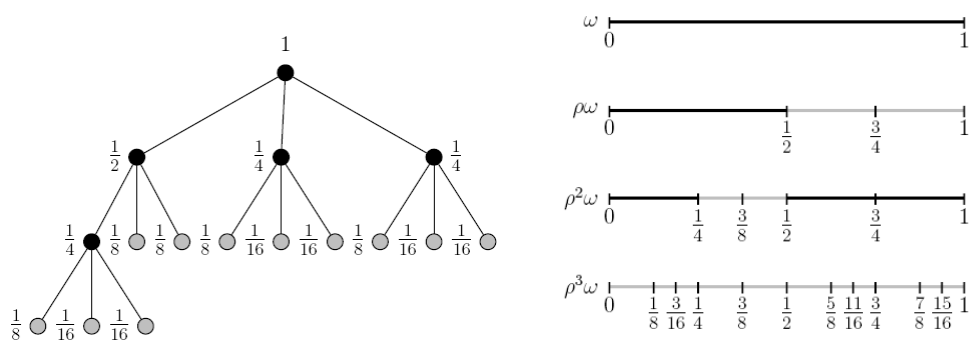
(c) III step

Figure 3.1: Correspondence between $\rho-$refinements of ω and Khodak's tree.

3.2.2   Asymptotic results on Khodak's algorithm

One of the main results from [20] provides asymptotic information on the number M_r of external nodes in Khodak's construction. Actually, these relations can be used to prove Theorem 3.1.12. However, in order to derive bounds for the discrepancy of the sequence $(\rho^n \omega)$ we need more precise information on the error terms in the asymptotic results given in [20]. Therefore, we will extend the analysis of this paper in Theorem 3.2.3. Before introducing it, let us give a fundamental definition, which allows us to distinguish two main cases in our discussion.

Definition 3.2.2.
We say that $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ *are* rationally related *if there exists a positive real number* Λ *such that* $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ *are integer multiples of* Λ, *that is*

$$\log\left(\frac{1}{p_j}\right) = n_j \Lambda, \quad \text{with } n_j \in \mathbb{Z} \text{ for } j = 1, \ldots, m.$$

Without loss of generality we can assume that Λ *is as large as possible which is equivalent to assume that* $\gcd(n_1, \ldots, n_m) = 1$. *Equivalently, all fractions* $(\log p_i)/(\log p_j)$ *are rational.*

Similarly we say that $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ *are* irrationally related *if they are not rationally related.*

Theorem 3.2.3.
Let M_r *be the number of the external nodes generated at the step corresponding to the parameter* r *in Khodak's construction, that is, the number of nodes in* $\mathcal{E}(r)$.

1. *If* $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ *are rationally related, let* $\Lambda > 0$ *be the largest real number for which* $\log\left(\frac{1}{p_j}\right)$ *is an integer multiple of* Λ *(for* $j = 1, \ldots, m$*). Then there exists a real number* $\eta > 0$ *and an integer* $d \geq 0$ *such that*

$$M_r = \frac{(m-1)}{rH} Q_1\left(\log\left(\frac{1}{r}\right)\right) + \mathcal{O}\left((\log r)^d r^{-(1-\eta)}\right), \qquad (3.7)$$

   where

$$Q_1(x) = \frac{\Lambda}{1 - e^{-\Lambda}} e^{-\Lambda\left\{\frac{x}{\Lambda}\right\}}$$

   and $\{y\}$ *is the fractional part of the real number* y. *Furthermore, the error term is optimal.*

2. *If* $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ *are irrationally related, then*

$$M_r = \frac{(m-1)}{rH} + o\left(\frac{1}{r}\right). \qquad (3.8)$$

In particular, if $m = 2$ and $\gamma = (\log p_1)/(\log p_2)$ is badly approximable then

$$M_r = \frac{(m-1)}{rH}\left(1 + \mathcal{O}\left(\frac{(\log\log(1/r))^{1/4}}{(\log(1/r))^{1/4}}\right)\right).$$

Moreover, if p_1 and p_2 are algebraic then there exists an effectively computable constant $\kappa > 0$ with

$$M_r = \frac{(m-1)}{rH}\left(1 + \mathcal{O}\left(\frac{(\log\log(1/r))^{\kappa}}{(\log(1/r))^{\kappa}}\right)\right).$$

The proof of this theorem requires some auxiliary results, so let us introduce them preliminarly.

Trigonometric Sums

Lemma 3.2.4.
Let $f(n) = \sum\limits_{i=1}^{k} c_i \cos(2\pi\theta_i n + \alpha_i)$, $c_i, \alpha_i, \theta_i \in \mathbb{R}$ be defined for non-negative integers n and suppose that f is not identically zero. Then there exists $\delta > 0$ such that $|f(n)| \geq \delta$ for infinitely many non-negative integers n.

Proof.
We have to distinguish two cases:

Case 1 $\theta_1, \ldots, \theta_k$ are rationally related.
There exist $\Lambda \in \mathbb{R} \setminus \{0\}$ and $k_i \in \mathbb{Z}$ such that $\theta_i = \Lambda k_i$. In this case, we can rewrite the function f as follows

$$f(n) = \sum_{i=1}^{k} c_i \cos(2\pi\Lambda n k_i + \alpha_i) = \sum_{i=1}^{k} c_i \cos(2\pi\{\Lambda n\}k_i + \alpha_i).$$

Hence, $f(n) = g(\{\Lambda n\})$ where $g(x) = \sum\limits_{i=1}^{k} c_i \cos(2\pi k_i x + \alpha_i)$ is a periodic non-zero function of period 1.

Case 1.1 If $\Lambda \in \mathbb{Q}$, then $\Lambda = \frac{p}{q}$ for some coprime integers $p, q \in \mathbb{Z}$ and the sequence $(f(n))$ attains periodically the set of values

$$g\left(\left\{\frac{pn}{q}\right\}\right), \quad n = 0, \ldots, q-1.$$

Since they are not all equal to zero there exists $\delta > 0$ such that

$$|f(n)| = |g(\{\Lambda n\})| \geq \delta$$

for infinitely many n. In particular, we can use a linear subsequence $qn+r$ for which $|f(qn+r)| \geq \delta$.

Case 1.2 If $\Lambda \notin \mathbb{Q}$, then the sequence $(\{\Lambda n\})$ is u.d. on $[0,1]$ and consequently dense in $[0,1]$. Hence, there again exists $\delta > 0$ such that

$$|f(n)| = |g(\{\Lambda n\})| \geq \delta$$

for infinitely many n.

Case 2 $\theta_1, \ldots, \theta_k$ are irrationally related.

Here we divide the θ_i in groups which are rationally related. Assume that we have s groups $\{\theta_i : i \in I_j\}$, $j = 1, \ldots, s$, and in each group we write

$$\theta_i = \Lambda_j k_i, \quad i \in I_j$$

with $k_i \in \mathbb{Z}$ and some $\Lambda_j \in \mathbb{R} \setminus \{0\}$.

In this case, we distinguish between three different subcases.

Case 2.1 $1, \Lambda_1, \ldots, \Lambda_s$ are linearly independent over $\mathbb{Q}$ (so $\Lambda_1, \ldots, \Lambda_s \notin \mathbb{Q}$).

We set $f_j(x) = \sum_{i \in I_j} c_i \cos(2\pi x k_i + \alpha_i)$ (where we assume w.l.o.g. that f_j is non-zero) and $g(x_1, \ldots, x_s) = \sum_{j=1}^{s} f_j(x_j)$. Then

$$f(n) = \sum_{j=1}^{s} f_j(\{n\Lambda_j\}) = g(\{n\Lambda_1\}, \ldots, \{n\Lambda_s\}).$$

By Kronecker's Theorem (Theorem 1.3.5), the sequence $(\{n\Lambda_1\}, \ldots, \{n\Lambda_s\})$ is dense in the cube $[0,1]^s$. Thus, it follows (as above) that there exists $\delta > 0$ such that $|f(n)| \geq \delta$ for infinitely many n.

Note that by same reasoning it follows that for every $\varepsilon > 0$ we have $|f(n)| \leq \varepsilon$ for infinitely many n. (Here we also use the fact that f has zero mean.) This observation will be used in Case 2.3.

Case 2.2 $1, \Lambda_1, \ldots, \Lambda_s$ are linearly dependent over $\mathbb{Q}$ and $\Lambda_1, \ldots, \Lambda_s \notin \mathbb{Q}$.

In this case there exist $q, p_1, \ldots, p_s \in \mathbb{Z}$ such that $q = p_1 \Lambda_1 + \ldots + p_s \Lambda_s$. Suppose (w.l.o.g.) that $p_1 > 0$ and consider the subsequence of integers $(p_1 n)$, then

$$\begin{aligned} f(p_1 n) &= \sum_{j=1}^{s} f_j(n\Lambda_j p_1) \\ &= f_1(n(q - \Lambda_2 p_2 - \cdots - \Lambda_s p_s)) + \sum_{j=2}^{s} f_j(n\Lambda_j p_1). \end{aligned}$$

By using the addition theorem for cosine and rewriting the sum accordingly, we obtain a representation of the form

$$f(p_1 n) = \sum_{j=2}^{s} \tilde{f}_j(n\Lambda_j p_j),$$

where $\tilde{f}_j$ are certain trigonometric polynomials. This means that we have eliminated Λ_1.

In this way we can proceed further. If $1, p_2\Lambda_2, \ldots, p_s\Lambda_s$ are linearly independent over $\mathbb{Q}$ then we argue as in Case 2.1. However, if $1, p_2\Lambda_2, \ldots, p_s\Lambda_s$ are linearly dependent over $\mathbb{Q}$ then we repeat the elimination procedure etc. Note that this elimination procedure terminates, since we assume that $\Lambda_1, \ldots, \Lambda_s \notin \mathbb{Q}$. Hence, we always end up in Case 2.1.

Case 2.3 $\Lambda_1, \ldots, \Lambda_s$ are not all irrationals.

Here we represent $f(n) = h_1(n) + h_2(n)$, where

$$h_1(n) = \sum_{j \in \{j : \Lambda_j \in \mathbb{Q}\}} f_j(n) \quad \text{and} \quad h_2(n) = \sum_{j \in \{j : \Lambda_j \notin \mathbb{Q}\}} f_j(n).$$

If h_1 is non-zero then we can argue as in Case 1.1. All appearing θ_i are rational and consequently there exists a linear subsequence $qn + r$ such that $|h_1(qn+r)| \geq \frac{3\delta}{2}$ for some $\delta > 0$. Next we reduce the sum $h_2(qn+r)$ to a sum of the form that is discussed in Case 2.1 (possibly we have to eliminate several terms as discussed in Case 2.2.). Consequently, it follows that there exist infinitely many n such that $|h_2(qn+r)| \leq \delta/2$. Hence we have $|f(n)| \geq \delta$ for infinitely many n.

If h_1 is zero for all non-negative integers we just have to consider h_2. But this case is precisely that of Case 2.2.

$\square$

Zerofree Regions

The purpose of this paragraph is to discuss zero-free regions of the equation

$$1 - p^{-s} - q^{-s} = 0 \tag{3.9}$$

where p, q are positive numbers with $p + q = 1$.

It is clear that $s = -1$ is a solution of (3.9) and that all its solutions have to satisfy $\Re(s) \geq -1$. Otherwise, we would have

$$|p^{-s} + q^{-s}| \leq |p^{-s}| + |q^{-s}| = e^{-\Re(s)\log(p)} + e^{-\Re(s)\log(q)} < e^{\log(p)} + e^{\log(q)} = p + q = 1.$$

Furthermore, it is easy to verify the following property

Proposition 3.2.5.
There are no solutions of (3.9), other than $s = -1$, on the line $\Re(s) = -1$ if and only if the ratio $\gamma = (\log p)/(\log q)$ is irrational.

Proof.
Let s be of the form $s = -1 + it$. Then we have

$$p^{-s} + q^{-s} = p^{1-it} + q^{1-it} = pp^{-it} + qq^{-it} = pe^{it\log\left(\frac{1}{p}\right)} + qe^{it\log\left(\frac{1}{q}\right)}.$$

Therefore, s is a solution of (3.9) if and only if

$$pe^{it\log\left(\frac{1}{p}\right)} + qe^{it\log\left(\frac{1}{q}\right)} = 1.$$

Since $p + q = 1$, then
$$e^{it\log\left(\frac{1}{p}\right)} = 1 = e^{it\log\left(\frac{1}{q}\right)}$$

necessarily holds. But the last relations imply that there exist $k, l \in \mathbb{Z}$ such that

$$t\log\left(\frac{1}{p}\right) = 2\pi k \quad \text{and} \quad t\log\left(\frac{1}{q}\right) = 2\pi l.$$

It follows that $\gamma = \frac{\log\left(\frac{1}{p}\right)}{\log\left(\frac{1}{q}\right)} = \frac{k}{l}$, which is a rational number.

