SOCIEDADE BRASILEIRA DE MATEMÁTICA

ENSAIOS MATEMÁTICOS
2009, Volume **17**, 1–104

The remarkable effectiveness of ergodic theory in number theory

Part I. Green-Tao theorem

by Alexander Arbieto, Carlos Matheus and Carlos G. Moreira

Part II. Elkies-McMullen theorem

by Carlos Matheus

Abstract. The main goal of this survey is the description of the fruitful interaction between Ergodic Theory and Number Theory via the study of two beautiful results: the first one by Ben Green and Terence Tao (about long arithmetic progressions of primes) and the second one by Noam Elkies and Curtis McMullen (about the distribution of the sequence $\{\sqrt{n}\}$ mod 1). More precisely, during the first part, we will see how the ergodic-theoretical ideas of Furstenberg about the famous Szemerédi theorem were greatly generalized by Green and Tao in order to solve the classical problem of finding arbitrarily long arithmetical progression of prime numbers, while the second part will focus on how Elkies and McMullen used the ideas of Ratner's theory (about the classification of ergodic measures related to unipotent dynamics) to compute *explicitly* the distribution of the sequence $\{\sqrt{n}\}$ on the unit circle.

2000 Mathematics Subject Classification: 37-02, 11-02.

Preface

The Nobel laureate physicist Eugene Wigner wrote (in 1960) an article entitled *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*. In this paper, he argues that the fruitful interaction between Mathematics and Physics often points towards profound advances in Physics and he claims that this is not just a coincidence: instead, this beautiful interaction reflects a deeper truth about both Mathematics and Physics.

In the present book, we will discuss a much more modest interaction between two areas of Mathematics, namely Ergodic Theory and Number Theory, leading to the solution of several interesting problems of Number Theory (see the next paragraph below). Of course, this interaction has a much more recent history when compared with the interaction of Mathematics and Physics, so that it would be exaggerated to call it "unreasonable". On the other hand, since the Ergodic Theory certainly sheds light into some deep questions in Number Theory, we believe it is fair to say that the interaction between these two areas of Mathematics is "remarkable" (so that this justifies the choice for the title of this book).

More concretely, one can quote the following results where some ideas from Ergodic Theory helped the understanding of a problem in Number Theory:

- Weyl estimates of *exponential sums* – the study of exponential sums $\sum\limits_{j=1}^{n} e^{2\pi i x_j}$ is related to the *equidistribution* properties of the real numbers $x_n (\mathrm{mod}\, 1)$;

- Lévy's constant $e^{\pi^2/12\ln 2} \simeq 3,27582291872...$ giving the asymptotic exponential growth $\sqrt[n]{q_n}$ of the denominators q_n of the continued fraction expansion of (Lebesgue) almost every real number x can be very clearly explained in terms of the Ergodic Theory of the so-called Gauss map $G : [0,1] \to [0,1]$, $G(x) = \left\{\frac{1}{x}\right\}$, where $\{.\}$ denotes the fractionary part of . (namely, via Pesin's formula relating the entropy and the Jacobian of G with respect to its natural absolutely continuous invariant measure).

- Margulis' solution (1989) of Oppenheim conjecture – the study of quadratic forms Q of n variables restricted to the integer lattice $\mathbb{Z}^n$ is an important subject of Number Theory; for instance, Lagrange theorem says that the image of the quadratic form $Q(a, b, c, d) = a^2 + b^2 + c^2 + d^2$ is precisely the set of all natural numbers. In general, when Q is positive definite, it is easy to see that $Q(\mathbb{Z}^n)$ is a discrete subset of the positive real numbers. On the other hand, the study of indefinite forms can be a little bit more sophisticate, e.g., the study of the values of $Q(a,b) = a^2 - db^2$ for some integer d involves the class field theory of $\mathbb{Q}(\sqrt{-d})$ as pointed out by Gauss. In general, when the coefficients of Q are commensurable, we see that the image $Q(\mathbb{Z}^n)$ is discrete, so that a natural question arises: what happens when the coefficients of Q are incommensurable? In two variables, we still have some discreteness phenomena (as the reader can check with the quadratic form $Q(a,b) = a^2 - ((1+\sqrt{5})/2)^2 b^2)$. Nevertheless, Oppenheim (1929) conjectured that this discreteness phenomena doesn't occur in higher dimensions: for $n \geq 3$, the image $Q(\mathbb{Z}^n)$ of $\mathbb{Z}^n$ by any indefinite quadratic form Q on n variables is dense in $\mathbb{R}$. This number-theoretical conjecture was solved by Margulis via the Ergodic Theory of certain homogenous flows on the space of lattices (and, in fact, to the best of our knowledge, all proofs of Oppenheim conjecture are ergodic-theoretical in nature).

- Elkies-McMullen theorem (2004) on the gaps of $\sqrt{n}$ (mod 1) – given a sequence of real numbers on the circle $S^1 = \mathbb{R}/\mathbb{Z}$, a natural question concerns the study of its distribution: namely, we can remove these points from the circle and look at the lengths of the intervals one gets in this process. For certain sequences of numbers (e.g., a sequence of randomly chosen real numbers), we know that the distribution law is given by an exponential function. However, for the sequence $\{n^\alpha\}$ with $0 < \alpha < 1$, we know that it is equidistributed on S^1 and it is conjectured that its distribution law is exponential for any $\alpha \neq 1/2$. Furthermore, some numerical experiments performed by Boshernitzan (around 1993) indicated an special distribution for the particular case $\alpha = 1/2$. In this direction, Elkies and McMullen used the Ergodic Theory of homogenous flows (and, in particular, Ratner's theorem) to explicitly compute the distribution of $\{\sqrt{n}\}$. A consequence of their result is the fact that $\{\sqrt{n}\}$ doesn't have an exponential distribution.

- Green-Tao theorem (2004) on existence of arbitrarily large arithmetic progressions of primes – a classical subject of Number Theory is the study of prime numbers. A particularly interesting problem is to understand the additive properties of the prime numbers. In this direction, B. Green and T. Tao were able to use the ergodic-theoretical ideas of Furstenberg's proof of Szemerédi theorem to show that there are arbitrarily large arithmetic progressions of primes.

- Einsiedler-Katok-Lindenstrauss theorem (2006) on Littlewood conjecture – a fascinating problem of Number Theory is the study of rational approximations of irrational numbers. Although we have a lot of information about the rational approximations of a single irrational number (due to the marvelous properties of the continued fraction algorithm), the situation becomes more delicate when one asks about simultaneous approximations of irrational numbers. In this sense, Littlewood (1930) proposed the following conjecture: $\liminf_{n\to\infty} n\{n\alpha\}\{n\beta\} = 0$ for all $\alpha, \beta \in \mathbb{R}$. Again using the Ergodic Theory of homogenous flows, Einsiedler, Katok and Lindenstrauss were able to give the following almost complete answer to Littlewood conjecture: the set of exceptional pairs (α, β) of Littlewood conjecture has Hausdorff dimension zero!

The initial plan of this book was to cover the Green-Tao, Elkies-McMullen and Einsiedler-Katok-Lindenstrauss theorems. However, due to the usual problem of limitation of space and time, we were forced to make a choice between these three beautiful results. Because Einsiedler-Katok-Lindenstrauss theorem is a little bit more difficult to explain from the technical point of view (in the author's opinion), we have chosen to discuss Green-Tao and Elkies-McMullen theorems.

More precisely, this book has two parts: the first part, by A. Arbieto, C. G. Moreira and C. Matheus, consists of the first two chapters and concerns Green and Tao theorem, while the last one, by C. Matheus, consists of the third chapter and concerns Elkies and McMullen theorem. The resume of the contents of these chapters is:

- in the first chapter (part I), we will make a historical review of the basic questions and theorems about the additive and multiplicative properties of integer numbers. In particular, we are going to see that there are several problems about the additive properties of prime numbers which are very easy to state but very difficult to solve: e.g., it is not known whether there are infinitely many pairs of twin prime numbers, i.e., pairs of prime numbers whose difference is 2 (this is known as the Twin Prime Conjecture) and it is still open the so-called Goldbach conjecture saying that every even natural number ≥ 4 is the sum of two prime numbers. Also, we will see that another classical conjecture (solved by Ben Green and Terence Tao) says that there are arbitrarily large arithmetical progressions of prime numbers. The biggest known progression of primes (to the best of the authors' knowledge) is

  $$6171054912832631 + k \cdot 81737658082080, 0 \leq k \leq 24,$$

  formed by 25 prime numbers (this arithmetical progression of primes was discovered in May 17, 2008 by Raanan Chermoni and Jaroslaw Wroblewski). Observe that this conjecture was solved by Ben Green and Terence Tao in 2004 and, in fact, Terence Tao was awarded a Fields Medal in 2006 partly due to this outstanding work with

Ben Green. The basic plan of the first chapter is to discuss some "preparatory" results in the direction of Green and Tao theorem such as Szemerédi theorem and its ergodic-theoretical proof by Furstenberg.

- in the second chapter (part I), we will present the ergodic-theoretical component of Green and Tao argument, namely, the proof of Green-Tao-Szemerédi theorem (via the introduction of the Gowers norms and the adaptation of Furstenberg proof of Szemerédi theorem). Once this theorem is proved, it remains to perform a number-theoretical argument to conclude the Green-Tao theorem from the Green-Tao-Szemerédi. However, in order to be coherent with the title of this book, we will *omit* the *very interesting* part related to the construction of a pseudorandom majorant of the (modified) von Mangoldt function (based on the works of Goldston and Yildirim).

- finally, in the last chapter (part II), we will completely change the subject from the additive properties of prime numbers to the Elkies-McMullen calculation of the distribution law of $\sqrt{n} \pmod 1$. In particular, we subdivide this chapter into three sections: the first two concerns the translation of the problem of computing the distribution of $\sqrt{n} \pmod 1$ into an ergodic-theoretical problem and the last section concerns the solution of the corresponding ergodic problem via Ratner theory of homogenous flows.

Evidently, as the reader can infer from this summary, the parts I and II are *completely independent*, so that the reader can chose where he/she wants to start reading the book.

Finally, we would like to apologize for the omission of Einsiedler-Katok-Lindestrauss theorem: as a form of compensation, C. Matheus would like to say that he's planning to include some notes about Einsiedler-Katok-Lindenstrauss theorem in his mathematical blog "Disquisitiones Mathematicae" (http://matheuscmss.wordpress.com/) in a near future.

Acknowledgments

The first two chapters are strongly based on our book [1] “Aspectos Ergódicos da Teoria dos Números” (in Portuguese) for a mini-course of the 26th Brazilian Mathematical Colloquium, which was, in turn, based on our discussions with Jimmy Santamaria in an Ergodic-Theory seminar at IMPA around Green-Tao theorem. In particular, we would like to thank Jimmy Santamaria for his decisive participation in this seminar. Also, we would like to thank Yuri Lima and Jimmy Santamaria for a careful revision of previous versions of this work.

Furthermore, this book was written during the pos-doctoral stay of Carlos Matheus at Collège de France (under the supervision of Jean-Christophe Yoccoz). In particular, he is thankful to Jean-Christophe for the several interesting mathematical (and non-mathematical) discussions, advices and friendship, and the two secretaries of the Mathematics section of College de France, Dominique Bidois and Veronique Sainz, for their invaluable support and friendship. Moreover, Carlos Matheus is grateful to his beloved wife Aline Gomes Cerqueira for her patience and support (both mathematical and non-mathematical).