$\square$

Another important property about the structure of the set of the solutions of (3.9) is the following (see [58]).

Proposition 3.2.6.
There exist two real numbers $\sigma_0 > -1$ and $\tau > 0$ such that in each box of the form

$$B_k = \{s \in \mathbb{C} : -1 \leq \Re(s) \leq \sigma_0, \ (2k-1)\tau \leq \Im(s) < (2k+1)\tau\}, \quad k \in \mathbb{Z} \setminus \{0\},$$

there is precisely one zero of (3.9), and there are no other zeros.

However, the positions of the zeros in B_k is by no means clear. Nevertheless, with the help of the continued fractional expansion of γ it is possible to construct infinitely many zeros s of (3.9) with $\Re(s) < -1 + \varepsilon$ (for every $\varepsilon > 0$). Therefore, it is natural to ask for zerofree regions of this equation. Actually, one has to assume some Diophantine conditions on γ to get precise information.

Lemma 3.2.7.

If γ is badly approximable then for every solution $s \neq -1$ of the equation

$$1 - p^{-s} - q^{-s} = 0$$

we have that

$$\Re(s) > \frac{c}{(\Im(s))^2} - 1$$

for some positive constant c.

Before proving the lemma, let us recall some basic notions of the theory of continued fractions, [42].

Definition 3.2.8.

The continued fractional expansion of a real number x is given by

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}}$$

where $a_0, a_1, \dots$ are positive integers. In a compact notation we can write

$$x = [a_0; a_1; a_2; a_3 \dots].$$

Definition 3.2.9.

An irrational number γ is said to be badly approximable if its continued fractional expansion $\gamma = [a_0; a_1; \dots]$ is bounded, i.e. there exists a positive constant D such that $\max\limits_{j \geq 1}(a_j) \leq D$.

Equivalently, we have the following property.

Proposition 3.2.10.

An irrational number γ is badly approximable if there exists a constant $d > 0$ such that

$$\left| \gamma - \frac{k}{l} \right| \geq \frac{d}{l^2} \tag{3.10}$$

for all non-zero integers k, l.

Proof of Lemma 3.2.7.

In order to make the presentation of the proof more transparent we make a shift by 1 and consider the equation

$$p^{1-s} + q^{1-s} = 1 \tag{3.11}$$

and show that all its non-zero solutions satisfy $\Re(s) > c/\Im(s)^2$ for some positive constant c that depends on γ.

Suppose that $s = \sigma + i\tau$ is a zero of (3.11) with $\sigma > 0$. Furthermore, we assume that $\sigma \leq \varepsilon$, where ε is a sufficiently small constant. Since $p + q = 1$ and $|p^{1-s}| = p^{1-\sigma} = p(1 + \mathcal{O}(\varepsilon)) > p$ and $|q^{1-s}| = q^{1-\sigma} = q(1 + \mathcal{O}(\varepsilon)) > q$ we can only have a solution if the arguments of p^{1-s} and q^{1-s} are small. (Actually they have to be of order $\mathcal{O}(\sqrt{\varepsilon})$ if ε is chosen sufficiently small). W.l.o.g. we write

$$\arg(p^{1-s}) = \tau \log(1/p) = 2\pi k + \eta_1 \quad \text{and} \quad \arg(q^{1-s}) = \tau \log(1/q) = 2\pi l - \eta_2$$

for some integers k, l and certain positive numbers η_1, η_2 (which are of order $\mathcal{O}(\sqrt{\varepsilon})$). More precisely, by doing a local expansion in (3.11) we obtain

$$\eta_2 = \frac{p}{q}\eta_1 + \mathcal{O}(\eta_1^2) \quad \text{and} \quad \sigma = \frac{p}{2qH}\eta_1^2 + \mathcal{O}(\eta_1^4).$$

In fact (3.11) is equivalent to

$$|p^{1-s}|(\cos(\arg(p^{1-s})) + i\sin(\arg(p^{1-s}))) + |q^{1-s}|(\cos(\arg(q^{1-s})) + i\sin(\arg(q^{1-s}))) = 1$$

and so

$$p^{1-\sigma}(\cos(\eta_1) + i\sin(\eta_1)) + q^{1-\sigma}(\cos(\eta_2) - i\sin(\eta_2)) = 1. \tag{3.12}$$

Therefore we have

$$p^{1-\sigma}\sin(\eta_1) - q^{1-\sigma}\sin(\eta_2) = 0$$

and by doing a local expansion it follows that

$$p\left(1 + \log\left(\frac{1}{p}\right)\sigma + \mathcal{O}(\sigma^2)\right) \cdot (\eta_1 + \mathcal{O}(\eta_1^2))$$
$$+ \quad q\left(1 + \log\left(\frac{1}{q}\right)\sigma + \mathcal{O}(\sigma^2)\right) \cdot (-\eta_2 + \mathcal{O}(\eta_2^2)) = 0.$$

Now, by taking in account that $\sigma \leq \epsilon$ and η_1, η_2 are both of the order $\mathcal{O}(\sqrt{\varepsilon})$, we have

$$p\eta_1 - q\eta_2 + \mathcal{O}(\eta_1^2) = 0$$

and so

$$\eta_2 = \frac{p}{q}\eta_1 + \mathcal{O}(\eta_1^2). \tag{3.13}$$

On the other hand, by (3.12) it follows also that

$$p^{1-\sigma}\cos(\eta_1) + q^{1-\sigma}\cos(\eta_2) = 0$$

and by doing a local expansion it follows that

$$p\left(1 + \log\left(\frac{1}{p}\right)\sigma + \mathcal{O}(\sigma^2)\right) \cdot \left(1 - \frac{\eta_1^2}{2} + \mathcal{O}(\eta_1^4)\right)$$
$$+ \quad q\left(1 + \log\left(\frac{1}{q}\right)\sigma + \mathcal{O}(\sigma^2)\right) \cdot \left(1 - \frac{\eta_2^2}{2} + \mathcal{O}(\eta_2^4)\right) = 1.$$

Now, by using the same argumentations of above, we have

$$\sigma\left(p\log\left(\frac{1}{p}\right) + q\log\left(\frac{1}{q}\right)\right) = \frac{p\eta_1^2}{2} + \frac{q\eta_2^2}{2} + \mathcal{O}(\eta_1^4)$$

and so by (3.13) we get

$$\sigma = \frac{p}{2H}\eta_1^2 + \frac{p^2}{2Hq}\eta_1^2 + \mathcal{O}(\eta_1^4) = \frac{p}{2qH}\eta_1^2 + \mathcal{O}(\eta_1^4).$$

Furthermore, we have

$$\begin{aligned}
\gamma &= \frac{\tau\log\frac{1}{p}}{\tau\log\frac{1}{q}} \\
&= \frac{2\pi k + \eta_1}{2\pi l - \eta_2} \\
&= \frac{k}{l}\left(1 + \frac{\eta_1}{2\pi k}\right)\left(1 + \frac{p\eta_1}{q2\pi l} + \mathcal{O}(\eta_1/l^2)\right) \\
&= \frac{k}{l} + \frac{1}{2\pi}\left(\frac{1}{l} + \frac{kp}{l^2 q}\right)\eta_1(1 + \mathcal{O}(\eta_1/l)).
\end{aligned}$$

This means that k/l is close to γ and by applying (3.10) it follows that

$$\eta_1 \geq \frac{d'}{|l|}$$

for some constant $d' > 0$. Consequently, we obtain $\sigma \geq d''/l^2$ (for some constant $d'' > 0$) which translates directly to $\sigma > c/\tau^2$ for some positive constant c.

$\square$

Next we consider the case of algebraic number p and q such that $\log(p)/\log(q)$ is irrational.

Lemma 3.2.11.

If $p, q \in]0, 1[$ are positive algebraic numbers with $p + q = 1$ and the property that $\log(p)/\log(q)$ is irrational. Then for every solution $s \neq -1$ of the equation

$$1 - p^{-s} - q^{-s} = 0$$

we have

$$\Re(s) > \frac{D}{(\Im(s))^{2C}} - 1$$

with effectively computable positive constants C, D.

The classical theorem of Gelfond-Schneider says that $\gamma = \log(p)/\log(q)$ is irrational for algebraic numbers p and q then γ is transcendental. Baker's theorem (see [3]) gives also effective bounds for Diophantine approximation of γ that will be used in the subsequent proof of Lemma 3.2.11. Before introducing Baker's theorem, let us recall that the *height* of an algebraic number is the maximum of the absolute values of the relatively prime integer coefficients in its minimal defining polynomial, while its *degree* is the degree of this polynomial.

Theorem 3.2.12 (Baker's Theorem).
Let $\gamma_1, \ldots, \gamma_n$ be non-zero algebraic numbers with degrees at most d and heights at most A. Further, $\beta_0, \beta_1, \ldots, \beta_n$ are algebraic numbers with degree at most d and heights at most B (≥ 2). Then for

$$\Lambda = \beta_0 + \beta_1 \log \gamma_1 + \ldots + \beta_n \log \gamma_n$$

we have either $\Lambda = 0$ or $|\Lambda| \geq B^{-C}$, where C is an effectively computable number depending only on n, d, and A.

Proof of Lemma 3.2.11.
We apply Theorem 3.2.12 to the algebraic numbers $\gamma_1 = p$ and $\gamma_2 = q$ and to the integers $\beta_0 = 0$, $\beta_1 = l$, and $\beta_2 = -k$. Then $B = \max\{|k|, |l|\}$. W.l.o.g. we may assume that $p > q$ which assures that we only have to consider cases with $|k| \leq |l|$. Thus

$$|l \log p - k \log q| > B^{-C}$$

and consequently

$$\left| \frac{\log p}{\log q} - \frac{k}{l} \right| > \left(\frac{1}{\log q} \right) \frac{B^{-C}}{l} > \left(\frac{1}{\log q} \right) \frac{1}{l^{1+C}}, \tag{3.14}$$

where C is effectively computable.

By using (3.14) instead of (3.10) in the proof of Lemma 3.2.7 we easily complete the proof of Lemma 3.2.11.

$\square$

Differentiating Asymptotic Expansions

For our analysis, we need a Tauberian theorem for the Mellin transform. A classical result in this direction is the following (cf. [44], [33]).

Theorem 3.2.13.

Suppose that $f(v)$ is a monotone function for $v \geq 0$ such that

$$F(v) = \int_0^v f(w)dw$$

is asymptotically given by

$$F(v) \sim \frac{v^{\lambda+1}}{(\lambda+1)} \quad as \quad v \to \infty,$$

for some $\lambda > -1$. Then

$$f(v) \sim v^\lambda \quad as \quad v \to \infty.$$

We make this result more precise in the next lemma.

Lemma 3.2.14.

Suppose that $f(v)$ is a non-negative increasing function for $v \geq 0$. Assume that

$$F(v) = \int_0^v f(w)dw$$

has the asymptotic expansion

$$F(v) = \frac{v^{\lambda+1}}{\lambda+1}\left(1 + \mathcal{O}\left(g(v)\right)\right) \quad as \quad v \to \infty,$$

where $\lambda > -1$ and $g(v)$ is a decreasing function that tends to zero as $v \to \infty$. Then

$$f(v) = v^\lambda \left(1 + \mathcal{O}\left(g(v)^{\frac{1}{2}}\right)\right) \quad as \quad v \to \infty.$$

Proof.
By the assumption we have that there exist $v_0, c > 0$ such that for all $v \geq v_0$ we have

$$\left|F(v) - \frac{v^{\lambda+1}}{(\lambda+1)}\right| \leq c|g(v)|\frac{v^{\lambda+1}}{(\lambda+1)}.$$

Now, set $h = |g(v)|^{\frac{1}{2}}v$. By monotonicity, for $v \geq v_0$ we get

$$\frac{F(v+h) - F(v)}{h} = \frac{1}{h}\int_v^{v+h} f(w)dw \geq \frac{1}{h}\int_v^{v+h} f(v)dw = f(v).$$

Consequently we have

$$
\begin{aligned}
f(v) \;\le\; & \frac{F(v+h) - F(v)}{h} \\
\le\; & \frac{1}{h}\left(\frac{(v+h)^{\lambda+1}}{\lambda+1} - \frac{v^{\lambda+1}}{\lambda+1}\right) + \frac{1}{h}\left(c|g(v+h)|\frac{(v+h)^{\lambda+1}}{(\lambda+1)} + c|g(v)|\frac{v^{\lambda+1}}{(\lambda+1)}\right) \\
\le\; & \frac{1}{h(\lambda+1)}\left(v^{\lambda+1} + (\lambda+1)v^{\lambda}h + \mathcal{O}(v^{\lambda-1}h^2) - v^{\lambda+1}\right) + \mathcal{O}\left(|g(v)|\frac{v^{\lambda+1}}{h}\right) \\
=\; & v^{\lambda} + \mathcal{O}\left(v^{\lambda-1}h\right) + \mathcal{O}\left(|g(v)|\frac{v^{\lambda+1}}{h}\right) \\
=\; & v^{\lambda} + \mathcal{O}\left(v^{\lambda-1}|g(v)|^{\frac{1}{2}}v\right) + \mathcal{O}\left(|g(v)|\frac{v^{\lambda+1}}{|g(v)|^{\frac{1}{2}}v}\right) \\
=\; & v^{\lambda} + \mathcal{O}\left(v^{\lambda}|g(v)|^{\frac{1}{2}}\right).
\end{aligned}
$$

$\square$

Proof of Theorem 3.2.3

Set $v = \frac{1}{r}$ and denote by $A(v)$ the number of internal nodes (root node included) in Khodak's construction with parameter $r = 1/v$, i.e.

$$
A(v) = \sum_{x:P(x)\ge\frac{1}{v}} 1.
$$

Hence, the number of external nodes generated at the step corresponding to the parameter r is

$$
M_r = (m-1)A(v) + 1. \tag{3.15}
$$

The key relation is that that $A(v)$ satisfies the following recurrence, (see [20, Lemma 2])

$$
A(v) = \begin{cases} 0 & v < 1 \\ 1 + \sum\limits_{j=1}^{m} A(p_j v) & v \ge 1 \end{cases}. \tag{3.16}
$$

For the asymptotic analysis of $A(v)$ and consequently that of M_r we distinguish between the rational and the irrational case.

Rational case

If the $\log(1/p_j)$'s are rationally related then $A(v)$ is constant for $v \in [e^{\Lambda n}, e^{\Lambda(n+1)}[$ (for every integer n). Hence, it suffices to study the behaviour of the sequence

$G(n) = A(e^{\Lambda n})$ which verifies the recurrence

$$G(n) = \begin{cases} 0 & n < 0 \\ 1 + \sum\limits_{j=1}^{m} G(n - n_j) & n \geq 0 \end{cases}$$

where $n_j = \dfrac{\log\left(\frac{1}{p_j}\right)}{\Lambda}$ for $j = 1, \ldots, m$. The generating function $g(z) = \sum\limits_{n \geq 0} G(n)z^n$ is then given by

$$\begin{aligned} g(z) &= \sum_{n \geq 0} G(n)z^n = \sum_{n \geq 0} \left(1 + \sum_{j=1}^{m} G(n - n_j) \right) z^n \\ &= \sum_{n \geq 0} z^n + \sum_{n \geq 0} \sum_{j=1}^{m} G(n - n_j)z^n \\ &= \frac{1}{1 - z} + \sum_{j=1}^{m} \sum_{n = -n_j}^{\infty} G(n)z^{n + n_j} \\ &= \frac{1}{1 - z} + \sum_{j=1}^{m} z^{n_j} \left(\sum_{n \geq 0} G(n)z^n \right) \\ &= \frac{1}{1 - z} + g(z) \sum_{j=1}^{m} z^{n_j} \end{aligned}$$

and so

$$g(z) = \frac{1}{(1 - z)f(z)},$$

where $f(z) = 1 - z^{n_1} + \cdots - z^{n_m}$. By Definition 3.2.2, it follows that $e^{-\Lambda}$ is a positive real root of f. Moreover, it is proved in [15] that if we denote by $\omega_1, \ldots, \omega_h$ all the other different roots with multiplicities μ_i of f, then $|\omega_i| > e^{-\Lambda}$ for $i = 1, \ldots, h$. (Here we use the assumption that $n_1, \ldots, n_m$ are coprime). Therefore, we have by partial fraction decomposition

$$\begin{aligned} g(z) &= \frac{a}{1 - z} + \frac{b}{z - e^{-\Lambda}} + \frac{c_1^{(1)}}{z - \omega_1} + \cdots + \frac{c_{\mu_1}^{(1)}}{(z - \omega_1)^{\mu_1}} + \cdots \\ &+ \frac{c_1^{(h)}}{z - \omega_h} + \cdots + \frac{c_{\mu_h}^{(h)}}{(z - \omega_h)^{\mu_h}} \\ &= \frac{a}{1 - z} + \frac{-be^{\Lambda}}{1 - ze^{\Lambda}} + \frac{-c_1^{(1)}\omega_1^{-1}}{1 - z\omega_1^{-1}} + \cdots + \frac{-c_{\mu_1}^{(1)}\omega_1^{-\mu_1}}{(1 - z\omega_1^{-1})^{\mu_1}} + \cdots \\ &+ \frac{-c_1^{(h)}\omega_h^{-1}}{1 - z\omega_h^{-1}} + \cdots + \frac{-c_{\mu_h}^{(h)}\omega_h^{-\mu_h}}{(1 - z\omega_h^{-1})^{\mu_h}} \end{aligned}$$

for certain constants $a, b, c_1^{(i)}, \ldots, c_{\mu_i}^{(i)}$ with $i = 1, \ldots, h$. By expanding in series we have

$$g(z) = \sum_{n=0}^{\infty} \left(a - (be^{\Lambda})e^{\Lambda n} + \sum_{i=1}^{h} P_i(n)\omega_i^{-n} \right) z^n,$$

where P_i are polynomials of degree smaller than μ_i. Consequently, the following holds

$$G(n) = a - (be^{\Lambda})e^{\Lambda n} + \sum_{i=1}^{h} P_i(n)\omega_i^{-n}.$$

It remains to determine the constants a, b.

$$\begin{aligned} a &= \lim_{z \to 1}(z - 1) \cdot g(z) \\ &= \lim_{z \to 1} \frac{(z - 1)}{(1 - z)(1 - z^{n_1} + \cdots - z^{n_m})} \\ &= -\frac{1}{1 - m}, \end{aligned}$$

and

$$\begin{aligned} b &= \lim_{z \to e^{-\Lambda}}(z - e^{-\Lambda}) \cdot g(z) \\ &= \lim_{z \to e^{-\Lambda}} \frac{(z - e^{-\Lambda})}{(1 - z)(1 - z^{n_1} + \cdots - z^{n_m})} \\ &= \frac{1}{1 - e^{-\Lambda}} \lim_{z \to e^{-\Lambda}} \left(\frac{1}{-n_1 z^{n_1 - 1} + \cdots - n_m z^{n_m - 1}} \right) \\ &= \frac{-\Lambda e^{-\Lambda}}{(1 - e^{-\Lambda})(n_1 e^{n_1 \Lambda} + \cdots + n_m e^{n_m \Lambda})} \\ &= \frac{-\Lambda e^{-\Lambda}}{H(1 - e^{-\Lambda})}, \end{aligned}$$

where in the last equality we used that $n_j = \dfrac{\log\left(\frac{1}{p_j}\right)}{\Lambda}$ for $j = 1, \ldots, m$ and the definition of entropy. Hence, it follows that

$$G(n) = \frac{\Lambda e^{\Lambda n}}{H(1 - e^{-\Lambda})} + \sum_{i=1}^{h} P_i(n)\omega_i^{-n} - \frac{1}{m - 1}.$$

Note that in view of (3.15) the constant term $-1/(m - 1)$ disappears when we translate the asymptotics of $G(n)$ to M_r. Next we study the error term (without the constant term $-1/(m - 1)$) in more detail. W.l.o.g. we can assume that $\omega_1, \ldots, \omega_k$ (with $k \leq h$) are those roots of $f(z)$ with smallest modulus

$$|\omega_i| = e^{-\Lambda(1 - \eta)} \tag{3.17}$$

(for some $\eta > 0$) such that $P_i \neq 0$ and the degrees of P_i are maximal and all equal to $d \geq 0$, for $1 \leq i \leq k$. This means that the difference between $G(n)$ and the asymptotic leading term is bounded by

$$\delta(n) = \left| G(n) - \frac{\Lambda e^{\Lambda n}}{H(1 - e^{-\Lambda})} + \frac{1}{m-1} \right| = \left| \sum_{i=1}^{h} P_i(n) \omega_i^{-n} \right| \leq C n^d e^{\Lambda(1-\eta)n}$$

for some constant $C > 0$. More precisely $\delta(n)$ can be written as

$$\delta(n) = \left| n^d \sum_{i=1}^{k} \widetilde{c}_i \, \omega_i^{-n} \right| + \mathcal{O}\left(n^{d-1} e^{\Lambda(1-\eta)n} \right),$$

with complex numbers $\widetilde{c}_i \neq 0$, $1 \leq i \leq k$. Since all roots of $f(z)$ are either real or appear in conjugate pairs of complex numbers we can rewrite the sum $\sum_{i=1}^{k} \widetilde{c}_i \, \omega_i^{-n}$ to

$$n^d e^{\Lambda(1-\eta)n} \sum_{i=1}^{k'} c_i' \cos(2\pi\theta_i n + \alpha_i)$$

with real numbers $c_i' \neq 0$, $1 \leq i \leq k'$. By Lemma 3.2.4 it follows that there exists $\delta > 0$ and infinitely many n such that $\left| \sum_{i=1}^{k'} c_i' \cos(2\pi\theta_i n + \alpha_i) \right| \geq \delta$. This shows that

$$\delta(n) \geq C' n^d e^{\Lambda(1-\eta)n}$$

for infinitely many n and some constant $C' > 0$. This means that the error term in (3.7) is optimal.

Finally, we have

$$G(n) = \frac{\Lambda e^{\Lambda n}}{H(1 - e^{-\Lambda})} - \frac{1}{m-1} + \mathcal{O}\left(n^d e^{\Lambda(1-\eta)n} \right)$$

for some $\eta > 0$. Obviously, this implies the representation (3.7) of Theorem 3.2.3. In fact, since $A(v) = G\left(\left[\frac{\log v}{\Lambda} \right] \right)$ we have

$$
\begin{aligned}
A(v) &= \frac{\Lambda}{H(1 - e^{-\Lambda})} \cdot e^{\Lambda\left[\frac{\log v}{\Lambda} \right]} - \frac{1}{m-1} + \mathcal{O}\left(\left[\frac{\log v}{\Lambda} \right]^d e^{\Lambda(1-\eta)\left[\frac{\log v}{\Lambda} \right]} \right) \\
&= \frac{\Lambda v}{H(1 - e^{-\Lambda})} \cdot e^{-\Lambda\left\{ \frac{\log v}{\Lambda} \right\}} - \frac{1}{m-1} + \mathcal{O}\left((\log v)^d v^{(1-\eta)} \right)
\end{aligned}
$$

and so (3.15) implies that

$$M_r = \frac{(m-1)}{rH} \cdot \frac{\Lambda}{(1 - e^{-\Lambda})} e^{-\Lambda\left\{ \frac{\log\left(\frac{1}{r} \right)}{\Lambda} \right\}} + \mathcal{O}\left((\log r)^d r^{-(1-\eta)} \right).$$

Irrational case

The analysis in the irrational case is much more involved. Instead of using power series we use the Mellin transform of the function $A(v)$ (see [63]), i.e.

$$A^*(s) = \int_0^\infty A(v) v^{s-1}\, dv.$$

By using the fact that the Mellin transform of $A(av)$ is $a^{-s} A^*(s)$, a simple analysis of recurrence (3.16) reveals that the Mellin transform $A^*(s)$ of $A(v)$ is given by

$$A^*(s) = \frac{-1}{s\left(1 - p_1^{-s} - \cdots - p_m^{-s}\right)}, \qquad \Re(s) < -1. \tag{3.18}$$

In fact, for $\Re(s) < -1$ we have that

$$
\begin{aligned}
A^*(s) &= \int_0^\infty A(v) v^{s-1} dv = \int_1^\infty \left(1 + \sum_{j=1}^m A(p_j v)\right) v^{s-1} dv \\
&= \int_1^\infty v^{s-1} dv + \sum_{j=1}^m \int_1^\infty A(p_j v) v^{s-1} dv \\
&= -\frac{1}{s} + \sum_{j=1}^m p_j^{-s} \int_1^\infty A(v) v^{s-1} dv \\
&= -\frac{1}{s} + \left(\sum_{j=1}^m p_j^{-s}\right) \int_0^\infty A(v) v^{s-1} dv \\
&= -\frac{1}{s} + \left(\sum_{j=1}^m p_j^{-s}\right) A^*(s),
\end{aligned}
$$

which implies the relation (3.18).

In order to find asymptotics of $A(v)$ as $v \to \infty$ one can directly use the Tauberian theorem (for the Mellin transform) by Wiener-Ikehara [44, Theorem 4.1]. For this purpose we have to check that $s_0 = -1$ is the only (polar) singularity on the line $\Re(s) = -1$ and that $(s+1)A^*(s)$ can be analytically extended to a region that contains the line $\Re(s) = -1$. However, in the irrational case this follows by a lemma of Schachinger [58]. In particular, one finds

$$A(v) \sim \frac{v}{H}$$

but this procedure does not provide any information about the error term.