Last, but not least, it is a great pleasure to acknowledge Étienne Ghys, Maria Eulália Vares and Vladas Sidoravicius, current editors of “Ensaios Matemáticos”, for their invitation to start this project (after our mini-course during the 26th Brazilian Math. Colloquium). In particular, we are grateful to Maria Eulália Vares whose strong encouragement and kind words kept us working on this project.

A. Arbieto, C. Matheus, C. Moreira,
Rio de Janeiro and Paris, Oct. 30, 2009

Contents

Part I

Green-Tao theorem

Chapter 1

Additive properties of prime numbers

1.1 Introduction

One of the oldest concepts in Mathematics is the the notion of *prime number*. By definition, an integer number p is *prime* if it is divisible only by 1 and by itself.

Prime numbers are important objects in Number Theory due the unique factorization theorem (saying that any integer number can be written as a product of prime numbers in an essentially unique way). Another elementary property of prime numbers is the fact that they are precisely the integer numbers p such that $\mathbb{Z}/p\mathbb{Z}$ is a *field*.

Obviously, due to the *multiplicative* character of the definition of prime numbers, it is fairly easy to extract its multiplicative properties. For instance, the product of two primes is certainly not a prime number and there are no *geometrical* progressions of primes of length ≥ 3 formed only by prime numbers.

On the other hand, the situation changes considerably when one poses some questions about the *additive* character of the prime numbers. For example, one can ask whether the sum of two prime numbers is still a prime number. Of course, the answer is: it depends. In fact, 2+3=5 is prime and 2+5=7 is prime, but neither 3+5=8 isn't prime nor 7+2=9. However, the Bertrand's postulate says that, for every natural number N, there exists a prime number between N and $2N = N + N$. In particular, this shows that the following question deserves a little bit of attention:

Are there arithmetical progressions of length ≥ 3 formed only by prime numbers? In the case of an affirmative answer to this question, how many of them exist once the length of the arithmetical progression is fixed?

We are going to see (in the first two chapters of this book) that this problem was solved by Ben Green and Terence Tao. However, before entering this issue, let us take a little trip around the world of the prime numbers in order to see some related questions about the additive properties of prime numbers and its partial solutions.

1.2 Classical problems about the additive properties of prime numbers

1.2.1 The twin prime conjecture

As the examples of the introduction showed us, the sum of a prime number with 2 is not always a prime number. However, we can ask whether there are infinitely many primes of this form. We say that p and $p+2$ are *twin primes* whenever both of them are prime numbers. Some examples of twin primes are: (3 and 5), (5 and 7), (11 and 13), (17 and 19), (29 and 31), (41 and 43), etc. One of the most famous open conjectures in Number Theory is the so-called *twin prime conjecture*:

Are there infinitely many twin prime numbers?

An important result due to Brun [2] gives a flavor of the difficulty of this conjecture: namely, Brun proved that, even in the case of the existence of infinitely many pairs of twin primes, it is a very hard task to locate them because they are very rare. More precisely, Brun's theorem says that the series of the inverse of the twin primes converges (to a certain number called Brun's constant):

$$(\frac{1}{3}+\frac{1}{5})+(\frac{1}{5}+\frac{1}{7})+(\frac{1}{11}+\frac{1}{13})+(\frac{1}{17}+\frac{1}{19})+...<+\infty.$$

Later on, we will reformulate this conjecture in a more analytical language.

1.2.2 Goldbach conjecture

In a letter addressed to Euler (in 1742), Goldbach asked whether every integer number ≥ 2 is the sum of 3 prime numbers. In his formulation, Goldbach assumed that 1 is a prime number, although nowadays this convention is not used anymore. In modern terms, *Goldbach's conjecture* can be reformulated as:

Can we write every even integer number $n \geq 4$ as a sum of 2 prime numbers?

Although this conjecture is fairly easy to state, Goldbach's conjecture is still one of the big challenges in Number Theory. Nowadays, we have several partial results, but none of them seem to extend to a full solution of this conjecture.

For example, Schnirelman [13] showed that every integer number can be written as a sum of prime numbers, where the quantity of terms can be bounded by ~ 300000 (a little bit far from 2, don't you think?).

A related conjecture (also called Goldbach conjecture) is:

Can we write every odd integer number $n \geq 9$ as a sum of 3 prime numbers?

In this direction, we have a famous result of Vinogradov [20] solving this conjecture for any sufficiently large odd integers (for instance, the conjecture holds for any odd integer $\geq 3^{3^{15}}$).

Another interesting result is Chen's theorem [3] showing that any sufficiently large even integer is the sum of a prime number and a *quasi-prime* number (i.e., an integer number with 2 prime factors at most).

A stronger version of Goldbach conjecture (called Levy's conjecture) is:

Can we write every odd integer $n \geq 7$ as a sum of a prime number p and the double $2q$ of another prime number q?

Later on, we will also reformulate these conjectures in an analytical language.

1.2.3 Some results about arithmetical progressions vs. prime numbers

A classical result in this subject is Dirichlet's theorem:

If a and b are relatively prime, then the arithmetical progression $a + nb$ contains infinitely many primes.

During the proof of this result, Dirichlet introduce the important concept of *L-series*. In particular, we will omit the proof of this beautiful theorem in order to keep the coherence with the purpose of this book.

Observe that Dirichlet theorem *doesn't* say that this arithmetical progression is *entirely* formed by prime numbers. A natural question arises: are there arithmetical progressions of infinite length formed only by prime numbers? The *negative* answer to this problem is provided by the following theorem of Lagrange and Waring:

Consider an arithmetical progression formed only by prime numbers of length k and ratio d. Then, d is necessarily divisible by every prime number smaller than k. In particular, there is no infinite arithmetical progression formed only by prime numbers.

1.3 Arithmetical progressions in certain subsets of $\mathbb{Z}$

The question of existence of arithmetical progressions (AP) of finite length of prime numbers can be extended as follows:

Let $A \subset \mathbb{Z}$ be a given infinite subset of integers. Are there APs of arbitrarily large length formed only by elements of A?

In a certain sense, we will see that the subset P of prime numbers is very "thin". Thus, we can start to attack the previous question by the "easy" case of "fat" subsets A, where there is a nice chance of finding AP, and then one could try to adapt the method to work with "thin" sets (such as the set of primes P). Of course, the central problem lies in the formalization of the definition of "thin" and "fat" subsets. In this section, we will revise some results in this direction, although we should stress that we are not going to follow the chronological order in the exposition of these theorems.

1.3.1 Van der Waerden theorem

Suppose that you are given a finite number, say k, of *colors* and you use them to color the integer numbers. In this process, you get a partition of the integer numbers into k disjoint subsets. Van der Waerden theorem says that:

At least one of these subsets is so "fat" that it should contain arbitrarily large arithmetical progressions.

In particular, if we take two colors and we give one color to the set of prime numbers P and the other color to the composite (i.e., non-prime) numbers, we obtain:

Either the set of prime numbers or the set of composite numbers possesses arbitrarily large arithmetical progressions.

Later in this chapter, we will see two proofs of Van der Waerden theorem (one of them is combinatorial and the other is ergodic-theoretical).

1.3.2 Szemerédi theorem

Logically, the subset of even integers possesses arithmetical progressions of arbitrarily large length (with ratio 2, for instance). Observe that the even integers occupy essentially 1/2 of any interval $[1, N] := \{n \in \mathbb{Z}; 1 \leq n \leq N\}$. Similarly, the odd integers share the same property and they also possess arbitrarily large arithmetical progressions. More generally, we can pick an integer number k and we can look at the subset of multiples of k: this subset essentially occupies $1/k$ of any interval $[1, N]$ and it has arbitrarily large arithmetical progressions.

Based on these simple remarks, we are tempted to say that a subset A is "fat" if it occupies a definite positive fraction of the intervals $[1, N]$ for a infinite sequence of Ns:

Definition 1.3.1. *Let $A \subset \mathbb{N}$. The (upper) density of A is:*

$$d(A) = \limsup_{N\to\infty} \frac{|[1, N] \cap A|}{N}.$$

Here, $|B|$ denotes the cardinality of a given $B \subset \mathbb{N}$.

Of course, this definition extends to subsets of integer numbers. The first theorem dealing with the existence of arithmetical progressions in "fat" subsets (i.e., subsets with positive density) is Roth's theorem [12] (1956):

If $A \subset \mathbb{Z}$ has positive density, then A contains infinitely many arithmetic progressions of length 3.

In general, the problem of existence of arbitrarily large arithmetical progressions in a positive density subset was solved by Szemerédi [14] (1975):

Theorem 1.3.1 (Szemerédi)**.** *If $A \subset \mathbb{Z}$ has positive density, then A has arbitrarily large arithmetical progressions.*

The extension of Szemerédi theorem to a more general context is the subject of the chapter 2 of this book (since, as stated, this theorem can't be applied to the subset of prime numbers). In the subsequent sections, we will give a proof of Szemerédi theorem assuming an important result of Ergodic Theory (namely, Furstenberg multiple recurrence theorem).

1.3.3 Prime number theorem

The basic obstruction to the application of Szemerédi theorem to the subset of prime numbers is provided by the famous prime number theorem:

Theorem 1.3.2 (Prime number theorem)**.** *It holds:*

$$\frac{|P \cap [1, N]|}{N} = \frac{1}{\log N}\,(1 + o(1)).$$

Here P is the subset of prime numbers and $o(1)$ is a quantity converging to zero when $N \to \infty$. In particular, $d(P) = 0$.

Although the prime numbers form a subset of zero density, the existence of infinitely many arithmetical progressions of length 3 formed only by prime numbers was showed in 1939 by Van der Corput (before Roth's theorem):

There are infinitely many arithmetical progressions of length 3 *composed only by prime numbers.*

Finally, in 2004, Ben Green and Terence Tao [9] proved the general result (about arbitrarily large arithmetic progressions formed only by primes). This theorem is the main object of the first two chapters of this book:

Theorem 1.3.3 (Green-Tao)**.** *The primes contains arbitrarily large arithmetical progressions.*

1.3.4 Erdös-Turán conjecture

We know that $\sum \frac{1}{n^2}$ converges (to $\pi^2/6$), but, in 1737, Euler showed that the serie of the inverse of the primes is divergent:

$$\sum_{p \text{ prime}} \frac{1}{p} = +\infty.$$

This means that the prime numbers are less sparse than the perfect squares.

The Erdös-Turán conjecture claims that any set with this property contains arbitrarily large arithmetical progressions (so that Green-Tao theorem solves a particular case of this conjecture):

Conjecture 1 (Erdös-Turán)**.** Let $A \subset \mathbb{N}$ be a subset such that

$$\sum_{n \in A} \frac{1}{n} = +\infty.$$

Then, A contains arbitrarily large arithmetical progressions.

This conjecture is completely open (in general): we don't know even when such subsets contain arithmetic progressions of length 3.

1.4 Proof of Van der Waerden theorem

During this section, we will present two proofs of Van der Waerden theorem:

Theorem 1.4.1 (Van der Waerden)**.** *For any coloring of the integers with a finite number of colors m, we can find arbitrarily large* monochromatic *(i.e., one color) arithmetic progressions.*

1.4.1 Combinatorial proof

In this subsection we prove Van der Waerden theorem via the coloring method in Combinatorics. In order to alleviate the notation, we denote the arithmetical progression $a, a+r, \dots, a+(k-1)r$ by $a+[0,k)r$, and we assume that one disposes of m colors to assign to the natural numbers from 1 to N.