For making our presentation as simple as possible we will restrict ourselves to the case $m = 2$ and we will also assume certain conditions on the Diophantine properties

of the irrational number
$$\gamma = \frac{\log p_1}{\log p_2}.$$
We use the simplified notation $p = p_1$ and $q = p_2$.

The principle idea to obtain error terms for $A(v)$, in this case, is using the formula for the inverse Mellin transfrom
$$A(v) = \frac{1}{2\pi i} \lim_{T \to \infty} \int_{\sigma-iT}^{\sigma+iT} A^*(s)v^{-s}ds, \qquad \sigma < -1, \tag{3.19}$$
and shifting the line of integration to the right. Of course, all polar singularities of $A^*(s)$, which are given by the solutions of the equation $p^{-s} + q^{-s} = 1$ and $s = 0$, give rise to a polar singularity of $A(v)$. Unfortunately, the order of magnitude of $A^*(s)$ is $\mathcal{O}(1/s)$. Hence the integral in (3.19) is not absolutely convergent. Therefore, it is convenient to smooth the problem and to study the function
$$A_1(v) = \int_0^v A(w)\, dw,$$
which is given by
$$
\begin{aligned}
A_1(v) &= \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} A^*(s) \int_0^v w^{-s}\, dw\, ds \\
&= \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} A^*(s) \cdot \frac{v^{-s+1}}{1-s}\, ds \\
&= \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{v^{-s+1}}{s(s-1)(1-p^{-s}-q^{-s})}\, ds, \qquad \sigma < -1.
\end{aligned}
$$
By Proposition 3.2.6 we know that all zeros of the equation $p^{-s} + q^{-s} = 1$ that are different from -1 satisfy $-1 < \Re(s) \le \sigma_0$ for some σ_0. Furthermore, there exists $\tau > 0$ such that in each box of the form
$$B_k = \{s \in \mathbb{C} : -1 < \Re(s) \le \sigma_0, \ (2k-1)\tau \le \Im(s) < (2k+1)\tau\}, \quad k \in \mathbb{Z} \setminus \{0\},$$
there is precisely one zero of $p^{-s} + q^{-s} = 1$ that we denote by s_k.

Now, in the evaluation of $A_1(v)$ we shift the line of integration to the right, namely to $\Re(s) = \sigma_1$ with $\sigma_1 > \max\{\sigma_0 + 1, 1\}$. Let us denote by $\mathcal{S}$ the set of all the singularities $s' \in \mathbb{C}$ of the integrand function $g(s) := A^*(s)\frac{v^{-s+1}}{1-s}$ such that $\sigma \le \Re(s') \le \sigma_1$, i.e.
$$\mathcal{S} = \{-1,\ 0,\ 1,\ s_k \text{ for } k \in \mathbb{Z} \setminus \{0\}\}.$$
Then by applying Cauchy's residue theorem, we get
$$A_1(v) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} g(s)\, ds = -\sum_{s' \in \mathcal{S}} Res\left(g(s),\ s = s'\right) + \frac{1}{2\pi i} \int_{\sigma_1-i\infty}^{\sigma_1+i\infty} g(s)\, ds. \tag{3.20}$$

So, we have to consider the following residues

- $Res\left(g(s),\ s=-1\right)=-\frac{v^2}{2H}$;

- $Res\left(g(s),\ s=s_k\right)=\frac{v^{1-s_k}}{s_k(s_k-1)\left(p^{-s_k}\log\left(\frac{1}{p}\right)+q^{-s_k}\log\left(\frac{1}{q}\right)\right)}$, for $k\in\mathbb{Z}\setminus\{0\}$;

- $Res\left(g(s),\ s=0\right)=v$;

- $Res\left(g(s),\ s=1\right)=\frac{1}{1-p^{-1}-q^{-1}}$.

Hence, by collecting all residues and using (3.20) we obtain

$$A_1(v)=\frac{v^2}{2H}-\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{v^{1-s_k}}{s_k(s_k-1)H(s_k)}-v-\frac{1}{1-p^{-1}-q^{-1}}$$
$$+\frac{1}{2\pi i}\int_{\sigma_1-i\infty}^{\sigma_1+i\infty}\frac{v^{-s+1}}{s(s-1)(1-p^{-s}-q^{-s})}\,ds,$$

where we put $H(s)=p^{-s}\log\left(\frac{1}{p}\right)+q^{-s}\log\left(\frac{1}{q}\right)$.

Now, it is easy to see that the integral can be estimated by

$$\frac{1}{2\pi i}\int_{\sigma_1-i\infty}^{\sigma_1+i\infty}\frac{v^{-s+1}}{s(s-1)(1-p^{-s}-q^{-s})}\,ds=\mathcal{O}\left(v^{-\sigma_1+1}\right).\qquad(3.21)$$

In fact, we can write

$$\frac{1}{2\pi i}\int_{\sigma_1-i\infty}^{\sigma_1+i\infty}\frac{v^{-s+1}}{s(s-1)(1-p^{-s}-q^{-s})}\,ds$$
$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{v^{-\sigma_1-it+1}}{(\sigma_1+it)(\sigma_1+it-1)(1-p^{-\sigma_1-it}-q^{-\sigma_1-it})}\,dt.$$

Since

$$\left|\frac{v^{-\sigma_1-it+1}}{(\sigma_1+it)(\sigma_1+it-1)(1-p^{-\sigma_1-it}-q^{-\sigma_1-it})}\right|$$
$$=\frac{|v^{-\sigma_1+1}|\cdot|v^{-it}|}{\sqrt{\sigma_1^2+t^2}\cdot\sqrt{(\sigma_1-1)^2+t^2}|1-p^{-\sigma_1-it}-q^{-\sigma_1-it}|}$$
$$\leq c\frac{v^{-\sigma_1+1}}{(1+t^2)}$$

for some constant $c>0$, then we have

$$\left|\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{v^{-\sigma_1-it+1}}{(\sigma_1+it)(\sigma_1+it-1)(1-p^{-\sigma_1-it}-q^{-\sigma_1-it})}\,dt\right|\leq\frac{c}{2\pi}\int_{-\infty}^{\infty}\left(\frac{v^{-\sigma_1+1}}{1+t^2}\right)\,dt$$
$$=\frac{c}{2}\cdot v^{-\sigma_1+1}$$

and so the relation (3.21) holds. Hence, we just have to deal with the sum of residues

$$\sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{v^{1-s_k}}{s_k(s_k - 1)H(s_k)}. \tag{3.22}$$

First, let us show that there exists $\delta > 0$ such that $|H(s_k)| \geq \delta$ for all $k \in \mathbb{Z} \setminus \{0\}$. Without loss of generality we can assume that $q < p$ and so $\frac{\log q}{\log p} < 1$. Since $s_k \in \mathcal{S}$ we have that

$$
\begin{aligned}
H(s_k) &= p^{-s_k} \log\left(\frac{1}{p}\right) + q^{-s_k} \log\left(\frac{1}{q}\right) \\
&= (1 - q^{-s_k}) \log\left(\frac{1}{p}\right) + q^{-s_k} \log\left(\frac{1}{q}\right) \\
&= \log\left(\frac{1}{p}\right)\left(1 - q^{-s_k}\left(1 - \frac{\log q}{\log p}\right)\right).
\end{aligned}
$$

So it follows that

$$|H(s_k)| = \log\left(\frac{1}{p}\right)\left|1 - q^{-s_k}\left(1 - \frac{\log q}{\log p}\right)\right| \geq \log\left(\frac{1}{p}\right)\left|1 - |q^{-s_k}| \cdot \left|1 - \frac{\log q}{\log p}\right|\right|.$$

Therefore, since $\left(1 - \frac{\log q}{\log p}\right) < 1$ and $\Re(s_k) < 1$, there exists $c_0 > 0$ such that

$$|H(s_k)| \geq c_0 \log\left(\frac{1}{p}\right) = \delta.$$

Thus, we do not have to care about this factor in (3.22) .

Next assume that γ is a badly approximable irrational number. Here Lemma 3.2.7 shows that all zeros $s_k \in \mathcal{S}$ satisfy $\Re(s_k) > -1 + c/\Im(s_k)^2$ for some constant $c > 0$. Hence it follows that $\Re(s_k) > -1 + c_1/k^2$ for some constant $c_1 > 0$ and we can estimate the sum of residues by

$$
\begin{aligned}
\left|\sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{v^{1-s_k}}{s_k(s_k-1)H(s_k)}\right| &\leq \left|\sum_{0<|k|\leq K} \frac{v^{1-s_k}}{s_k(s_k-1)H(s_k)}\right| + \left|\sum_{|k|>K} \frac{v^{1-s_k}}{s_k(s_k-1)H(s_k)}\right| \\
&\leq \sum_{0<|k|\leq K} \frac{v^{1-\Re(s_k)}}{|\Re(s_k)(\Re(s_k)-1)|\,\delta} + \sum_{|k|>K} \frac{v^{1-\Re(s_k)}}{|\Re(s_k)(\Re(s_k)-1)|\,\delta} \\
&\leq C_1 v^{2-c_1/K^2} \sum_{0<|k|\leq K} \frac{1}{k^2} + C_2 v^2 \sum_{|k|>K} \frac{1}{k^2} \\
&\leq C_3 v^2 \left(v^{-c_1/K^2} + \frac{1}{K}\right).
\end{aligned}
$$

where C_1, C_2, C_3 are appropriate positive constants.

Thus, by choosing $K = \sqrt{c_1(\log v)/(\log\log v)}$, we obtain the upper bound

$$\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{v^{1-s_k}}{s_k(s_k-1)H(s_k)} = \mathcal{O}\left(v^2\frac{\sqrt{\log\log v}}{\sqrt{\log v}}\right)$$

and consequently

$$A_1(v) = \frac{v^2}{2H}\left(1+\mathcal{O}\left(\frac{\sqrt{\log\log v}}{\sqrt{\log v}}\right)\right).$$

Finally, by an application of Lemma 3.2.14, the previous relation implies

$$A(v) = \frac{v}{H}\left(1+\mathcal{O}\left(\frac{(\log\log v)^{1/4}}{(\log v)^{1/4}}\right)\right).$$

Similarly, when p and q are algebraic, we deal with the case when all solutions of the equation $p^{-s}+q^{-s}=1$ (that are different from -1) satisfy $\Re(s_k) > -1 + \frac{D}{\Im(s_k)^{2C}}$ for some positive constants C, D (see Lemma 3.2.11). Then with the same procedure as above we get

$$\left|\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{v^{1-s_k}}{s_k(s_k-1)H(s_k)}\right| \leq C_4 v^2\left(v^{-c_2K^{-2C}} + \frac{1}{K}\right)$$

for some constant $C_4 > 0$. Hence, if we choose $K = (c_2(\log v)/(\log\log v))^{1/(2C)}$, we obtain (after a second application of Lemma 3.2.14)

$$A(v) = \frac{v}{H}\left(1 + \frac{(\log\log v)^{\kappa}}{(\log v)^{\kappa}}\right).$$

where $\kappa = \frac{1}{4C}$. This completes the proof of Theorem 3.2.3.

$\square$

3.2.3   Discrepancy bounds in the rational case

First, we want to study the rational case. So in this subsection we are going to consider a partition ρ of $[0,1]$ consisting of m intervals of lengths $p_1,\ldots,p_m$ such that $\log\left(\frac{1}{p_1}\right),\ldots,\log\left(\frac{1}{p_m}\right)$ are rationally related.

By Theorem 3.2.3 we know that asymptotically

$$M_{r_n} = \frac{c'}{r_n} + \mathcal{O}\left((\log r_n)^d r_n^{-(1-\eta)}\right), \quad r_n = e^{-\Lambda n}, \tag{3.23}$$

for some $\eta > 0$ and some integer $d \geq 0$, where $c' = \frac{(m-1)\Lambda}{H(1-e^{-\Lambda})}$ and the error term is optimal. Recall also that $k(n) = M_{r_n}$, which gives an asmyptotic expansion for $k(n)$ of the form

$$k(n) \sim \frac{(m-1)\Lambda}{H(1-e^{-\Lambda})}\, e^{\Lambda n}. \tag{3.24}$$

Theorem 3.2.15.

Suppose that the lengths of the intervals of a partition ρ are $p_1, \ldots, p_m$ and suppose that $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ are rationally related. Furthermore, let $\eta > 0$ and $d \geq 0$ be given as in Theorem 3.2.3. Then the discrepancy of the sequence of partitions $(\rho^n \omega)$ is bounded by

$$D_n = \begin{cases} \mathcal{O}\left((\log k(n))^d k(n)^{-\eta}\right) & \text{if } 0 < \eta < 1, \\ \mathcal{O}\left((\log k(n))^{d+1} k(n)^{-1}\right) & \text{if } \eta = 1, \\ \mathcal{O}\left(k(n)^{-1}\right) & \text{if } \eta > 1. \end{cases} \tag{3.25}$$

Moreover, there exist $\delta > 0$ and infinitely many n such that

$$D_n \geq \begin{cases} \delta\,(\log k(n))^d k(n)^{-\eta} & \text{if } 0 < \eta < 1, \\ \delta\,(\log k(n))^d k(n)^{-1} & \text{if } \eta = 1, \\ \delta\,k(n)^{-1} & \text{if } \eta > 1. \end{cases} \tag{3.26}$$

Proof.

For notational convenience we set

$$\Delta_n = \sup_{0 \leq y \leq 1} \left| \sum_{i=1}^{k(n)} \chi_{[0,y[}\left(t_i^{(n)}\right) - k(n)y \right|,$$

where $t_i^{(n)}$ are the points defining the partition $\rho^n \omega$. Then we have $D_n \leq 2\Delta_n / k(n)$, since $\Delta_n / k(n) = D_n^*$ and Theorem 1.1.10 holds.

Fix a step in the algorithm corresponding to a certain parameter r of the form $r = e^{-n\Lambda}$ for some integer $n \geq 0$, and consider an interval $A = [0, y[\subset [0, 1]$. We want to estimate the number of elementary intervals belonging to $\mathscr{E}_r$ which are contained in A. For this purpose, let us fix another parameter $\bar{r}$, of the form $\bar{r} = e^{-\bar{n}\Lambda}$ with an integer $0 \leq \bar{n} \leq n$, corresponding to a previous step in Khodak's construction. At this previous step, we have $M_{\bar{r}}$ intervals I_j generated by the construction. Now, the lenghts of the intervals I_j are given by $\lambda(I_j)$ and we have that

$$p_{min}\bar{r} \leq \lambda(I_j) < \bar{r}, \quad \text{for} \quad j = 1, \ldots, M_{\bar{r}}, \tag{3.27}$$

since the lengths of the intervals $\mathscr{E}_{\bar{r}}$ correspond to the values $P(y)$ of the external nodes y in $\mathcal{E}(\bar{r})$.

Suppose that precisely the first h of these intervals I_j are contained in A, so $U = I_1 \cup \ldots \cup I_h \subset A$. Now, we want to estimate the number of elementary intervals in $\mathscr{E}_r$ contained in each I_j. Khodak's construction shows that this equals precisely the number of external nodes in the subtree of the node x that is related to the interval I_j.

An important feature of Khodak's construction is that subtrees of $\mathcal{T}(r)$ rooted at an internal node $x \in \mathcal{I}(r)$ are parts of a self-similar infinite tree and therefore they are constructed in the same way as the whole tree. So, one just has to replace r by $\frac{r}{P(x)}$. Hence, by using this remark in (3.23), the number N_{I_j} of subintervals of I_j (corresponding to the value r) equals

$$N_{I_j} = M_{\frac{r}{\lambda(I_j)}} = \frac{c'}{r}\lambda(I_j) + \mathcal{O}\left(|\log r|^d \frac{\lambda(I_j)^{1-\eta}}{r^{1-\eta}}\right).$$

Therefore, we have that the number N_U of elementary intervals in $\mathscr{E}_r$ contained in U is

$$N_U = N_{I_1} + \ldots + N_{I_h} = \frac{c'}{r}(\lambda(I_1) + \ldots + \lambda(I_h)) + \mathcal{O}\left(\frac{|\log r|^d}{r^{1-\eta}}\sum_{j=1}^{h}\lambda(I_j)^{1-\eta}\right).$$

By using (3.27) and the fact that $h \leq M_{\overline{r}} = \mathcal{O}(1/\overline{r})$ we obtain

$$N_U = \frac{c'}{r}(\lambda(I_1) + \ldots + \lambda(I_h)) + \mathcal{O}\left(|\log r|^d \frac{h\overline{r}^{(1-\eta)}}{r^{(1-\eta)}}\right)$$

$$= \frac{c'}{r}(\lambda(I_1) + \ldots + \lambda(I_h)) + \mathcal{O}\left(|\log r|^d \frac{\overline{r}^{(-\eta)}}{r^{(1-\eta)}}\right).$$

Since the total number of intervals equals $M_r = c'/r + \mathcal{O}(|\log r|^d r^{-1+\eta})$ it follows that

$$N_U - M_r\lambda(U) = \mathcal{O}\left(|\log r|^d \frac{\overline{r}^{(-\eta)}}{r^{(1-\eta)}}\right) + \mathcal{O}\left(\frac{|\log r)|^d}{r^{1-\eta}}\right) = \mathcal{O}\left(|\log r|^d \frac{\overline{r}^{(-\eta)}}{r^{(1-\eta)}}\right).$$

Since $N_A - M_r\lambda(A) = (N_U - M_r\lambda(U)) + (N_{A\setminus U} - M_r\lambda(A \setminus U))$ it remains to study the difference

$$N_{A\setminus U} - M_r\lambda(A \setminus U) = \left(N_{A\setminus U} - M_{r/\lambda(I_{h+1})}\frac{\lambda(A \setminus U)}{\lambda(I_{h+1})}\right)$$
$$+ \left(M_{r/\lambda(I_{h+1})}\frac{\lambda(A \setminus U)}{\lambda(I_{h+1})} - M_r\lambda(A \setminus U)\right).$$

The second term can be directly estimated by

$$\left|M_{r/\lambda(I_{h+1})}\frac{\lambda(A \setminus U)}{\lambda(I_{h+1})} - M_r\lambda(A \setminus U)\right| = \mathcal{O}\left(|\log r|^d \frac{\overline{r}^{(1-\eta)}}{r^{(1-\eta)}}\right),$$

whereas the first term is bounded by

$$\left|N_{A\setminus U} - M_{r/\lambda(I_{h+1})}\frac{\lambda(A \setminus U)}{\lambda(I_{h+1})}\right| \leq \Delta_{n-\overline{n}}.$$

since $\frac{r}{\overline{r}} = e^{-\Lambda(n-\overline{n})}$.

Summing up and taking the supremum over all sets $A = [0, y[$, we obtain the recurrence relation

$$\Delta_n \leq \Delta_{n-\overline{n}} + \mathcal{O}\left(|\log r|^d \frac{\overline{r}^{(-\eta)}}{r^{(1-\eta)}}\right).$$

We now set $\overline{n} = 1$ and recall that $r = e^{-\Lambda n}$ and also $\overline{r} = e^{-\Lambda \overline{n}} = e^{-\Lambda}$. Thus, by the previous relation we get

$$\Delta_n \leq \Delta_{n-1} + \mathcal{O}\left(n^d e^{\Lambda n(1-\eta)}\right). \tag{3.28}$$

We distinguish between three cases.

1. $0 < \eta < 1$. In this case we get

$$\Delta_n = \mathcal{O}\left(\sum_{k \leq n} k^d e^{\Lambda k(1-\eta)}\right) = \mathcal{O}\left(n^d e^{\Lambda n(1-\eta)}\right).$$

By taking into account also the relation (3.24), it follows that

$$D_n \leq 2\frac{\Delta_n}{k(n)} = \mathcal{O}\left((\log k(n))^d \frac{k(n)^{1-\eta}}{k(n)}\right) \mathcal{O}\left((\log k(n))^d k(n)^{-\eta}\right).$$

2. $\eta = 1$. In this case we get

$$\Delta_n = \mathcal{O}\left(\sum_{k \leq n} k^d\right) = \mathcal{O}(n^{d+1})$$

and consequently

$$D_n \leq 2\frac{\Delta_n}{k(n)} = \mathcal{O}\left((\log k(n))^{d+1} k(n)^{-1}\right).$$

3. $\eta > 1$. Here we have

$$\Delta_n = \mathcal{O}\left(\sum_{k \leq n} k^d e^{-\Lambda k(\eta-1)}\right) = \mathcal{O}(1)$$

which rewrites to $D_n = \mathcal{O}\left(k(n)^{-1}\right)$.

In order to give a lower bound of the discrepancy it is sufficient to handle the case $0 < \eta \leq 1$. In fact, if $\eta > 1$ we just use the trivial lower bound $D_n \geq \frac{1}{k(n)}$ which meets the upper bound. For the remaining case $0 < \eta \leq 1$ we consider the interval

$A = [0, p_1[$. We also recall (see the proof of Theorem 3.2.3) that we can write M_r, for $r = r_n = e^{-\Lambda n}$, as

$$M_r = c' \, e^{\Lambda n} + \delta_n,$$

where δ_n has a representation of the form

$$\delta_n = n^d e^{\Lambda n(1-\eta)} \sum_{i=1}^{k} c_i \cos(2\pi\theta_i n + \alpha_i) + \mathcal{O}\left(n^{d-1} e^{\Lambda n(1-\eta)}\right).$$

Similarly to the above we obtain

$$
\begin{aligned}
N_A - M_r \lambda(A) &= M_{r/p_1} - M_r p_1 \\
&= c' \, e^{\Lambda n} p_1 + \delta_{n-n_1} - p_1 \left(c' \, e^{\Lambda n} + \delta_n\right) \\
&= \delta_{n-n_1} - p_1 \delta_n \\
&= n^d e^{\Lambda n(1-\eta)} \left(\sum_{i=1}^{k} c_i \cos(2\pi\theta_i n + \alpha_i - 2\pi\theta_i n_1) \right. \\
&\quad \left. - p_1 \sum_{i=1}^{k} c_i \cos(2\pi\theta_i n + \alpha_i) \right) + \mathcal{O}\left(n^{d-1} e^{\Lambda n(1-\eta)}\right).
\end{aligned}
$$

By applying Lemma 3.2.4 it follows that there exist $\delta > 0$ and infinitely many n with

$$|N_A - M_r \lambda(A)| \geq \delta n^d e^{\Lambda n(1-\eta)}.$$

Consequently

$$D_n \geq \frac{1}{M_r} |N_A - M_r \lambda(A)| \geq \frac{\delta n^d e^{\Lambda n(1-\eta)}}{e^{\Lambda n}} = \delta n^d e^{-\Lambda n \eta}$$

for some $\delta > 0$. This completes the proof of the lower bound (3.26).

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

3.2.4   Discrepancy bounds in the irrational case

As mentioned above, the case when $\log\left(\frac{1}{p_1}\right), \ldots, \log\left(\frac{1}{p_m}\right)$ are irrationally related is much more difficult to handle, since the error term in the asymptotic expansion for M_r is not explicit in general, (see (3.8) in Theorem 3.2.3). Nevertheless, we can provide upper bounds in some cases of interest.

Suppose that $m = 2$, set $p = p_1$ and $q = p_2$ and $\gamma = (\log p)/(\log q)$. First, let us show that

Proposition 3.2.16.

The number of intervals $k(n)$ of the partition $\rho^n \omega$ is asymptotically given by

$$k(n) \sim \frac{m-1}{H} \exp\left(\sqrt{2n \log\left(\frac{1}{p}\right) \log\left(\frac{1}{q}\right)}\right).$$

Proof.

Let r be the parameter in Khodak's construction that corresponds to the step n, then $M_r = k(n)$. By (3.8) in Theorem 3.2.3 we have that

$$k(n) \sim \frac{m-1}{H} \cdot \frac{1}{r}. \tag{3.29}$$

Note that there is a one-to-one correspondence between the probability of each node x, that is $P(x) = p^k q^l$ and the non-negative integral lattice points (k, ℓ). So the number n of steps corresponding to the value r is approximatively given by the cardinality of the set

$$
\begin{aligned}
\{x \in \mathcal{T}(r) : P(x) \geq r\} &= \{(k, l) : p^k q^l \geq r\} \\
&= \left\{(k, l) : \frac{1}{p^k q^l} \leq \frac{1}{r}\right\} \\
&= \left\{(k, l) : k \log\left(\frac{1}{p}\right) + l \log\left(\frac{1}{q}\right) \leq \log\left(\frac{1}{r}\right)\right\}
\end{aligned}
$$

Now, the equation $k \log p + \ell \log q = \log r$ has at most one solution in integer pairs (k, ℓ). Hence, we have

$$n \sim \frac{1}{2} \cdot \frac{\log\left(\frac{1}{r}\right)}{\log\left(\frac{1}{p}\right)} \cdot \frac{\log\left(\frac{1}{r}\right)}{\log\left(\frac{1}{q}\right)} = \frac{\left(\log\left(\frac{1}{r}\right)\right)^2}{2 \log\left(\frac{1}{p}\right) \log\left(\frac{1}{q}\right)}$$

and so

$$\frac{1}{r} \sim \exp\left(\sqrt{2n \log\left(\frac{1}{p}\right) \log\left(\frac{1}{q}\right)}\right).$$

The conclusion follows by using this relation in (3.29). $\qquad\square$

In Theorem 3.2.3 we have considered the case when γ is badly approximable and the case when p and q are algebraic. By using these results we can show the following theorem for the discrepancy in the irrational case.