Definition 1.4.1. *Let $c : \{1,\dots,N\} \to \{1,\dots,m\}$ be a coloring. Given $k \geq 1$, $d \geq 0$ and $a \in \{1\dots,N\}$, a* fan *of radius k, degree d with base point a is a d-tuple of arithmetic progressions $(a+[0,k)r_1,\dots,a+[0,k)r_d)$ where $r_1,\dots,r_d > 0$. For each $1 \leq i \leq d$, the progressions $a+[1,k)r_i$ are called* spokes *of the fan. We say that a fan is* polychromatic *if its base point and its spokes are* monochromatic*, i.e., there are distinct colors $c_0, c_1, \dots, c_d$ such that $c(a) = c_0$ and $c(a+jr_i) = c_i$ for $j = 1,\dots,k$ and $i = 1,\dots,d$.*

Remark 1.4.1. *Observe that, by the distinction between the colors, if we have m colors, it is not possible to construct a polychromatic fan whose degree is $\geq m$.*

Of course, we see that the van der Waerden theorem is a direct consequence of the following result:

Theorem 1.4.2. *Let $k, m \geq 1$. Then, there exists N such that any coloring of $\{1,\dots,N\}$ with m colors contains a monochromatic arithmetic progression of length k.*

Proof. The argument consists into a double induction scheme. Firstly, we make an inductive argument on k: observe that the case $k = 1$ is trivial, so that we can take $k \geq 2$ and we can assume that the theorem holds for $k-1$. Secondly, we perform an induction on d, i.e., we will show the following claim by induction: given d, there exists N such that for any coloring of $\{1,\dots,N\}$ with m colors, we have either a monochromatic arithmetic progression of length k or a polychromatic fan of radius k and degree d. Note that the case $d = 0$ is trivial and once we prove this claim for $d = m$, one can use the remark 1.4.1 in order to obtain the desired monochromatic arithmetic progression of length k (so that the double inductive argument is complete).

Let us take $d \geq 1$ and suppose that this claim is true for $d-1$. Let $N = 4kN_1N_2$, where N_1 and N_2 are large integers to be chosen later, and consider $A = \{1,\dots,N\}$. Fix $c : \{1,\dots,N\} \to \{1,\dots,m\}$ a coloring of A. Obviously, $\{bkN_1+1,\dots,bkN_1+N_1\}$ is a subset of A with N_1 elements for each $b = 1,\dots,N_2$. By our inductive hypothesis on k and d, if N_1 is sufficiently large, we can find either a monochromatic arithmetic progression of length k or a polychromatic fan of radius k and degree $d-1$.

Of course, if we find a monochromatic arithmetic progression of length k inside $\{bkN_1+1,\dots,bkN_1+N_1\}$ for some $b = 1,\dots,N_2$, we are done.

Thus, one can suppose that we find a polychromatic fan inside $\{bkN_1+1,\dots,bkN_1+N_1\}$ for every $b=1,\dots,N_2$. In other words, for each $b=1,\dots,N_2$, we have $a(b),r_1(b),\dots,r_{d-1}(b)\in\{1,\dots,N_1\}$ and *distinct* colors $c_0(b),c_1(b),\dots,c_{d-1}(b)\in\{1,\dots m\}$ such that $c(bkN_1+a(b))=c_0(b)$ and $c(bkN_1+a(b)+jr_i(b))=c_i(b)$ for every $j=1,\dots,k-1$ and $i=1,\dots,d-1$. We say that these are the first and second properties of the fan associated to b. In particular, the map

$$b\to(a(b),r_1(b),\dots,r_{d-1}(b),c_0(b),\dots,c_{d-1}(b))$$

is a *coloring* with $m^dN_1^d$ colors of the set $\{1,\dots,N_2\}$. Using again our inductive hypothesis on k, if N_2 is sufficiently large, there exists some arithmetic progression $b+[0,k-1)s$ which is monochromatic with respect to this new coloring, say that its color has the form $(a,r_1,\dots,r_{d-1},c_1,\dots,c_{d-1})$. Up to reversing the position of the progression, we can suppose that s is negative.

At this point, the idea is to convert this huge progression of *identical* polychromatic fans of degree $d-1$ (in the sense that their combinatorial type is fixed by the coloring $(a,r_1,\dots,r_{d-1},c_1,\dots,c_{d-1})$) in a new polychromatic fan with degree d in order to close the inductive argument. Let $b_0=(b-s)kN_1+a\in\{1,\dots,N\}$ and consider:

$$(b_0+[0,k)skN_1,b_0+[0,k)(skN_1+r_1),\dots,b_0+[0,k)(skN_1+r_{d-1})).$$

We affirm that this is a fan of radius k, degree d and base point b_0.

Indeed, let us verify that the spokes are monochromatic. In the first spoke we have $c(b_0+jskN_1)=c((b+(j-1)s)kN_1+a)$ by direct substitution. By the first property of the fan associated to $b+(j-1)s$, it follows that $c((b+(j-1)s)kN_1+a)=c_0(b+(j-1)s)=c_0(b)$ (since the arithmetic progression $b+[0,k-1)s$ is monochromatic if $1\le j\le k-1$). Similarly, in an arbitrary spoke, using the second property of the fans, we have that, if $1\le j\le k-1$ and $1\le t\le d$, then

$$c(b_0+j(skN_1+r_t))=c((b+(j-1)s)kN_1+a+jr_t)=c_t(b+(j-1)s)=c_t.$$

If the base point b_0 has the same color of a spoke, we found a monochromatic arithmetic progression of length k. Otherwise, the base point has a distinct color from the spokes, so that we found a polychromatic fan of radius k and degree d. This ends the inductive step and, *a fortiori*, the proof of the theorem. □

1.4.2 Dynamical proof

An useful tool in Dynamical Systems is the so-called *symbolic dynamics* consisting on the study of a specific map called *shift*. In the sequel, we will introduce the precise definition of the shift map and we will see how this important tool was applied by Furstenberg and Weiss to give a proof of van der Waerden theorem.

Let $A = \{a_1, \dots, a_k\}$ be a finite alphabet. Consider the set Ω of all infinite words obtained from the letters of this alphabet:

$$\Omega = \{(x_1, x_2, \dots, x_n, \dots) \; ; \; x_i \in A, \forall\, i\}.$$

This set has a natural structure of metric space with respect to the following distance: given $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$, define

$$d(x, y) := \frac{1}{l} \text{ if } l \text{ is the smallest integer such that } x_l \neq y_l.$$

The shift map $T : \Omega \to \Omega$ is:

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots).$$

It is a simple exercise to show that the shift map is continuous with respect to the distance d.

From these concepts, Furstenberg proved the van de Waerden theorem via the following topological dynamical theorem (whose complete proof is presented in the appendix to this chapter):

Theorem 1.4.3 (Topological Multiple Recurrence - Furstenberg and Weiss)**.** *Let $T : X \to X$ be a continuous dynamical system on a compact metric space X. For all $k \in \mathbb{N}$ and $\varepsilon > 0$, there exist $x \in X$ and $n \in \mathbb{N}$ such that $d(T^{in}(x), x) < \varepsilon$ for every $i = 1, \dots, k$. Moreover, given any dense subset $Z \subset X$, we can take $x \in Z$.*

Assuming this result, let us see how one can prove van der Waerden theorem. Let $A = \{c_1, \dots, c_s\}$ be the set of colors and $z = (z_1, z_2, z_3, \dots) \in A^{\mathbb{N}}$ a given coloring of $\mathbb{N}$, where $z_i \in A$ is the color of the integer i. Consider $T : A^{\mathbb{N}} \to A^{\mathbb{N}}$ the shift map. From the definition of the distance d, we have that, for $x, y \in A^{\mathbb{N}}$ and $m, l \in \mathbb{N}$, it holds $d(T^m(x), T^l(y)) < 1$ if and only if $x_{m+1} = y_{l+1}$.

In particular, for a given coloring $z \in A^{\mathbb{N}}$, an arithmetic progression $m, m+n, \dots, m+kn$ is monochromatic if and only if $z_m = z_{m+n} = \cdots = z_{m+kn}$, that is, if and only if:

$$\begin{aligned} d(T^{m-1}(z), T^{m-1+in}(z)) &= d(T^{m-1}(z), T^{in}(T^{m-1}(z))) \\ &< 1, \text{ for } i = 1, \dots, k. \end{aligned}$$

Taking $X = \overline{\{T^m(z)\}_{m=0}^{\infty}}$, we see that X is a compact metric space, T is a continuous dynamical system on X and the subset $Z = \{T^m(z)\}_{m=0}^{\infty}$ is dense in X. Thus, the van der Waerden theorem follows directly from the topological multiple recurrence theorem (Theorem 1.4.3).

1.5 Furstenberg theorem and its application to Szemerédi theorem

In this section we present a proof of Szemerédi theorem inspired by the dynamical proof of van der Waerden theorem: firstly, we will make a quick revision of basic elements in Ergodic Theory, then we will state a deep *multiple recurrence* result of Furstenberg, and finally we will get Szemerédi theorem as a consequence of Furstenberg theorem.

1.5.1 Crash course in Ergodic Theory

Ergodic Theory studies the *statistics* of the dynamics of a (measurable) map $T : X \to X$, where X is a probability space, from the point of view of a T-invariant probability measure μ (i.e., for any measurable subset A we have $\mu(A) = \mu(T^{-1}(A)))$.

Usually, the mere existence of a T-invariant probability measure gives us a lot of information about the statistics of generic orbits of T (i.e., the subsets $\{T^n(x)\}_{n=0}^{\infty}$, for almost every $x \in X$ with respect to μ). For instance, Poincaré recurrence theorem says that if $T : X \to X$ is μ-invariant and $\mu(A) > 0$, then for μ-almost every $x \in A$, there exists $n(x) \geq 1$ such that $T^{n(x)}(x) \in A$. Consequently, there exists N such that

$$\mu(A \cap T^{-N}(A)) > 0.$$

In particular, we see that, independently of the size of a given subset, if it has positive measure, then there are plenty of orbits starting at this subset and coming back to it infinitely many times. In the case of a topological probability space, one can reformulate Poincaré recurrent theorem as:

Let $T : X \to X$ be a dynamical system of a probability space (X, μ). Assume that X is also a compact metric space and μ is T-invariant. Then, μ-almost every point is recurrent, i.e., for a μ-generic point x, there exists a sequence $n_k \to \infty$ of natural numbers such that $d(T^{n_k}(x), x)) \to 0$ when $k \to \infty$.

After knowing that the dynamics T enjoys nice statistical properties *once* it has an invariant probability measure, a natural question arises: which dynamical systems possess invariant probability measures? When the space X is compact and T is continuous, the answer is *yes.* The idea of the proof of this fact is quite simple: take *any* probability measure (e.g., a Dirac measure at some arbitrary point) and let us analyze the evolution of this measure under the dynamics (that is, by the action of the iterates of T). By picking the average of the measures obtained up to a large iterate N, we hope to get an *almost insensitive* probability measure with respect to the action of T (as N grows). Thus, the natural argument is to consider an accumulation point of this sequence of almost insensitive measures and (by crossing fingers) one can expect that the limit probability measure is the desired T-invariant probability.

Now we put some details into the previous rough scheme. Firstly, since our scheme involves the process of taking limits of probability measures, let us introduce a notion of convergence of probability measures. Since the space of (Radon) measures is the dual space of continuous functions, it is natural to use the weak-* topology (because the Functional Analysis results can help us with the compactness issues we are going to face in a few moments).

Definition 1.5.1. *We say that a sequence of measures μ_k on X converges (weakly-*) to μ whenever, for any continuous function $f : X \to \mathbb{R}$, it holds*

$$\int_X f d\mu_k \to \int_X f d\mu.$$

Because this is the so-called weak-* topology on the concrete space of Radon measures, we can apply Banach-Alaoglu theorem to obtain:

The space of probability measures on X is compact with respect to the weak- topology.*

Now let us come back to the question of existence of invariant measures. Let η be an *arbitrary* probability measure. The action of T on η occurs by *push-forward*, i.e., $((T^n)^*\eta)(A) := \eta(T^{-n}(A))$ for every measurable subset A. A simple observation is: a given probability η is T-invariant if and only if $T^*\eta = \eta$.