Theorem 3.2.17.

Suppose that the lenghts of the intervals of a partition ρ of $[0, 1]$ are p and q and let

$\gamma = \frac{\log p}{\log q}$. *If $\gamma \notin \mathbb{Q}$ and it is badly approximable, then the discrepancy of $(\rho^n \omega)$ is bounded by*

$$D_n = \mathcal{O}\left(\left(\frac{\log\log\left(k(n)\right)}{\log\left(k(n)\right)}\right)^{\frac{1}{4}}\right), \quad \text{as } n \to \infty.$$

Furthermore, if p and q are algebraic and $\gamma \notin \mathbb{Q}$ then

$$D_n = \mathcal{O}\left(\left(\frac{\log\log\left(k(n)\right)}{\log\left(k(n)\right)}\right)^{\kappa}\right), \quad \text{as } n \to \infty,$$

where $\kappa > 0$ is an effectively computable constant (see Theorem 3.2.3).

Proof.

We use a procedure similar to the proof of Theorem 3.2.15 but now we consider the asymptotic expansion

$$M_r = \frac{c''}{r} + \mathcal{O}\left(\frac{1}{r}\left(\frac{\log\log\frac{1}{r}}{\log\frac{1}{r}}\right)^{\xi}\right)$$

where $c'' = (m-1)/H$. Moreover, we have that $\xi = \frac{1}{4}$ when γ is badly approximable, while $\xi = \kappa$ when p, q are algebraic (see Theorem 3.2.3).

Fix a step n in the algorithm corresponding to a certain parameter r and consider an interval $A = [0, y[\subset [0, 1]$. We want to estimate the number of elementary intervals belonging to $\mathscr{E}_r$ which are contained in A. For this purpose, let us fix another parameter $\bar{r}$, corresponding to a step $\bar{n}$ in Khodak's construction such that $0 \leq \bar{n} \leq n$. At this previous step, we have $M_{\bar{r}}$ intervals I_j generated by the construction. Now, the lenghts of the intervals I_j are given by $\lambda(I_j)$ and we have that the relation (3.27) holds.

Suppose that precisely the first h of these intervals I_j are contained in A, so $U = I_1 \cup \ldots \cup I_h \subset A$. We want to estimate the number of elementary intervals in $\mathscr{E}_r$ contained in I_j. Similarly to the rational case we have that the number N_{I_j} of subintervals of I_j (corresponding to the value r) equals

$$N_{I_j} = M_{\frac{r}{\lambda(I_j)}} = \frac{c''}{r}\lambda(I_j) + \mathcal{O}\left(\frac{\lambda(I_j)}{r}\left(\frac{\log\log\left(\frac{\lambda(I_j)}{r}\right)}{\log\left(\frac{\lambda(I_j)}{r}\right)}\right)^{\xi}\right).$$

By using (3.27) and the fact that $h \leq M_{\bar{r}} = \mathcal{O}\left(\frac{1}{\bar{r}}\right)$, it follows that

$$N_U - M_r\lambda(U) = \frac{c''}{r}(\lambda(I_1) + \ldots + \lambda(I_h)) + \mathcal{O}\left(\frac{h\overline{r}}{r}\left(\frac{\log\log\frac{\overline{r}}{r}}{\log\frac{\overline{r}}{r}}\right)^\xi\right)$$

$$- \frac{c''}{r}(\lambda(I_1) + \ldots + \lambda(I_h)) + \mathcal{O}\left(\frac{1}{r}\left(\frac{\log\log\frac{1}{r}}{\log\frac{1}{r}}\right)^\xi\right)$$

$$= \mathcal{O}\left(\frac{1}{r}\left(\frac{\log\log\frac{1}{r}}{\log\frac{1}{r}}\right)^\xi\right).$$

For the remaining interval $A \setminus U$ we use the bounds $N_{A\setminus U} \leq M_{r/\lambda(I_{h+1})} = \mathcal{O}(\overline{r}/r)$ and $\lambda(I_{h+1}) = \mathcal{O}(\overline{r})$ to end up with the upper bound

$$D_n = \mathcal{O}\left(\left(\frac{\log\log\frac{1}{r}}{\log\frac{1}{r}}\right)^\xi\right) + \mathcal{O}\left(\overline{r}\right).$$

Hence, by choosing

$$\overline{r} = \left(\frac{\log\log\frac{1}{r}}{\log\frac{1}{r}}\right)^\xi$$

we finally obtain

$$D_n = \mathcal{O}\left(\left(\frac{\log\log\frac{1}{r}}{\log\frac{1}{r}}\right)^\xi\right).$$

This completes the proof of the theorem.

$\square$

Note that the upper bounds for the discrepancy we obtained are worse than $k(n)^{-\beta}$ for any $\beta > 0$. Actually, it seems that we cannot do really better in the irrational case. This is due to the fact that $\liminf_{k\neq 0} \Re(s_k) = -1$ where s_k, $k \neq 0$, runs through all the zeros of the equation $p^{-s} + q^{-s} = 1$ different from $s_0 = -1$. Indeed, it seems that the continued fraction expansion of $\gamma = (\log p)/(\log q)$ could be used to obtain more explicit upper bounds. However, since they are all rather poor it is probably not worth working them out in detail.

Moreover, the case $m > 2$ is even more involved, as we can see by comparing with the discussion of [26]. In fact, in this paper the authors study the asymptotic structure of the main parameters of interest for digital trees. They represent collections of words over some finite alphabet, so the parsing tree constructed by Khodak's algorithm can be included in this class. In particular, in this paper digital trees are assumed to

be under the simplest of all probabilistic models; namely, the memoryless source, where letters of words are drawn independently according to a fixed distribution. If $\mathcal{A} = \{a_1, \ldots, a_m\}$ is the alphabet, the model is determined by the basic quantities $p_j = P(a_j)$ with $p_1 + \cdots + p_m = 1$. As it turned out by an analysis based on Mellin transform, quantifying the main paramters of digital trees (such us expected number of internal nodes, expected path lenght, etc.) is strongly dependent on the location of poles in the complex plane of the fundamental Dirichlet series associated with the p_j's, which is given by

$$\Delta(s) = \frac{1}{1 - p_1^s - \cdots - p_m^s}.$$

Neverthless, the results obtained in this paper relatively to the aperiodic case, which corresponds to our irrational one, shows clearly how the geometry of the set of poles of $\Delta(s)$ depends on the approximation properties of the ratios $\frac{\log p_i}{\log p_j}$ and therefore how much complicated is the study of this case.

3.3   Applications

In this section, we intend to present some examples and applications of the results of this chapter. In particular, we want to stress the application to fractals because the technique introduced in this chapter on $[0,1]$ allows to get discrepancy bounds for the elementary discrepancy of u.d. sequence of partitions of fractals belonging to a class wider than one considered in Section 2.2.

3.3.1   Kakutani's sequences

The procedure introduced in the Section 3.2 can be used to obtain bounds of the discrepancy for classical Kakutani's sequences of partitions. In fact, if we fix $\alpha \in]0,1[$ then the corresponding Kakutani's sequence of partitions of $[0,1]$ is constructed by successive α-refinements of the trivial partition $\omega = \{[0,1]\}$. So according to the notation used in the previous sections, in this case we have that $p_1 = \alpha$ and $p_2 = 1-\alpha$.

For a Kakutani's sequence of parameter α we have that $\log\left(\frac{1}{\alpha}\right)$ and $\log\left(\frac{1}{1-\alpha}\right)$ are rationally related if and only if $\frac{\log \alpha}{\log(1-\alpha)} \in \mathbb{Q}$ (see Definition 3.2.2).

Let us denote by $\alpha_{n,m}$ the unique solution in $]0,1[$ of the following equation

$$\frac{\log x}{\log(1-x)} = \frac{n}{m}$$

with $n, m \in \mathbb{N}$. Since the function

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

is continuous and strictly decreasing on $]0, 1[$ and attains all positive values, the countable set

$$\left\{ \alpha_{n,m} = f^{-1} \left(\frac{n}{m} \right), \ n, m \in \mathbb{N} \right\}$$

is dense in $]0, 1[$.

The density of the values of the parameter α for which $\log \left(\frac{1}{\alpha} \right)$ and $\log \left(\frac{1}{1-\alpha} \right)$ are rationally related shows that we have interesting bounds of the discrepancy for a countable set of Kakutani's sequences (see Theorem 3.2.15). On the other hand, there are much more values of α in $]0, 1[$ for which $\log \left(\frac{1}{\alpha} \right)$ and $\log \left(\frac{1}{1-\alpha} \right)$ are irrationally related and for which the discrepancy bounds are weaker. Neverthless, our technique allows to get quantitive results about the discrepancy of a large class of Kakutani's sequences, not known in the existing literature.

3.3.2 LS-sequences

LS-sequences are a special class of sequences of partitions constructed by successive ρ-refinements of the trivial partition $\omega = [0, 1]$. We have already introduced this sequences in Section 3.2, but let us recall their definition.

Definition 3.3.1.
Let us fix two positive integers L and S and let $0 < \alpha < 1$ be the real number given by the equation $L\alpha + S\alpha^2 = 1$. The LS-sequence is the sequence of partitions obtained by successive ρ-refinements of ω, when ρ consists of L subintervals of $[0, 1]$ of length α and S subintervals of length α^2.

For instance, if $L = S = 1$ then $\alpha = \frac{\sqrt{5}-1}{2}$ and we obtain the so-called *Kakutani-Fibonacci sequence*. This term was used first in [10], since this is a particular Kakutani's sequence and the sequence $(k(n))_{n \in \mathbb{N}}$ of the number of intervals of the n-th partition is the sequence of Fibonacci numbers. The Kakutani-Fibonacci sequence is important because it is the only Kakutani's sequence for which the exact discrepancy is known (apart from the trivial Kakutani's sequence with $\alpha = \frac{1}{2}$). Here we have $p_1 = \alpha$ and $p_2 = 1 - \alpha = \alpha^2$ and consequently, by using our techinique, we have

$$\log \left(\frac{1}{\alpha} \right) = n_1 \Lambda \quad \text{and} \quad \log \left(\frac{1}{\alpha^2} \right) = n_2 \Lambda$$

with $\Lambda = -\log \alpha$, $n_1 = 1$ and $n_2 = 2$. By following the lines of the proof of Theorem 3.2.3 and in particular (3.17) we can explicitely get the value of η. In fact, since the roots of the equation $1 - z - z^2 = 0$ are given by $z_1 = \frac{\sqrt{5}-1}{2} = \alpha = e^{-\Lambda}$ and

$z_2 = \frac{-\sqrt{5}-1}{2}$, it follows that $d = 0$ and

$$\eta = 1 + \frac{\log|z_2|}{\Lambda} = 1 + \frac{\log\left|\frac{-\sqrt{5}-1}{2}\right|}{-\log\left(\frac{\sqrt{5}-1}{2}\right)} = 2.$$

According to Theorem 3.2.15, this shows that the discrepancy is of the order $\mathcal{O}\left(1/k(n)\right)$ and therefore it is optimal.

In the general case set $m = L + S$. Of course we are in the rational case since $p_i = \alpha$ or $p_i = \alpha^2$ for $i = 1, \ldots, m$. More precisely, according to Definition 3.2.2, we have $\Lambda = \log(1/\alpha)$ and $n_i \in \{1, 2\}$ corresponding to $p_i = \alpha^{n_i}$. The zeros of the equation

$$1 - Lz - Sz^2 = 0$$

are given by $z_1 = \frac{-L+\sqrt{L^2+4S}}{2S} = \alpha$ and $z_2 = \frac{-L-\sqrt{L^2+4S}}{2S}$. Hence,

$$\eta = 1 + \frac{\log\left|\frac{-L-\sqrt{L^2+4S}}{2S}\right|}{\Lambda} = 1 + \frac{\log\left(\frac{L+\sqrt{L^2+4S}}{2S}\right)}{\Lambda}.$$

Consequently, we have $\eta < 1$ if and only if $\frac{L+\sqrt{L^2+4S}}{2S} < 1$ or if $S > L + 1$. Similarly we have $\eta = 1$ if and only if $S = L + 1$ and $\eta > 1$ if and only if $S < L + 1$. This is in perfect accordance with the results of Carbone [10]. The discrepancy bounds are (of course) also of the same kind.

3.3.3   Sequences related to Pisot numbers

Let us introduce a class of sequences having optimal discrepancy. This kind of sequences is related to Pisot numbers, so let us recall some classical definitions.

Definition 3.3.2.