Next, let us consider the sequence of probabilities

$$\mu_k = \frac{1}{k}\sum_{i=0}^{k-1}(T^i)^*\eta.$$

In other words, we are taking temporal averages of the measures obtained by the push-forward of η under the first $k-1$ iterates. By compactness, there is a convergent subsequence μ_{n_k} accumulating some probability measure μ. We claim that μ is invariant. In fact, we have the following equalities (explained below):

$$\begin{aligned}
T^*\mu &= T^*(\lim \mu_{n_k})\\
&= \lim(T^*(\mu_{n_k}))\\
&= \lim(\frac{1}{n_k}\sum_{i=0}^{n_k-1}(T^{i+1})^*(\eta))\\
&= \lim(\frac{1}{n_k}(\sum_{i=0}^{n_k}(T^i)^*(\eta) - \eta + (T^{n_k})^*\eta))\\
&= \lim \frac{1}{n_k}\sum_{i=0}^{n_k}(T^i)^*(\eta)\\
&= \mu.
\end{aligned}$$

In the second equality, we used the fact that the push-forward operator T^* is continuous in the weak-* topology. This is true because T is continuous: indeed, suppose that $\mu_k \to \mu$ weakly-* and fix a continuous $f : X \to \mathbb{R}$, so that we have that $f \circ T$ is also continuous and, *a fortiori*,

$$\int_X f d(T^*\mu_k) = \int_X f \circ T d\mu_k \to \int_X f \circ T d\mu = \int_X f d(T^*\mu).$$

In the fifth equality, we observe that, for every continuous $f : X \to \mathbb{R}$, by compactness of X, we have:

$$\frac{1}{n_k}\int_X f d\mu \to 0 \text{ and } \frac{1}{n_k}\int_X f d((T^{n_k})^*\mu) = \frac{1}{n_k}\int_X f \circ T^{n_k} d\mu \to 0.$$

Hence, the two last parts of this sum go to zero (weakly-*).

A concrete interesting example for our purposes is the shift map T on $X = \{0,1\}^{\mathbb{N}}$. Considering a Dirac measure associated to a point $x \in X$, that is, $\delta_x(A) = 0$ if $x \notin A$ and $\delta_x(A) = 1$ if $x \in A$, then we know that the sequence $\mu_k = \frac{1}{k}\sum_{j=0}^{k-1}\delta_{T^j(x)}$ accumulates (weakly-*) some probability measure and any such accumulation point is an invariant measure of the shift map.

1.5.2 Furstenberg theorem

Going back to Poincaré recurrence theorem, given a positive measure subset A, one can ask whether there is some *structure* on the set of return times to A. More precisely, we know that this set is infinite, but, does it have any arithmetic structure? For instance, can it coincide with the set of prime numbers?[1] In this direction, a beautiful result of Furstenberg (called Furstenberg multiple recurrence theorem) gives us a precise answer:

[1] Actually, Birkhoff theorem says that the density of the set of return times is positive, so that this set can't coincide with the prime numbers.

Theorem 1.5.1 (Furstenberg's multiple recurrence theorem)**.** *Let $T : X \to X$ be a μ-invariant map, $k \geq 3$ an integer and $\mu(A) > 0$. Then, there exists N such that*

$$\mu(A \cap T^{-N}(A) \cap \cdots \cap T^{-(k-1)N}(A)) > 0.$$

This deep theorem is the heart of Furstenberg ergodic-theoretical proof of Szemerédi theorem. Unfortunately, the complete proof of this result would lead us too far away from the scope of this book, so that we will content ourselves with a proof of this result into two important *representative* cases.

The first case is the *Bernoulli system.* Again, T is the (full) shift map on the set $X = A^{\mathbb{N}}$, where A is a finite alphabet. Next, we pick $p_1, \dots, p_r$ non-negative real numbers such that $\sum p_i = 1$. This provides a probability measure on A and, by taking the associated product measure, we get a probability measure on X. The dynamical system T equipped with this probability measure is called Bernoulli system.

By definition, the *product σ-algebra* is generated by the *cylinders subsets* one obtains by fixing a finite number n of coordinates, i.e., a cylinder is a subset of the form $C = \{w \in Z; w_{i_1} = j_1, \dots, w_{i_n} = j_n\}$ and its Bernoulli measure is $\mu(C) = p_{j_1} \dots p_{j_n}$. After extending this definition to the whole σ-algebra, it is a simple task to show that the Bernoulli measure is invariant by the shift map (in fact, it suffices to prove that $\mu(B) = \mu(T^{-1}(B))$ when B is a cylinder and this fact is easy to check).

In the same manner, since the cylinders generate the σ-algebra, it suffices to show Furstenberg multiple recurrence theorem to every cylinder subset. Consider $C_0, C_1, \dots, C_k$ some cylinders and observe that, for a sufficiently large integer n, the fixed coordinates in the definitions of the cylinders $T^{-nl}(C_l)$ are *distinct.* Hence, we obtain:

$$\mu(C_0 \cap T^{-n}(C_1) \cap \cdots \cap T^{-kn}(C_k)) = \mu(C_0)\mu(C_1) \dots \mu(C_k) > 0.$$

This proves Furstenberg theorem in this first example.

Another example is a *periodic* system, i.e., a dynamical system T such that $T^p = T$ for some p. In this case, Furstenberg theorem is totally trivial. A slightly less trivial dynamics (still along this line of reasoning) is the following *quasi-periodic* example: the space $X = S^1 = \mathbb{R}/\mathbb{Z}$ is the circle, μ is the Lebesgue measure and $T(x) = x + \alpha (\text{mod } 1)$ is a rotation of α (of course, the name quasi-periodic is motivated by the fact that T is periodic if and only if α is a rational number and the dynamics of T is close to periodic when α is irrational by taking approximations of α by rational numbers).

Given a measurable subset A with $\mu(A) > 0$, note that the function $\int 1_A(x+y) d\mu(x)$ is continuous on y. Hence, for all $\varepsilon > 0$, there exists δ

such that, if $|y| < \delta$, then $\mu(A \cap (A - y)) > \mu(A) - \varepsilon$. Hence

$$\mu(A \cap (A - y) \cap (A - 2y) \cap \cdots \cap (A - ky)) > \mu(A) - (k + 1)\varepsilon.$$

Choosing $\varepsilon < \frac{\mu(A)}{k+1}$ and fixing the corresponding δ, we can define $D_\delta = \{n \geq 1; n\alpha \in (-\delta, \delta)(mod\ 1)\}$. Observe that, if $n \in D_\delta$, it holds:

$$\mu(A \cap T^{-n}(A) \cap \cdots \cap T^{-nk}(A)) > \mu(A) - (k + 1)\varepsilon > 0.$$

This proves Furstenberg theorem in this second example.

Notice that the first example belongs to the class of *weak-mixing* systems, i.e., the class of dynamical systems verifying the following equality

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} (\mu(A \cap T^{-n}B) - \mu(A)\mu(B))^2 = 0$$

for any two measurable subsets A and B. By an adaptation of the ideas used to treat the case of Bernoulli shifts, it is possible to show that Furstenberg theorem holds for any weak-mixing (i.e., pseudorandom) system.

On the other hand, the second example belongs to the class of *compact* systems, i.e., the class of dynamical systems such that, for any function $f \in L^2(\mu)$, the closure of $\{f, Tf, T^2f, \ldots, T^nf, \ldots\}$ is compact in L^2. Again, by an adaptation of the ideas used to deal with the case of quasiperiodic rotations of the circle, it is possible to prove that Furstenberg theorem also holds for any compact (i.e., structured) system.

In general, Furstenberg theorem is a consequence of the so-called *Furstenberg structure theorem* saying that we can decompose an arbitrary dynamical system into several levels (i.e., *factors*) along a *tower of extensions* such that each level (factor) is weakly mixing or compact and any two distinct factors doesn't correlate much. In fact, since these factors don't interact and we know Furstenberg theorem for weakly-mixing and compact factors, we are morally done. Evidently, the existence of this tower of extensions is a highly non-trivial fact beyond the scope of this book (so that we will end our comments on Furstenberg theorem here). However, it is worth to point out that this structure theorem will reappear later (on chapter 2) in a *finitary context* during the proof of Green-Tao theorem.

1.5.3 Szemerédi theorem via multiple recurrence

Once we have Furstenberg multiple recurrence theorem in our toolbox, one can quickly give a proof of Szemerédi theorem using the shift dynamics (in the spirit of the ergodic proof of van der Waerden theorem).

Put $X = \{0, 1\}^{\mathbb{N}}$ and let $T : X \to X$ be the shift. Take $(x_n) = (1_A(n))$, where $1_A(x)$ is the characteristic function of A, and consider

$\mu_k = \frac{1}{k}\sum_{j=0}^{k-1}\delta_{T^j(x)}$. Then, as we already know, up to passing to a subsequence, we can assume that $\mu = \lim \mu_k$ is a T-invariant probability measure.

Define $Y = \{(y_n); y_1 = 1\}$. Since Y is a compact subset, we have that $\mu(Y) = \lim \mu_k(Y) = \lim \frac{1}{k}|A \cap [1,k]| > 0$ (by hypothesis). Thus, by Furstenberg multiple recurrence theorem, it follows that there exists N such that $\mu(Y \cap T^{-N}(Y) \cap \cdots \cap T^{-(k-1)N}(Y)) > 0$. In particular, there is $z \in Y \cap T^{-N}(Y) \cap \cdots \cap T^{-(k-1)N}(Y)$. I.e., there exists some integer x such that $x, x+N, \dots, x+(k-1)N \in A$. This ends the proof of Szemerédi theorem.

1.6 Quantitative Szemerédi theorem

In this section, we will see some reformulations of Szemerédi theorem 1.3.1 (which are very useful for our future purposes).

Let's start with the remark that Szemerédi theorem is *equivalent* to the following statement:

For any $k \geq 1$ and $0 < \delta \leq 1$, there exists a large integer $N_{SZ}(k,\delta) \geq 1$ such that, for every $N \geq N_{SZ}$, any subset $A \subset \{1, \dots, N\}$ of cardinality $|A| \geq \delta N$ contains some arithmetic progression of length k.

Logically, this statement is *a priori* certainly stronger than Szemerédi theorem.

On the other direction, we will use an abstract non-sense argument: admitting that the statement is false for a certain pair (k, δ), we affirm that there is a subset $Y \subset \mathbb{N}^*$ satisfying $|Y \cap \{1, 2, \dots, n\}| \geq \delta r$, $\forall\, r \in \mathbb{N}^*$ so that Y doesn't contain any arithmetical progression of length k.