- *An* algebraic number *is a number which is a root of a non-zero polynomial in one variable with rational (or equivalently, integer) coefficients.*

- *Given an algebraic number α, there is a unique monic polynomial with rational coefficients of least degree that has α as a root. This polynomial is called* minimal polynomial *of α.*

- *If an algebraic number α has its minimal polynomial of degree n, then the algebraic number is said to be of degree n.*

- *The* conjugates *of an algebraic number α are the other roots of its minimal polynomial.*

- *An* algebraic integer *is an algebraic number which is a root of a monic polynomial with integer coefficients.*

Definition 3.3.3.
A Pisot number β is an algebraic integer larger than 1 with the property that all its conjugates have modulus smaller than 1.

A prominent example of Pisot numbers are the real roots of a polynomial of the form

$$z^k - a_1 z^{k-1} - a_2 z^{k-2} - \cdots - a_k = 0, \tag{3.30}$$

where a_j are positive integers with $a_1 \geq a_2 \geq \cdots \geq a_k$, see [7]. In this case the polynomial in (3.30) is also irreducible over the rationals.

Suppose now that ρ is a partition of $m = a_1 + a_2 + \cdots + a_k$ intervals, where a_j intervals have length α^j, $1 \leq j \leq k$, $\alpha = 1/\beta$, and β is the Pisot number related to the polynomial (3.30). Note that we have

$$a_1 \alpha + a_2 \alpha^2 + \cdots + a_k \alpha^k = 1.$$

Since all conjugates of α have now modulus larger than 1 it follows that $\eta > 1$. This means that the order of magnitude of the discrepancy is optimal, namely $1/k(n)$. Moreover, we can note that LS-sequences are a special instance for $k = 2$, $a_1 = L$ and $a_2 = S$ with $L \geq S$.

Note that in the Pisot case all complex zeros of the polynomial are simple, since the polynomial is irreducible over the rationals. However, this is not necessarily true in less restrictive cases than Pisot numbers. For example, let $\alpha = 1/5$ and consider one interval of length $\alpha = 1/5$, 16 intervals of lengths $\alpha^2 = 1/25$ and 20 intervals of lengths $\alpha^3 = 1/125$. Since $\alpha + 16\alpha^2 + 20\alpha^3 = 1$ we have a proper partition ρ. Here the roots of the polynomial $z + 16z^2 + 20z^3 = 1$ are $z_1 = \alpha = 1/5$ and $z_2 = z_3 = -1/2$ (which is a double root). Hence, we obtain $\eta = 1 - (\log 2)/(\log 5) = 0.56932\ldots < 1$ and $d = 1$. Consequently the discrepancy is bounded by

$$D_n = \mathcal{O}((\log k(n)) \, k(n)^{-\eta}),$$

and this upper bound is optimal.

3.3.4 The rational case on fractals

The same procedure of ρ-refinements can be used to obtain an extension of the results introduced in Section 2.2 to fractals defined by similarities which do not have

the same ratio and which satisfy the OSC. In fact, we will describe an analogous of the method of successive ρ−refinements which allows to produce sequences of partitions on this new class of fractals. Actually, we will introduce a new correspondence between nodes of the tree associated to Khodak's algorithm and the subsets belonging to the partitions generated on the fractal.

Let $\varphi = \{\varphi_1, \ldots, \varphi_m\}$ be a system of m similarities on $\mathbb{R}^h$ which have ratios $c_1, \ldots, c_m \in]0, 1[$ respectively and which verify the OSC. Let F be the attractor of φ and let s be its Hausdorff dimension. Moreover, we will consider the normalized s-dimensional Hausdorff measure P on the fractal F, that is given by (2.3).

Start with a tree having a root node of probability 1, which corresponds to the fractal F, and m leaves corresponding to the m imagines of F through the m similarities, i.e. $\varphi_1(F), \ldots, \varphi_m(F)$. The probability of each node is given by the probability of the corresponding subset, that is $p_i = P(\varphi_i(F)) = c_i^s$. At each iteration we select the leaves having the highest probability and grow m children out of each of them. On the fractal this corresponds to apply successively the m similarities only to those subsets of F having the highest probability at this certain step. By iterating this procedure we obtain a tree associated to the sequence of partitions on the fractal F, which is the same tree generated by Khodak's algorithm.

Let us denote by (π_n) the sequence of partitions of F constructed by this technique, i.e.

$$\pi_n = \left\{ \psi_{j_{k(n)}} \psi_{j_{(k(n)-1)}} \cdots \ \psi_{j_1}(F) : j_1, \ldots, j_{k(n)} \in \{1, \ldots, m\} \right\}.$$

where $k(n)$ is the number of sets constructed at the step n.

Let us denote by $\mathscr{E}_n$ the collection of the $k(n)$ sets E_i^n belonging to the partition π_n and by $\mathscr{E}$ the union of the families $\mathscr{E}_n$, by varying n. The sets of the class $\mathscr{E}$ are called *elementary sets* and they are exactly the ones defined in Subsection 2.2.1.

In Lemma 2.2.2 and Lemma 2.2.1, the class $\mathscr{E}$ is proved to be determining and consisting of P-continuity sets. Now, if we choose a point $t_i^{(n)}$ in each $E_i^n \in \pi_n$, we can consider the elementary discrepancy of this set of points on the fractal, i.e.

$$D_n^{\mathscr{E}} = \sup_{E \in \mathscr{E}} \left| \frac{1}{k(n)} \sum_{i=1}^{k(n)} \chi_E \left(t_i^{(n)} \right) - P(E) \right|.$$

By using a procedure similar to the one used in the proof of the Theorem 3.2.15 we can prove the following theorem.

Theorem 3.3.4.

Let (π_i) be the sequence of partitions of F just constructed. Assume that $\log\left(\frac{1}{p_1}\right),\ldots$
$\ldots,\log\left(\frac{1}{p_m}\right)$ are rationally related. Then we have the following bounds for the elementary discrepancy

$$D_n^{\mathscr{E}} = \begin{cases} \mathcal{O}\left((\log k(n))^d k(n)^{-\eta}\right) & \text{if } 0 < \eta \le 1, \\ \mathcal{O}\left(k(n)^{-1}\right) & \text{if } \eta > 1. \end{cases} \tag{3.31}$$

Furthermore, both upper bounds are best possible.

Proof.

Fix a step in the algorithm corresponding to a certain parameter r of the form $r = e^{-n\Lambda}$ for some integer $n \ge 0$. First we observe that the number $N_E^{(n)}$ of elementary sets if $\mathscr{E}_n$ that are contained in a fixed elementary set E is given by $M_{r/P(E)}$. This fact implies that

$$N_E^{(n)} = \frac{c'}{r}P(E) + \mathcal{O}\left(|\log r|^d\, r^{-1+\eta} P(E)^{1-\eta}\right). \tag{3.32}$$

This proves (3.31) directly for $\eta \le 1$ and also shows that this bound is optimal.

If $\eta > 1$ then we argue recursively. The elementary set E is either contained in $\mathscr{E}_1 = \{\varphi_1(F),\ldots,\varphi_m(F)\}$, which means that we can use (3.32) for $P(E) \in \{p_1,\ldots,p_m\}$, or it is part of $E_j = \varphi_j(F)$ for some j. In the latter case we can rewrite $N_E^{(n)} - k(n)P(E)$ to

$$N_E^{(n)} - k(n)P(E) = \left(N_E^{(n)} - k(n-1)\frac{P(E)}{P(E_j)}\right) + \left(k(n-1)\frac{P(E)}{P(E_j)} - k(n)P(E)\right),$$

which leads to a recurrence of the form

$$\Delta_n^{\mathscr{E}} = \sup_{E \in \mathscr{E}} \left| N_E^{(n)} - k(n)P(E) \right| \le \Delta_{n-1}^{\mathscr{E}} + \mathcal{O}\left(n^d e^{\Lambda n(1-\eta)}\right).$$

Here we have

$$\Delta_n^{\mathscr{E}} = \mathcal{O}\left(\sum_{k \le n} k^d e^{-\Lambda k(\eta-1)}\right) = \mathcal{O}(1)$$

and consequently $D_n^{\mathscr{E}} = \mathcal{O}(1/k(n))$ (which is also optimal).

$\square$

In particular it follows that the sequence of partitions (π_n) is u.d. with respect to P. Actually, this remains true in the irrationally related case but we can only derive effective upper bounds for the discrepancy in very specific cases.

Chapter 4

Conclusions and open problems

The main objectives of this work were to construct new classes of u.d. sequences of partitions and of points on fractals and on $[0, 1]$ and to study their discrepancy. In fact with reference to fractals, we were able to find a general algorithm for producing sequences of van der Corput type on fractals generated by an IFS consisting of similarities which have the same ratio and which satisfy the OSC. Secondly, we got bounds for the discrepancy of a class of generalized Kakutani's sequences of partitions on $[0, 1]$ constructed through the recent technique of successive $\rho-$refinements. Moroever, these last results enabled us to introduce a new family of u.d. sequences on a wider class of fractals. Although the techniques we used are direct and explicit procedures, the quantitive analysis of the distribution properties of the sequences constructed was proved to be more involved than we expected. Indeed, we now want to discuss major problems as far as the study of discrepancy is concerned.

A first problem was finding a unifying approach to the discrepancy on the class of fractals considered in this work. In fact, the only kind of discrepancy which makes sense for all the fractals generated by IFS and satisfying the OSC is the so-called elementary discrepancy, as we have already noted in Subsection 2.2.2. We got estimates for the elementary discrepancy of the sequences generated by our techniques in Theorem 2.2.5 and Theorem 3.3.4. In both cases the particular properties given by our procedures to the sequences allowed to obtain significative estimates for their discrepancy with a quite direct proof. Neverthless, it is not surprising that these properties are no more sufficient to provide estimates for another kind of discrepancy. This is due to the fact that the elementary sets are intrisically related to the construction of the fractals belonging to our class and not to the specific geometry of each fractal. In literature, there are few papers devoted to u.d. sequences on fractals

and to estimates of the discrepancy, see [18, 17, 27]. In these articles, the various types of discrepancy considered depend heavily on the geometric features of the particular fractal considered. At the moment the choice of elementary discrepancy seems to be the only one which allows to overcome the problem of the peculiar complexity of each fractal and to give explicit results for the fractals of the whole class taken in consideration. So a still open problem is finding a definition of discrepancy, different from the elementary one, which is general at least for the fractals of our class. An attempt in this direction has been proposed by Albrecher, Matoušek and Tichy in [2], but it concerns the average discrepancy.

This work also leaves open several questions concerning the discrepancy of the generalized Kakutani's sequences constructed by the $\rho-$refinements technique on $[0, 1]$. As we have already said in Section 3.2, the natural problem of studying the behaviour of the discrepancy D_n (1.7) of the sequence of partitions constructed as n tends to infinity was posed in [69]. In this thesis we partially answered to that question. Indeed, we were able to find a class of partitions ρ such that the speed of convergence to zero of the discrepancy of the sequences generated by successive $\rho-$refinements is quite high (see Theorem 3.2.15). The strategy used to get bounds for the discrepancy of these sequences exploits the correspondence between the procedure of $\rho-$refinements and Khodak's algorithm. Although this analogy is crucial in our analysis, it is not sufficient to give effective estimates in the irrational case, too. In fact, as we have already discussed in Subsection 3.2.4, the irrational case is too closely related to the Diophantine approximation properties of the quotients $\frac{\log p_i}{\log p_j}$ where the p_i's are the lenghts of the starting partition ρ. This problem is mainly caused by the necessity of having an explicit error term in the asymptotic expansion (3.8), which is not provided by the results related to Khodak's algorithm in [20]. In particular, we gave more precise information on the error term when the initial probabilities are only two $p, q = 1 - p$ and $\gamma = \log p / \log q$ is badly approximable or when p and q are algebraic numbers. In these two cases, the correspondent upper bounds for the discrepancy are weaker than the one in the rational case. Consequently, even if the continued fractional expansion of γ could be useful for getting more explicit upper bounds for discrepancy, the relative estimates will probably be again poor. Moreover, as we mentioned in Subsection 3.2.4 the paper [26] shows the difficulties presented by the irrational case when ρ consists of more than two intervals. Therefore it could be interesting to develop different approaches to the problem, which allow to improve the results at least for a class of instances of the irrational case.

A further interesting problem is finding explicit algorithms to provide low discrepancy sequences of points associated to a low discrepancy u.d. sequence of partitions constructed by successive $\rho-$refinements. In Subsection 1.2.2 we analyzed this problem in general and reported the main result about this question developed in [69]. In fact, in this paper it has been proved that a random reordering of the points of a u.d. sequence of partitions gives with probability one a u.d. sequence of points (see Theorem 1.2.8). Neverthless, this important theoretical result does not give any information about the discrepancy of the sequences of points obtained by the random reordering. Some results in this direction have been already obtained concerning to the LS-sequences. In fact, in [10] the author presents an explicit procedure for associating to LS-sequences with low discrepancy sequences of points having low discrepancy, too. It would be of great interest trying to extend this algorithm to the whole class of sequences of partitions constructed by successive $\rho-$refinements included in the rational case.

Acknowledgments

It is my pleasure to thank the many people who made this thesis possible.