In order to prove this affirmation, let us first prove that the non-existence of $N_{SZ}(k,\delta)$ implies that, for each $n \in \mathbb{N}^*$, there exists a subset $X_n \subset \{1, 2, \dots, n\}$ verifying $|X_n \cap \{1, 2, \dots, k\}| \geq \delta k$ for $1 \leq k \leq n$ such that X_n doesn't contain any arithmetic progression of length k. In fact, let $\varepsilon_n = \max\limits_{1 \leq k \leq n}((\lceil \delta k \rceil - 1)/k) < \delta$. We claim that, if N is sufficiently large and $A \subset \{1, 2, \dots, N\}$ has cardinality $|A| \geq \delta N$, then there exists $m \leq N - n$ such that $|A \cap \{m+1, \dots, m+k\}| \geq \delta k$, for each $1 \leq k \leq n$. Indeed, if this is not the case, there are $s \in \mathbb{N}^*$, $k_1, k_2, \dots, k_s \in \{1, 2, \dots, n\}$ such that $N \geq k_1 + k_2 + \cdots + k_s > N - n$ and, for $1 \leq r \leq s$, it holds $|A \cap \big(\sum\limits_{j<r} k_j, \sum\limits_{j \leq r} k_j\big]| < \delta k_r$, so that $\dfrac{1}{k_r}\,|A \cap \big(\sum\limits_{j<r} k_j, \sum\limits_{j \leq r} k_j\big]| \leq \varepsilon_n < \delta$, and, *a fortiori*, $\delta N \leq |A| \leq n + \varepsilon_n \cdot N$, an absurd when $N > \dfrac{n}{\delta - \varepsilon_n}\cdot$ Observe that this proves the desired affirmation about the existence of X_n (assuming that $N_{SZ}(k,\delta)$ doesn't exist) because, by the previous discussion, it suffices to consider an appropriate translation of a certain subset of any $A \subset$

$\{1,2,\dots,N\}$ with cardinality $|A|\geq\delta N$ such that A doesn't contain any arithmetic progression of length k.

Next, once we know about the existence of these special subsets X_n, let us cook up the desired Y. To do so, for each $r\in\mathbb{N}^*$, let $\pi_r\colon 2^{\mathbb{N}}\to 2^{\{1,2,\dots,r\}}$ be given by $\pi_r(A)=A\cap\{1,2,\dots,r\}$. We construct inductively some sets $Y_1,Y_2,Y_3,\dots$ with $Y_r\subset\{1,2,\dots,r\}$ for each $r\in\mathbb{N}^*$ such that $Y_{r+1}\cap\{1,2,\dots,r\}=Y_r$, $\forall\, r\in\mathbb{N}^*$ in the following way. Firstly, we put $Y_1:=\{1\}\subset X_n$, for all $n\in\mathbb{N}^*$. Next, given Y_r, $r\in\mathbb{N}^*$ such that $Y_r=\pi_r(X_n)$ for infinitely many $n\in\mathbb{N}$, there exists $Y_{r+1}\subset\{1,2,\dots,r+1\}$ with $Y_{r+1}\cap\{1,2,\dots,r\}=Y_r$ such that $Y_{r+1}=\pi_{r+1}(X_n)$ for infinitely many $n\in\mathbb{N}$ (indeed, if $\pi_r(X_n)=Y_r$, there are only two possibilities for $\pi_{r+1}(X_n)$). Now, it is easy to see that $Y=\bigcup\limits_{n\in\mathbb{N}^*}Y_n$ verifies $\pi_r(Y)=Y_r$, $\forall\, r\in\mathbb{N}^*$, so that $\pi_r(Y)=\pi_r(X_n)$ for infinitely many $n\in\mathbb{N}$. In particular, $|Y\cap\{1,2,\dots,r\}|\geq\delta r$, $\forall\, r\in\mathbb{N}^*$ and Y doesn't contain any arithmetic progression of length k.

After this, we introduce a more analytical and finitary language in order to get another reformulation of Szemerédi theorem (Theorem 1.3.1). Keeping this goal in mind, we recall the following definition:

Definition 1.6.1. *Let $f:A\to\mathbb{C}$ be an arbitrary function where A is a finite set. Then, the expectation of f is*

$$\mathbb{E}(f)=\mathbb{E}(f(n);n\in A)=\frac{1}{|A|}\sum_{n\in A}f(n).$$

Given $f:(\mathbb{Z}/N\mathbb{Z})\to\mathbb{R}$ any function, we can define the shift $T^nf:(\mathbb{Z}/N\mathbb{Z})\to\mathbb{R}$ of this function f by an integer $n\in\mathbb{Z}/N\mathbb{Z}$ (or $n\in\mathbb{Z}$) via the formula $T^nf(x):=f(x+n)$.

Using this notation, we can reformulate Szemerédi theorem as follows:

Theorem 1.6.1 (Szemerédi theorem – quantitative version)**.** *For every integer number $k\geq 1$ and real number $0<\delta\leq 1$, there are $N_0(k,\delta)$ a large integer and $c(k,\delta)>0$ a small real number such that, for any $N\geq N_0(k,\delta)$ a large prime number and any $f:\mathbb{Z}/N\mathbb{Z}\to\mathbb{R}^+$ with $0\leq f\leq 1$, $\mathbb{E}(f|\mathbb{Z}/N\mathbb{Z})\geq\delta$, it holds*

$$\mathbb{E}(\prod_{j=0}^{k-1}T^{jr}f(x)|x,r\in\mathbb{Z}/N\mathbb{Z})\geq c(k,\delta).$$

Remark 1.6.1. *Concerning this quantitive version of Szemerédi theorem, we will construct in the appendix to this chapter some examples due to F. Behrend of some subsets S of the interval $[1,N]$ such that $|S|\geq N^{1-\frac{2\sqrt{2\log 2}+\varepsilon}{\sqrt{\log N}}}$ and S doesn't contains arithmetic progressions of length*

3. *Moreover, by a slight modification of the scheme of Behrend's argument, we will see that, concerning the behavior of $c(k,\delta)$ above with respect to δ, one can't expect that $c(k,\delta)$ has a polynomial behavior in the variable δ (i.e., $c(k,\delta)\geq\delta^{C_k}$ for some $C_k>0$): indeed, we will show that $c(3,\delta)\leq\delta^{c\log(1/\delta)}$.*

Observe that the statement of theorem 1.6.1 provides (*a priori*) a much stronger conclusion than the usual Szemerédi theorem. In fact, while the usual Szemerédi allows only to conclude the existence of *one* k-AP (i.e., arithmetic progression of length k), the quatitative version permits to infer the existence of $c(k,\delta)N^2$ k-APs (at least). However, although the quantitative Szemerédi theorem is apparently better than the usual one, we claim that the theorems 1.3.1 and 1.6.1 are *equivalents*.

We start by showing that the usual Szemerédi theorem follows from its quantitative version. Fix k, δ and take N a large prime number. Let us suppose that $A\subset\{1,\dots,N\}$ has cardinality $|A|\geq\delta N$ (this is plausible since A has positive density). Pick N' a prime number between kN and $2kN$ (its existence is assured by Bertrand's postulate). Consider $\{1,\dots,N\}$ as a subset of $\mathbb{Z}/N'\mathbb{Z}$ and denote by A' the subset of $\mathbb{Z}/N'\mathbb{Z}$ corresponding to A.

By our choices, we have $E(1_{A'}|\mathbb{Z}/N\mathbb{Z})\geq\delta/2k$. By the quantitative version of Szemerédi theorem, it follows that:

$$E(\prod_{j=0}^{k-1}T^{jr}1_{A'}(x)|x,r\in\mathbb{Z}/N'\mathbb{Z})\geq c(k,\delta/2k).$$

Rewriting this expression, we get:

$$|\{(x,r)\in(\mathbb{Z}/N'\mathbb{Z})^2;x,x+r,\dots,x+(k-1)r\in A'\}|\geq c(k,\delta/2k)(N')^2.$$

Since $N'\geq kN$ and $A'\subset\{1,\dots,N\}$, we have that $1\leq x\leq N$ and $-N\leq r\leq N$. Observe that the contribution of the (degenerate) progressions with $r=0$ is N (at most). Removing these degenerate progressions and taking N large, the right-hand side is still positive, so that A must contain a progression $x,x+r,\dots,x+(k-1)r$.

Now let us prove that the usual Szemerédi theorem implies its quantitative version. Recall that we know that the usual version is equivalent to its finitary version, i.e., the existence of the integer $N_{SZ}(k,\delta)$, $\forall\,k\in\mathbb{N}^*$, $\delta>0$ with the previously discussed properties. Thus, our task is reduced to the proof of the following proposition:

Proposition 1.6.1. *Suppose that the integer $N_{SZ}\big(k,\frac{\delta}{2}\big)$ with the previous properties does exist. Then, there are $N_0\in\mathbb{N}$ and $\alpha(k,\delta)>0$ such that, if $N\geq N_0$ for all $A\subset\{1,2,\dots,N\}$ with $|A|\geq\delta N$, there are $\alpha(k,\delta)\,N^2$ k-APs, i.e., arithmetic progressions of length k (at least).*

Proof. Let $m_0 = N_{SZ}(k, \delta/2)$. Then, for every $m \geq m_0$, any subset of $\{1, 2, \ldots, m\}$ with cardinality $\geq \delta m/2$ contains some arithmetic progression of length k. Let N be a large integer. For each $1 \leq r \leq \lfloor N/m_0 \rfloor$, we divide $\{1, 2, \ldots, N\}$ into r arithmetic progressions of ratio r, e.g., $\{1 \leq n \leq N \mid n \equiv a (\text{mod } r)\}$, for each a with $0 \leq a \leq r - 1$. Each of these APs has $\lfloor N/r \rfloor$ elements (at least), and, therefore, they can be decomposed into a union of $\lfloor \lfloor N/r \rfloor / m_0 \rfloor$ disjoint arithmetic progressions of ratios r, lengths $\geq m_0$ (and almost equal), so that their diameters belong to the interval $[r(m_0 - 1), r(2m_0 - 1)]$.

Now, if $A \subset \{1, 2, \ldots, N\}$ satisfies $|A| \geq \delta N$, we have that, for each r, $\#\{0 \leq a \leq r - 1 \mid \#A \cap \{1 \leq n \leq N \mid n \equiv a(\text{mod } r)\} \geq \frac{3\delta}{4} \lfloor N/r \rfloor\} \geq \frac{\delta r}{4 - 3\delta}$ (since $t < \frac{\delta}{4 - 3\delta} \Rightarrow t + \frac{3\delta}{4}(1 - t) < \delta$).

On the other hand, because $t < \frac{\delta}{4 - 2\delta} \Rightarrow t + \frac{\delta}{2}(1 - t) < \frac{3\delta}{4}$, if $\#A \cap \{1 \leq n \leq N \mid A \equiv a(\text{mod } r)\} \geq \frac{3\delta}{4} \lfloor N/r \rfloor$ for a certain r, then, it follows that $\frac{\delta}{4 - 2\delta} \cdot \lfloor \lfloor N/r \rfloor / m_0 \rfloor$ of the created arithmetic progressions of length $\geq m_0$ (at least) intersects A with a relative proportion of $\delta/2$ (at least), so that *a fortiori*, it should contain a k-AP. This gives us $\sum_{r=1}^{\lfloor N/m_0 \rfloor} \frac{\delta r}{4 - 3\delta} \cdot \frac{\delta}{4 - 2\delta} \lfloor \lfloor N/r \rfloor / m_0 \rfloor > \beta(\delta, m_0) N^2$ k-APs (at least) contained in A (for a large N), where $\beta(\delta, m_0) = \delta^2 / 64 m_0^2$. Of course, some of these APs can be double-counted sometimes (for different choices of r), but once we fix the diameter d of the AP, r must be a divisor of d between $\frac{d}{2m_0 - 1}$ and $\frac{d}{k - 1}$, i.e., $r = \frac{d}{r'}$, where $k - 1 \leq r' \leq 2m_0 - 1$. Consequently, there are $2m_0 - k + 1$ possibilities for r' (at most) and *a fortiori* for r, so that each AP is counted $2m_0 - k + 1$ times (at most). Hence, A contains at least $\alpha(k, \delta) N^2$ k-APs, where $\alpha(k, \delta) = \frac{\delta^2}{64 m_0^2 (2m_0 - k + 1)}$. □

Remark 1.6.2. *The basic difference between the proof of the quantitative Szemerédi theorem and its previous versions is the* finitary *nature of the arguments (allowing to explicit bounds on N_{SZ}). The proofs of the other versions are* infinitary *arguments (they use to some extend the Axiom of Choice) only permits us to show the existence of N_{SZ} without any bound on its magnitude. The strategy of the proof of the quantitative Szemerédi is used during the proof of Green-Tao theorem as we are going to see in the next chapter (we also recommend [16]).*

1.7 Further results

During this section, we indicate (without proofs) equivalent analytic formulations of some of the conjectures cited above. After that, we present some further results (historically more recent than Green-Tao theorem). Finally, we will make some comments about the nature of the integer $N_0(k,\delta)$.

1.7.1 von Mangoldt function

In order to reformulate some number-theoretical theorems and conjectures, we need to introduce the so-called von Mangoldt function.

Definition 1.7.1. *The von Mangoldt function $\Lambda : \mathbb{Z} \to \mathbb{R}^+$ is given by $\Lambda(n) = \log p$ if $n = p^r$ (for some $r \geq 1$), and $\Lambda(n) = 0$ otherwise.*

In these terms, observe that the unique factorization theorem can be expressed as:

$$\log n = \sum_{d|n} \Lambda(d). \tag{1.1}$$

Now we recall the definition of expectation of a function on a finite set:

Definition 1.7.2. *Given $f : X \to \mathbb{R}$ and $A \subset X$ a finite set, we define the (conditional) expectation of f with respect to A via the formula:*

$$\mathbb{E}(f(n)|n \in A) = \mathbb{E}(f|A) = \frac{1}{|A|}\sum_{n\in A} f(n).$$

In this setting, the prime number theorem can be seen as an estimate for the conditional expectations of the von Mangoldt function:

Theorem 1.7.1. *The prime number theorem is equivalent to*

$$\mathbb{E}(\Lambda|[1,N]) = 1 + o(1).$$

Proof. By the definition of the von Mangoldt function, we have:

$$\begin{aligned} N\mathbb{E}(\Lambda|[1,N]) &= \sum_{p\leq N}[\frac{\log N}{\log p}]\log p \leq \log N \sum_{p \leq N} 1 \\ &= \log N \cdot (|\text{primes in } [1,N]|). \end{aligned}$$

Dividing by N, we get that the prime number theorem implies the statement about the expectations of Λ.

On the other hand, if $1 < M < N$, then

$$\begin{aligned} |\text{primes in } [1,N]| &= |\text{primes in } [1,M]| + \sum_{M<p\leq N} 1 \\ &\leq |\text{primes in } [1,M]| + \sum_{M<p\leq N} \frac{\log p}{\log M} \\ &< M + \frac{1}{\log M} N\mathbb{E}(\Lambda|[1,N]). \end{aligned}$$

Now, if N is large, we have $1 < M = \frac{N}{\log^2 N} < N$. Combining this inequality with the previous estimate, we obtain:

$$|\text{primes in } [1,N]| < \frac{N}{\log^2 N} + \frac{N\mathbb{E}(\Lambda|[1,N])}{\log N - 2\log\log N}.$$

Therefore,

$$\frac{|\text{primes in } [1,N]|}{N} < \mathbb{E}(\Lambda|[1,N])(\frac{1}{\log N - 2\log\log N}) + \frac{1}{\log^2 N}.$$

This ends the proof because $\frac{\log x}{\log x - 2\log\log x} \to 1$ when $x \to \infty$. □

In fact, the precise knowledge of the expectations of von Mangoldt function[2] actually implies several famous conjectures. Let us give a list of these conjectures (without further details):

- The *Riemann hypothesis*[3] is equivalent to the following claim:

$$\mathbb{E}(\Lambda|[1,N]) = 1 + O(N^{-1/2}\log^2 N).$$

- The *twin prime conjecture* follows from the following affirmation:

$$\liminf_{N\to\infty} \mathbb{E}(\Lambda(n)\Lambda(n+2)|1 \leq n \leq N) > 0.$$

- The *Goldbach conjecture* is equivalent to:

$$\mathbb{E}(\Lambda(n_1)\Lambda(n_2)|n_1, n_2 \in [1,N] \text{ and } n_1 + n_2 = N) > 0 \ \forall\ N \text{ even}.$$

- The *odd Goldbach conjecture* is equivalent to:

$$\mathbb{E}(\Lambda(n_1)\Lambda(n_2)\Lambda(n_3)|n_1, n_2, n_3 \in [1,N] \text{ and } n_1+n_2+n_3 = N) > 0 \quad \forall\ N \text{ odd}.$$

[2] I.e., an explicit estimate for the speed of convergence to zero of the term $o(1)$ in Thereom 1.7.1.

[3] This famous conjecture, one of the seven Millennium prize problems of Clay Mathematical Institute (who offers 1 million dollars for its solution), says that the (non-trivial) zeros of Riemann zeta function $\zeta(s)$ (a complex-analytic function related to the prime numbers obtained by analytic continuation of $\sum_{n=1}^{\infty} 1/n^s$, $\Re(s) > 1$) are located in the line $\Re(s) = 1/2$.

1.7.2 Constellations of primes and polynomial progressions

A well-known set with a nice notion of primality is the set of Gaussian integers $\mathbb{Z}[i] := \{a + bi; a, b \in \mathbb{Z}\}$. Here, by definition, p is a Gaussian prime if it is *only* divisible by $\pm1, \pm i, \pm p$ and $\pm ip$.

A *shape* on $\mathbb{Z}[i]$ is a finite set $(v_j)_{j\in J} \in (\mathbb{Z}[i])^J$ of distinct Gaussian integers. A *constellation* in $\mathbb{Z}[i]$ with this shape is any J-tuple $(a + rv_j)_{j\in J} \in (\mathbb{Z}[i])^J$ of distinct Gaussian integers (where $a \in \mathbb{Z}[i]$ and $r \in \mathbb{Z}[i]$).

The notion of constellation extends the concept of arithmetic progressions to the context of Gaussian integers. The abundance of arbitrarily shaped constellations formed by Gaussian primes was proved by T. Tao [17]:

Let $(v_j)_{j\in J}$ be an arbitrary shape of Gaussian integers. Then, the set of Gaussian primes contains infinitely many constellations with this prescribed shape.

On the other hand, an alternative generalization of the concept of arithmetical progressions is: since any arithmetical progression has the form $x + P_1(m), \dots, x + P_k(m)$ where $P_i(m) = (i - 1)m$, one can extend this notion by simply allowing $P_i \in \mathbb{Z}[m]$ to be some integer-valued polynomials with $P_i(0) = 0$ (for each $i = 1, \dots, k$). These type of generalized progression are called *polynomial progressions.*

The existence of infinitely many polynimial progressions formed by prime numbers was showed by T. Tao and T. Ziegler [18]:

Let $P_1, \dots P_k$ be integer-valued polynomials with $P_i(0) = 0$. Given $\varepsilon > 0$, there are infinitely many integers x and m such that $x + P_i(m)$ are prime numbers for any $i = 1, \dots, k$ and $1 \le m \le x^\varepsilon$.

1.7.3 Gaps in the set of prime numbers

In a certain sense, all the theorems presented in this chapter have a common feature of searching patterns in the set P of prime numbers. In this direction, a natural problem is to know how sparse can P be.

In order to investigate this issue, one can denote by p_n the n-th prime number, so that the size of the n-th gap of P is $p_{n+1} - p_n$. The prime number theorem says that the *average* size of this gap is morally $\log p_n$. We define Δ as the smallest number such that there are infinitely many gaps of size less than $(\Delta + \varepsilon)$ times the average gap, i.e.,

$$\Delta = \liminf_{n\to\infty} \left(\frac{p_{n+1} - p_n}{\log p_n} \right).$$

It was conjectured that $\Delta = 0$ and this fact was recently proved by D. Goldston, J. Pintz and C. Yıldırım [7]. Also, in a more recent work, D. Goldston, J. Pintz and C. Yıldırım [8] were able to show the slightly stronger theorem:

$$\liminf_{n\to\infty}\left(\frac{p_{n+1}-p_n}{\sqrt{\log p_n}(\log\log p_n)^2}\right) < \infty.$$

In these works, the authors proposed a method to show the existence of large primes which are very close[4].

Remark 1.7.1. *Just to stress the difficulty of the twin prime conjecture, let us observe that this conjecture is much stronger than the result $\Delta = 0$ of Goldston, Pintz and Yıldırım (whose proof is highly non-trivial!).*

1.7.4 Magnitude of $N_0(k,\delta)$

Concerning the size of $N_0(k,\delta)$ in the quantitative version of Szemerédi theorem, we have the following results:

- T. Gowers showed that $N_0(k,\delta) \leq 2^{2^{\delta^{-c_k}}}$, where $c_k = 2^{2^{k+9}}$;
- R. Rankin proved that $N_0(k,\delta) \geq \exp(C(\log\frac{1}{\delta})^{1+\lfloor\log_2(k-1)\rfloor})$;
- J. Bourgain proved that $N_0(3,\delta) \leq 2^{C\delta^{-2}\log(1/\delta)}$;
- it is expected that $N_0(k,\delta) \leq 2^{c_k\delta^{-1}}$, but this is an open problem related to the Erdös-Turán conjecture.

1.8 Appendix to Chapter 1

1.8.1 Proof of Theorem 1.4.3

A preliminary observation is: if, for some k and ε, the first part of theorem holds for a certain x, then the same statement is true for an entire small neighborhood of x, and, *a fortiori*, the theorem works with some element of any fixed dense subset Z. Hence, it suffices to show the first part of the theorem to get a full proof of it.

Next, we notice that Zorn's lemma implies that one can assume that X is *minimal*, i.e., X doesn't possess any *proper* closed subset Y such that $T(Y) \subset Y$. Observe that, in this situation, the subsets $\{T^m(x)\}_{m=0}^{\infty}$ are dense in X, so that the theorem is true for $k = 1$ (since, by denseness, there exists some $n \in \mathbb{N}$ with $d(T^n(x), x) < \varepsilon$).

[4] Actually, they can show that the existence of infinitely many primes at a bounded distance *assuming* a conjecture of Elliot-Halberstam.

At this stage, the proof proceeds by induction (on k). Suppose that the theorem holds for some $k \geq 1$, i.e., for all $\varepsilon > 0$ there exists $x \in X$ and $n \in \mathbb{N}$ such that $d(T^{in}(x), x) < \varepsilon$ for each $i = 1, \dots, k$. We claim that the set of such points x is actually dense in X.

Indeed, let $U \subset X$ be an arbitrary open subset and pick $B \subset U$ a small ball of radius strictly less than ε. Define $B_m = (T^m)^{-1}(B)$, so that these subsets form an open cover of X (by the minimality assumption). Using the compactness of X, we can extract a finite subcover $\{B_{m_1}, \dots, B_{m_r}\}$. Let $\delta > 0$ be the Lebesgue number of this open subcover, that is, a number such that any ball of radius δ is contained inside some element of the subcover. Take x and n such that $d(T^{in}(x), x) < \delta$ for $i = 1, \dots, k$ (whose existence is assured by the inductive hypothesis) and denote by D the ball of center x and radius δ. Then, by our choice of δ, there exists j such that $D \subset B_{m_j}$. In particular, $T^{m_j}(D) \subset B$, that is, the elements $T^{m_j}(T^{in}(x))$ belong to the ball of radius ε centered on $T^{m_j}(x) \in U$. This proves our denseness claim.