I would like to begin by expressing my gratitude to my supervisor Prof. Aljoša Volčič. He was the first who believed in my capabilities and encouraged me during all these years. His devotion and passion for research were contagious for me and his suggestions were precious for my scientific and personal growth.

A special thanks to Prof. Michael Drmota who led me carefully during my first study experience abroad. He has always showed interest in my work and his intuitions opened my mind to many interesting research topics. I feel privileged for having the opportunity to work with him.

I am also grateful to Ingrid who supported me with her wise advices, providing me encouragenments in several moments.

An acknowledgement goes to all the professors who I met during my studies. Each of them contributed to improve my knowledge and so myself. In particular, I want to thank my high school math teacher Prof. Paola Armentano. Her love for mathematics was crucial in the development of my passion for this subject.

I am very thankful to all my collegues who provided a fun working environment. In particular, I am grateful to my room-mates Marina, Marianna and Viviana for having always been close to me with patience and for having divided with me joy and stress in these years. I want also to express my gratitude to Mascia for her deep friendship and for the inconditioned help that she has always given to me. Moreover, a warm thanks is for Caterina, Rosa Maria and Raffaella who have supported me since my first day in this university. Thanks also to Michele whose suggestions have been important for my training.

I am surely indebted to Mario who has made available his support in several ways during these years and to Gibbì who revealed to be an irreplaceable friend.

Apart from the work environment, a great thanks goes to all my friends of "Post Cresima" who every day give me for gift a piece of their hearth. I would like to thank especially Ilaria who played a such important role along these years in providing en-

Bibliography

[1] R.L. Adler and L. Flatto. *Uniform distribution of Kakutani's interval splitting procedure.* Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, **38** (no.4): 253–259, 1977.

[2] H. Albrecher, J. Matoušek, and R.F. Tichy. *Discrepancy of point sequences on fractal sets.* Publ. Math. Debrecen, **56** (no. 3-4): 233–249, 2000.

[3] A. Baker. *Transcendental Number Theory.* Cambridge University Press, London, 1975.

[4] V. Bergström. *Einige Bemerkungen zur Theorie der diophantischen Approximationen.* Fysiogr. Sälsk. Lund. Förh., **6** (no. 13): 1–19, 1936.

[5] P. Bohl. *Über ein in der Theorie der säkularen Störungen vorkommendes Problem.* J. reine angew. Math., **135**: 189–283, 1909.

[6] E. Braaten and G. Weller. *An improved low-discrepancy sequence for multidimensional quasi Monte Carlo integration.* J. Comput. Phys., **33**: 249–258, 1979.

[7] A. Brauer. *On algebraic equations with all but one root in the interior of the unit circle.* Math. Nachr., **4**: 250–257, 1951.

[8] M.D. Brennan and R. Durret. *Splitting intervals.* Annals of Probability, **14** (no.3): 1024–1036, 1986.

[9] N.G. De Bruijn and K.A. Post. *A remark on uniformly distributed sequences and Riemann integrability.* Indag. Math., **30**: 149–150, 1968.

[10] I. Carbone. *Discrepancy of LS-sequences of partitions.* submitted.

[11] I. Carbone and A. Volčič. *A von Neumann theorem for uniformly distributed sequences of partitions.* To appear in Rend. Circolo Mat. Palermo.

[12] I. Carbone and A. Volčič. *Kakutani's splitting procedure in higher dimension.* Rend. Ist. Matem. Univ. Trieste, **XXXIX**: 119–126, 2007.

[13] H. Chaix and H. Faure. *Discrépance et diaphonie en dimension un.* Acta Arith., **63** (no. 2): 103–141, 1993.

[14] F. Chersi and A. Volčič. *λ-equidistributed sequences of partitions and a theorem of the de Bruijn-Post type.* Ann. Mat. Pura Appl. (IV), **162**: 23–32, 1992.

[15] V. Choi and M. J. Golin. *Lopsided trees, I: Analyses.* Algorithmica, **31** (no. 3): 240–290, 2001.

[16] K.L. Chung. *A course in probability theory.* Academic Press Inc., San Diego, CA, third edition, 2001.

[17] L.L. Cristea, F. Pillichshammer, G. Pirsic, and K. Scheicher. *Discrepancy estimates for point sets on the s-dimensional Sierpiński carpet.* Quaest. Math., **27** (no. 2): 375–390, 2004.

[18] L.L. Cristea and R.F. Tichy. *Discrepancies of point sequences on the Sierpiński carpet.* Math. Slovaca, **53** (no. 4): 351–367, 2003.

[19] M. Drmota and M. Infusino. *On the discrepancy of some generalized Kakutani's sequences of partitions.* submitted.

[20] M. Drmota, Y.A. Reznik, and W. Szpankowski. *Tunstall Code, Khodak Variations, and Random Walks.* IEEE Trans. Inf. Th., **56**: 2928–2937, 2010.

[21] M. Drmota. and R.F. Tichy. *Sequences, discrepancies and applications.* Lecture Notes in Mathematics 1651. Springer-Verlag, Berlin, 1997.

[22] K. J. Falconer. *The Geometry of fractal Sets.* Cambridge Tracts in Mathematics 85. Cambridge University Press, Cambridge, 1986.

[23] K. J. Falconer. *Fractal Geometry-Mathematical Foundations and Applications.* John Wiley & Sons Ltd., Chichester, 1990.

[24] H. Faure. *Discrépances de suites associées à un système de numération (en dimension un).* Bull. Soc. Math. France, **109** (no. 2): 143–182, 1981.

[25] H. Faure. *Discrepancy and diaphony of digital $(0,1)$-sequences in prime base.* Acta Arith., **117** (no. 2): 125–148, 2005.

[26] P. Flajolet, M. Roux, and B. Vallee. *Digital Trees and Memoryless Sources: from Arithmetics to Analysis.* To appear in DMTCS proc.

[27] P.J. Grabner and R.F. Tichy. *Equidistribution and Brownian motion on the Sierpiński gasket.* Monatsh. Math., **125**: 147–164, 1998.

[28] J. H. Halton. *On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals.* Numer. Math., **2**: 84–90, 1960.

[29] J. M. Hammersley. *Monte Carlo methods for solving multiple problems.* Ann. New York Acad. Sci., **86**: 844–874, 1960.

[30] Z. Hedrlín. *On integration in compact metric spaces.* Comm. Mat. Univ. Carolinae, **2** (no. 4): 17–19, 1961.

[31] P. Hellekalek and G. Larcher. *Random and Quasi-random point sets.* Lecture Notes in Mathematics 138. Springer-Verlag, New York, 1998.

[32] G. Helmberg. *Abstract theory of uniform distribution.* Compos. Math., **16**: 72–82, 1964.

[33] P. Henrici. *Applied and Computational Complex Analysis*, volume 2. John Wiley and Sons, New York, 1977.

[34] E. Hewitt and L. J. Savage. *Symmetric measures on cartesian products. Trans. Amer. Math. Soc.*, **80**: 470–501, 1955.

[35] E. Hlawka. *Folgen auf kompakten Räumen.* Abh. Math. Sem. Univ. Hamburg, **20**: 223–241, 1956.

[36] E. Hlawka. *Folgen auf kompakten Räumen II.* Mat. Nachr., **18**: 188–202, 1958.

[37] E. Hlawka. *Funktionen von beschränckter Variation in der Theorie der Gleichverteilung.* Ann. Mat. Pura Appl., **54**: 325–333, 1961.

[38] E. Hlawka. *Uniform distribution modulo 1 and numerical analysis.* Compos. Math., **16**: 92–105, 1964.

[39] J.E. Hutchinson. *Fractals and self-similarity.* Indiana Univ. Math. J., **30**: 713–747, 1981.

[40] M. Infusino and A. Volčič. *Uniform distribution on fractals.* Uniform Distribution Theory, **4** (no. 2): 47–58, 2009.

[41] S. Kakutani. *A problem of equidistribution on the unit interval* [0, 1]. In *Measure theory (Proc. Conf., Oberwolfach, 1975)*, pages 369–375, Lecture Notes in Math. **541**. Springer, Berlin, 1976.

[42] A. Ya. Khinchin. *Continued Fractions.* Dover Publications Inc., Mineola, NY, 1997.

[43] G.L. Khodak. *Connection Between Redundancy and Average Delay for Fixed-Lenght Coding.* All-Union Conference on Problems of Theoretical Cybernetics (Novosibirsk, USSR, 1969), **12**. (in Russian).

[44] J. Korevaar. *A century of complex Tauberian theory.* Bull. Amer. Math. Soc., **39**: 475–531, 2002.

[45] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience Publ., New York, 1974.

[46] J.C. Lootgieter. *Sur la répartition des suites de Kakutani (I).* Ann. Inst. Henty Poincaré, **Vol. XIII** (no. 4): 385–410, 1977.

[47] J.C. Lootgieter. *Sur la répartition des suites de Kakutani (II).* Ann. Inst. Henty Poincaré, **Vol. XIV** (no. 3): 279–302, 1978.

[48] B. B. Mandelbrot. *The Fractal Geometry of Nature.* W. H. Freeman and Co., San Francisco, Calif., 1982.

[49] P.A.P. Moran. *Additive functions of intervals and Hausdorff measure.* Proc. Camb. Phil. Soc., **42**: 15–23, 1946.

[50] W.J. Morokoff and R.E. Caflisch. *Quasi-random sequences and their discrepancies.* SIAM J. Sci. Comput., **15**: 1251–1279, 1994.

[51] H. Niederreiter. *On the existence of uniformly distributed sequences in compact spaces.* Compos. Math., **25**: 93–99, 1972.

[52] H. Niederreiter. *Random number generation and Quasi-Monte Carlo methods.*, volume **Vol. 63**. SIAM Conf. Ser. Appl. Math., Philadelphia, PA, 1992.

[53] E. Pap, editor. *Handbook of measure theory.* North-Holland, Amsterdam, 2002.

[54] K.R. Parthasarathy. *Probability measures on metric spaces.* Academic Press, 1967.

[55] R. Pyke and W. R. van Zwet. *Weak convergence results for the Kakutani interval splitting procedure.* Ann. Probab., **32** (no. 1A): 380–423, 2004.

[56] K.F. Roth. *On irregularities of distribution.* Mathematika, **1**: 73–79, 1954.

[57] S. Salvati and A. Volčič. *A quantitative version of de Bruijn-Post theorem.* Math. Nachr., **229**: 161–173, 2001.

[58] W. Schachinger. *Limiting distributions for the costs of partial match retrievals in multidimensional tries.* Random Structures and Algorithms, **17**: 428–459, 2000.

[59] C. Schlier. *On scrambled Halton sequences.* Appl. Numer. Math., **58**:1467–1478, 2008.

[60] K. Schmidt. *Über einen Zusammenhang zwischen gleichverteilten Punkt- und Maßfolgen.* J. Reine Angew. Math., **244**: 94–96, 1970.

[61] W.M. Schmidt. *Irregularities of distribution.VII.* Acta Arithmetica, **21**: 45–50, 1972.

[62] W. Sierpiński. *Sur la valeur asymptotique d'une certaine sommme.* Bull Intl. Acad. Polon. Sci.(Cracovie), **A**: 9–11, 1910.

[63] W. Szpankowski. *Average Case Analysis of Algorithms on Sequences.* Wiley-Interscience, New York, 2001.

[64] T. van Aardenne-Ehrenfest. *Proof of the impossibility of a just distribution of an infinite sequence of points over an interval.* Proc. Kon. Ned. Akad. v. Wetensch., **48**: 266–271, 1945.

[65] T. van Aardenne-Ehrenfest. *On the impossibility of a just distribution.* Proc. Kon. Ned. Akad. v. Wetensch., **52**: 734–739, 1949.

[66] J. G. van der Corput. *Verteilungsfunktionen.* Proc. Ned. Akad. v. Wet., **38**: 813–821, 1935.

[67] J. G. van der Corput and C. Pisot. *Sur la discrépance modulo un.* Indag. Math., **1**: 260–269, 1939.

[68] W.R. van Zwet. *A proof of Kakutani's conjecture on random subdivision of longest intervals.* Annals of Probability, **6** (no. 1): 133–137, 1978.

[69] A. Volčič. *A generalization of Kakutani's splitting procedure.* To appear in Annali Mat. Pura e Appl.

[70] H. Weyl. *Über die Gibbssche Erscheinung und verwandte Konvergenzphänomene.* Rend. Circ. Mat. Palermo, **30**: 377–407, 1910.

[71] H. Weyl. *Über ein Problem aus dem Gebiete der diophantischen Approximationen.* Nachr. Ges. Wiss. Göttingen, *Math.-phys. Kl.*, pages 234–444, 1914.

[72] H. Weyl. *Über die Gleichverteilung von Zählen mod. Eins.* Math. Ann., **77**: 313–352, 1916.