Now, let's go back to the proof of the theorem. Fix $\varepsilon > 0$. By the inductive hypothesis, there are x_0 and n_0 such that $d(T^{in_0}x_0, x_0) < \varepsilon/2$ for $i = 1, \dots, k$. Taking x_1 such that $T^{n_0}(x_1) = x_0$, we have $d(T^{(i+1)n_0}x_1, x_0) < \varepsilon/2$ for $i = 1, \dots, k$. Hence, $d(T^{in_0}(x_1), x_0) < \varepsilon/2$ for $i = 1, \dots, k+1$.

By continuity, there exists $\varepsilon_1 < \varepsilon$ such that $d(y, x_1) < \varepsilon_1$ implies $d(T^{in_0}(y), x_0) < \varepsilon/2$ for $i = 1, \dots, k+1$. By our denseness claim, there are y_1 and n_1 such that $d(y_1, x_1) < \varepsilon_1/2$ and $d(T^{in_1}(y_1), y_1) < \varepsilon_1/2$ for $i = 1, \dots, k$. By the triangular inequality, we have:

$$d(T^{in_0}(T^{(i-1)n_1}(y_1)), x_0) < \varepsilon_2 \text{ for } i = 1, \dots, k+1.$$

Proceeding in this way (taking x_2 such that $T^{n_1}(x_2) = y_1$, etc.), we find a sequence of points $x_2, x_3, \dots \in X$ and a sequence of natural numbers $n_2, n_3, \dots$ such that, for each l, we have:

$$\begin{array}{rcll} d(T^{in_{l-1}}(x_l), x_{l-1}) & < & \varepsilon/2 & \\ d(T^{i(n_{l-1}+n_{l-2})}(x_l), x_{l-2}) & < & \varepsilon/2 & \\ & \cdots & & \\ d(T^{i(n_{l-1}+\cdots+n_0)}(x_l), x_0) & < & \varepsilon/2 & \text{for } i = 1, \dots, k+1. \end{array}$$

By compactness, there are $l > m$ such that $d(x_l, x_m) < \varepsilon/2$. By the triangular inequality, we have:

$$d(T^{i(n_{l-1}+\cdots+n_m)}(x_l), x_l) < \varepsilon \text{ , for } i = 1, \dots, k+1.$$

Therefore, it suffices to take $x = x_l$ and $n = n_{l-1} + \cdots + n_m$ to conclude the proof of the theorem.

1.8.2 F. Behrend's example

As we announced in the remark 1.6.1, firstly we will construct some examples of subsets S of natural numbers $\leq N$ such that S doesn't contain any arithmetic progression of length 3 and its cardinality is $|S| \geq N^{1-\frac{2\sqrt{2\log 2}+\varepsilon}{\sqrt{\log N}}}$; after that, we will adapt this technique to study the behavior of the function $c(3,\delta)$.

Given $d \geq 2$, $n \geq 2$ and $k \leq n(d-1)^2$, consider $S_k(n,d)$ the subset of all integer of the form

$$x = a_1 + a_2(2d-1) + \cdots + a_n(2d-1)^{n-1}$$

whose digits a_i in the basis $(2d-1)$ are subject to the constraints

$$0 \leq a_i < d \quad \text{ and } \quad \|x\|^2 := a_1^2 + \cdots + a_n^2 = k.$$

Notice that $S_k(n,d)$ *doesn't* contain 3-APs (arithmetic progressions of length 3): in fact, if this is not the case, there are $x, x', x'' \in S_k(n,d)$ such that $x + x' = 2x''$, so that

$$\|x + x'\| = \|2x''\| = 2\sqrt{k}$$

and

$$\|x\| + \|x'\| = 2\sqrt{k}.$$

Thus, since the equality in the triangular inequality $\|x + x'\| \leq \|x\| + \|x'\|$ can only occur when the vectors $(a_1, \ldots, a_n)$ and $(a_1', \ldots, a_n')$ are proportional, we see that $x = x' = x''$ (because the norms of these vectors are the same by hypothesis).

On the other hand, there are d^n vectors $(a_1, \ldots, a_n)$ satisfying the constraint $0 \leq a_i < d$ and there are $n(d-1)^2 + 1$ possible values of k. Consequently, for some $k = K_0$, the subset $S_k(n,d)$ must have cardinality (at least)

$$\frac{d^n}{n(d-1)^2+1} > \frac{d^{n-2}}{n}.$$

Since every element of $S_k(n,d)$ has modulus $< (2d-1)^n$, if we define

$$\nu(N) := \max\{|S|\,;\, S \subset [1,N],\ S \text{ without any 3-AP}\},$$

it holds

$$\nu((2d-1)^n) > d^{n-2}/n.$$

Now, for a fixed $\varepsilon > 0$ and for a given (large) N, we choose $n = \lfloor\sqrt{\frac{2\log N}{\log 2}}\rfloor$ and d satisfying

$$(2d-1)^n \leq N < (2d+1)^n,$$

so that

$$\begin{aligned}\nu(N) &\geq \nu((2d-1)^n) > \frac{d^{n-2}}{n} > \frac{(N^{1/n}-1)^{n-2}}{2^{n-2}n} \\ &= \frac{N^{1-(2/n)}}{2^{n-2}n}(1-N^{-1/n})^{n-2} \\ &> \frac{N^{1-(2/n)}}{2^{n-1}n} = N^{1-(2/n)-\frac{\log n}{\log N}-\frac{(n-1)\log 2}{\log N}} \\ &> N^{1-\frac{2\sqrt{2\log 2}+\varepsilon}{\sqrt{\log N}}}.\end{aligned}$$

Next, we will slightly modify this reasoning to study the behavior of $c(3,\delta)$: fix $d, n \geq 1$ integers (to be chosen later) and define $\phi : \{1,\dots,N\} \to \{0,\dots,2d-1\}^n$ by

$$\phi(x) := (\lfloor x/(2d-1)^i \rfloor \mathrm{mod}(2d-1))_{i=0}^{n-1}.$$

For each $1 \leq k \leq n(d-1)^2$, consider again the subsets

$$S_k(n,d) := \{(x_1,\dots,x_n) \in \{0,\dots,d-1\}^n \ : \ x_1^2+\dots+x_n^2 = k\}$$

and define $A_k(n,d) := \phi^{-1}(S_k(n,d))$. As we already know, $S_k(n,d)$ is free from 3-APs (*except* for the trivial 3-APs $\{x,x,x\}$). This implies that $A_k(n,d)$ can only contain arithmetic progressions $(n, n+r, n+2r)$ when r is a multiple of $(2d-1)^n$. In particular, the maximal number of 3-APs in $A_k(n,d)$ is $N^2/(2d-1)^n$. On the other hand, when $\phi(x) \in \{0,\dots,d-1\}^n$, the probability of x to belong to $A_k(n,d)$ is $\frac{1}{n(d-1)^2+1}$. Therefore, we get the following lower bound for the cardinality of $A_k(n,d)$:

$$|A_k(n,d)| \geq \frac{c}{nd^2}2^{-n}N.$$

Taking $n = c\log(1/\delta)$ and $d = \delta^{-c}$, we obtain that, for some k, the set $A_k(n,d)$ satisfies $|A_k(n,d)| \geq \delta^c N$ and the maximum number of 3-APs inside $A_k(n,d)$ is $\delta^{c\log(1/\delta)}N^2$. In other words, $c(3,\delta) \leq \delta^{c\log(1/\delta)}$.

Remark 1.8.1. *Actually, a careful analysis of Behrend's construction above shows that* $\nu(N) \gg \frac{1}{\log^{1/4} N} \cdot N^{1-\frac{2\sqrt{2\log 2}}{\sqrt{\log N}}}$. *Recently, this lower bound was improved by M. Elkin [4] who proved that*

$$\nu(N) \gg \log^{1/4} N \cdot N^{1-\frac{2\sqrt{2\log 2}}{\sqrt{\log N}}}.$$

For a brief exposition of M. Elkin estimate see the paper [10] of B. Green and J. Wolf.

Chapter 2

Green-Tao-Szemerédi theorem

2.1 Introduction

The main goal of this chapter is the discussion of some of the ideas behind the proof of Green-Tao theorem.

Roughly speaking, the proof has two main steps:

- firstly one generalizes Szemerédi theorem to the more general context of *pseudorandom measures*: this is the content of the Green-Tao-Szemerédi theorem (see the section 2.2 for further details);
- secondly, one shows the existence of pseudorandom measures on the set of primes (along the lines of the works of Goldston-Yıldırım).

Once these two facts are established, we will see that the Green-Tao theorem follows immediately (see the section 2.2).

However, before entering (in section 2.4) into the details of the previous outline, we will try to motivate the concepts and tactics of the proof of the Green-Tao-Szemerédi theorem via a complete proof of Roth theorem (which corresponds to the particular case $k = 3$ of Szemerédi theorem) in the section 2.3, while (unfortunately) we will skip completely the discussion of the results related to the work of Goldston-Yıldırım.

The organization of this chapter is:

- In section 2.2 we will present in more details the scheme of the proof of Green-Tao theorem; in particular, we will state precisely the Green-Tao-Szemerédi theorem and Goldston-Yıldırım theorems; finally, we will prove Green-Tao theorem *assuming* these two results.

- In section 2.3 we give a proof of Roth theorem about the existence of infinitely many arithmetic progressions of length 3 (i.e., 3-AP) in subsets of positive density to serve as an inspiring model for the proof of the Green-Tao-Szemerédi theorem.

- Ending this chapter, in section 2.4, we will prove the Green-Tao-Szemerédi theorem.

Closing this brief introduction, let us remark that when we reach the end of this chapter, the Green-Tao theorem will be proved *except* for the results of Goldston and Yıldırım whose beautiful proof will be omitted because it is purely number-theoretical (and, thus, outside the scope of this "ergodic-theoretical" book).

2.2 Strategy of the proof of the Green-Tao theorem

During this chapter, we will fix $k \geq 3$ the length of the arithmetic progression (AP) of primes (we are searching for) and we take $N := |\mathbb{Z}_N|$ a (large) prime number, so that the elements $1, \dots, N-1$ can be inverted in $\mathbb{Z}_N$. We write $o(1)$ to denote any quantity converging to zero when $N \to \infty$ and $O(1)$ to denote any quantity staying bounded when $N \to \infty$. At certain moments of the text, the $o(1)$ (resp., $O(1)$) quantities will converge to zero (resp., stay bounded) *depending on certain extra parameters* (e.g., j, ε). In this case, we will put these extra parameters as subscript indices (e.g., $o_{j,\varepsilon}(1)$). Moreover, we will abreviate any quantity of the form $O(1)X$ (resp., $o(1)X$) as $O(X)$ (resp., $o(X)$).

Keeping these notations in mind, let us put Green-Tao theorem in an appropriate context. Recall that this theorem says that for any $k \geq 3$, there are infinitely many k-APs (i.e., arithmetic progressions of length k) formed (only) by prime numbers. Moreover, we know that Szemerédi theorem ensures the existence of several k-APs in subsets of integers with *positive density*. However, we already know that Szemerédi theorem *doesn't* imply Green-Tao theorem because the set of prime numbers has density zero. Nevertheless, the idea of Green and Tao is:

- although we can't apply *directly* Szemerédi theorem, we can *modify* it to work with certain subsets with a *weakly random* (additive) behavior[1] (or, more precisely, *pseudorandom*); this result will be called *Green-Tao-Szemerédi* in this book;

[1] Of course, the point here consists into the choice of a nice definition of pseudorandomness in order to include the set of prime numbers.

- this reduces our task to show that the set of prime numbers has a pseudorandom behaviour; this fact follows more or less directly from the works of Goldston and Yıldırım.

In the sequel, we give the details for these items. To do so, we start with the definition of *pseudorandomness*:

Definition 2.2.1. - *We say that* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *is a measure if* $\mathbb{E}(\nu) = 1 + o(1)$.

- *A measure* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *satisfies the* (m_0, t_0, L_0)-linear forms condition *if, for any family of* $m \leq m_0$ *linear forms* $\psi_i : \mathbb{Z}_N^t \to \mathbb{Z}_N$, $t \leq t_0$, *say* $\psi_i(x) = \sum L_{ij} x_j + b_i$, *where* L_{ij} *are rational numbers with numerators and denominators* $\leq L_0$ *(in absolute value), no t-tuple* $(L_{ij})_{1 \leq j \leq t}$ *is a rational multiple of any other and for arbitrary* $b_i \in \mathbb{Z}$, *then:*

$$\mathbb{E}(\nu(\psi_i(x)) \cdots \nu(\psi_m(x)) | x \in \mathbb{Z}_N^t) = 1 + o_{m_0, t_0, L_0}(1).$$

- *A measure* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *satisfies the* m_0-correlation condition *if, for every* $m \leq m_0$, *there are weights* $\tau_m : \mathbb{Z}_N \to \mathbb{R}^+$ *such that* $\mathbb{E}(\tau_m^q) = O_{m,q}(1)$ *(moment conditions) for all* $1 \leq q < \infty$ *and*

$$\mathbb{E}(\nu(x + h_1) \cdots \nu(x + h_m) | x \in \mathbb{Z}) \leq \sum_{i < j \leq m} \tau_m(h_i - h_j),$$

for any $h_i \in \mathbb{Z}_N$ *(not necessarily distinct).*

- *A measure* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *is* k-pseudorandom *if* $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ *satisfy the* $(k \cdot 2^{k-1}, 3k - 4, k)$ *linear forms condition and the* 2^{k-1} *correlation condition.*

This definition may be strange at first, but it is essentially based on the works of Goldston-Yıldırım, where one studies majorants for modified versions of the von Mangoldt function (which, as we saw, is intimately related to the prime numbers). Intuitively, the two conditions above mean that the set of integers supporting ν has weakly random (additive) arithmetic properties. The main advantage of these conditions is the fact that they are only slightly weaker than the usual randomness condition, so that Szemerédi theorem can be generalized (see Theorem 2.2.1 below), although they are sufficiently weak to include the case of the prime numbers (even though these two facts are very far from trivial).

Once we have a precise definition of pseudorandomness, we can state one of the main results of the work of Green-Tao [9, Theorem 3.5]:

Theorem 2.2.1 (Green-Tao-Szemerédi)**.** *Let $k \geq 3$ and $0 < \delta \leq 1$. Then, for any k-pseudorandom measure ν, it holds that, for every $f : \mathbb{Z}_N \to \mathbb{R}^+$ such that $0 \leq f(n) \leq \nu(n)$ and $\mathbb{E}(f) \geq \delta$, we have:*

$$\mathbb{E}(f(n)f(n+r)\cdots f(n+(k-1)r)|n, r \in \mathbb{Z}_N) \geq c(k,\delta) - o_{k,\delta}(1).$$

Remark 2.2.1. *Taking $\nu \equiv \nu_{const} \equiv 1$ in this statement, we recover Szemerédi theorem as a corollary. However, it turns out we are going to* use *Szemerédi theorem in the proof of this more general result, so that the proof of Green-Tao-Szemerédi theorem* doesn't *give a new proof of Szemerédi theorem.*

The proof of Theorem 2.2.1 is based on *Furstenberg* ideas (i.e., multiple recurrence theorem) and the so-called *Gowers norms*. For the moment, we postpone the proof of Theorem 2.2.1 to section 2.4.

Since we aim to apply Theorem 2.2.1 to conclude the Green-Tao theorem, let us see how one can construct pseudorandom measures related to the prime numbers. We recall the definition:

Definition 2.2.2. *The* von Mangoldt function *is*

$$\Lambda(n) := \begin{cases} \log p \text{ if } n = p^m \\ 0 \text{ otherwise} . \end{cases}$$

We remember that this function is (essentially) supported on the prime numbers (since the contribution of the powers of prime numbers is very small), so that we can say that this function works as a "characteristic function" of the primes. In terms of this function, we know that the *prime number theorem* can be reformulated as $\mathbb{E}(\Lambda(n)) = 1 + o(1)$ (see chapter 1). In order to be able to use Green-Tao-Szemerédi theorem in the context of prime numbers, we need to find a k-pseudorandom measure ν such that $\nu(n) \geq c(k)\Lambda(n)$. However, it is known that such measures *don't exist*![2]

In order to overcome this technical difficulty, Green and Tao use the "*W-trick*". Let $w = w(N) \to \infty$ be a parameter growing very slowly with N and let $W = \Pi_{p \leq w(N);\ \text{p prime}} p$. The *modified* von Mangoldt function is:

$$\widetilde{\Lambda}(n) = \begin{cases} \frac{\phi(W)}{W} \log(Wn+1) \text{ if } Wn+1 \text{ is prime} \\ 0 \text{ otherwise} . \end{cases}$$

Now we have a function still seeing (some) prime numbers such that it doesn't see neither the powers of primes nor the annoying *non-uniformity*

[2] Basically this occurs because the prime numbers (and the von Mangoldt function) are concentrated, for any $q > 1$ integer, on the $\phi(q)$ residual classes $a(\text{mod } q)$ with $(a, q) = 1$ (where $\phi(q)$ is the Euler totient function), while any pseudorandom measure must be equidistributed along *all* congruence classes modulo q; since the quocient $\phi(q)/q$ can be made arbitrarily small, we see that the von Mangoldt function doesn't have pseudorandom majorants.

due to the presence of products of small primes. If $w(n)$ has a sufficiently slow growth [3], say $w(n) << \log\log n$, we see that Dirichlet theorem implies:

$$\frac{1}{N}\sum_{n\leq N}\widetilde{\Lambda}(n) = 1 + o(1).$$

In other words, $\widetilde{\Lambda}$ is a measure. In this situation, the second key result of the work of Green and Tao [9, Proposition 9.1] (based on the results of Goldston and Yıldırım) is:

Theorem 2.2.2. *If $\epsilon_k = 1/(k+4)!2^k$ and N is a very large prime number, then there exists ν a k-pseudorandom measure such that $\nu(n) \geq 2^{-k-5}k^{-1}\widetilde{\Lambda}(n)$ for $\epsilon_k N \leq n \leq 2\epsilon_k N$.*

Following a long-standing tradition in (analytical) Number Theory, the proof of Theorem 2.2.2 (as any important fact about prime numbers) uses zero-free regions of the so-called Riemann zeta function (besides other machineries). In particular, the nature of the argument is purely number-theoretical. Because we want to keep the coherence between the title and the content of this book, we will completely skip the proof of this beautiful result (referring the curious reader to the original paper of Green and Tao) and we will admit its validity during the rest of our discussion.

Finally, assuming momentarily the two key theorems 2.2.1 and 2.2.2, we can show the Green-Tao theorem.

2.2.1 Proof of Green and Tao theorem

Suppose that the theorems 2.2.1 and 2.2.2 are valid.

If

$$f(n) = \frac{1}{k2^{k+5}}\widetilde{\Lambda}(n)1_{[\epsilon_k N, 2\epsilon_k N]}$$

then

$$\mathbb{E}(f) = \frac{1}{Nk2^{k+5}}\sum_{\epsilon_k N\leq n\leq 2\epsilon_k N}\widetilde{\Lambda}(n) = \frac{1}{k2^{k+5}}\epsilon_k(1+o(1)).$$

Since the theorem 2.2.2 guarantees the existence of a k-pseudorandom majorant of $2^{-k-5}k^{-1}\widetilde{\Lambda}$ in $[\epsilon_k N, 2\epsilon_k N]$, we can apply the Green-Tao-Szemerédi theorem 2.2.1 to conclude that:

$$\mathbb{E}(f(n)\cdots f(n+(k-1)r)|n, r\in\mathbb{Z}_N) \geq c(k, k^{-1}2^{-k-3}\epsilon_k) - o(1).$$

Because the case $r = 0$ contributes with a factor of $O(\frac{1}{N}\log^k N) = o(1)$, we obtain the existence of a k-AP of primes in $\mathbb{Z}_N$ (for a large N). Moreover, since $\epsilon_k < 1/k$ and $k \geq 3$, we see that this k-AP is a legitimate k-AP in $\mathbb{Z}$ (in the sense that it doesn't wrap around in $\mathbb{Z}_N$).

[3] Although we ask for slow growth of $w(n)$, it is possible to check (at the end of the proof of Green-Tao theorem) that it suffices to take $w(n)$ a very large *constant* depending only on k (but not on N).

2.2.2 Some comments

Once we reduced the Green-Tao theorem to the two key theorems 2.2.1 and 2.2.2, we will dedicate the rest of this chapter to the proof of the ergodic part of the argument, namely, the Green-Tao-Szemerédi theorem 2.2.1 (while we skip the number-theoretical part corresponding to the theorem 2.2.2).

However, in order to illustrate the ideas behind the Green-Tao-Szemerédi theorem (which can be technical and hard in a first reading), we will present a proof of Roth theorem (i.e., Szemerédi theorem in the particular case $k = 3$) and, in the sequel, we prove the Green-Tao-Szemerédi theorem (closing the chapter).

2.3 Proof of Roth theorem

Define $\Lambda_3(f, g, h) = \mathbb{E}(f(n)g(n+r)h(n+2r)|n, r \in \mathbb{Z}_N)$. In this language, Roth theorem can be reformulated as:

Theorem 2.3.1. *For all non-negative* $f : \mathbb{Z}_N \to \mathbb{R}$ *with*

$$0 < \delta \leq \|f\|_{L^1(\mathbb{Z}_N)} \leq \|f\|_{L^\infty(\mathbb{Z}_N)} \leq 1$$

it holds

$$\Lambda_3(f, f, f) \geq c(3, \delta) - o_\delta(1).$$

In other words, we want some *lower bounds* on $\Lambda_3(f, f, f)$. We begin with the simple remark that it is fairly easy to get *upper bounds*: e.g., by Young inequality,

$$|\Lambda_3(f, g, h)| \leq \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r},$$

if $1 \leq p, q, r \leq \infty$ and $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} \leq 2$.

On the other hand, we are only interested in *lower* bounds for Λ_3 and, *a priori*, upper bounds are not useful for our interests. However, we can decompose f into a "good" part $g = \mathbb{E}(f)$ and a"bad" part $b = f - \mathbb{E}(f)$. Using the multilinearity of Λ_3, we can use this decomposition to split up $\Lambda_3(f, f, f)$ into eight pieces:

$$\Lambda_3(f, f, f) = \Lambda_3(g, g, g) + \cdots + \Lambda_3(b, b, b).$$

By hypothesis, $\mathbb{E}(f) \geq \delta$, so that the first term is $\Lambda_3(g, g, g) \geq \delta^3$. Thus, some good upper estimates of the remaining terms (e.g., if the sum of the seven remaining terms is smaller than δ^3) would lead us to a proof of Roth theorem.

Unfortunately, the previous upper bound (via Young inequality) isn't sufficient to conclude the argument, unless δ is close to 1, say $\delta > 2/3$

(a non-realistic case). But, we can refine this upper bound argument via Harmonic Analysis, or, more precisely the *Fourier transform*:

$$\widehat{f}(\xi) := \mathbb{E}(f(x)e_N(-x\xi) : x \in \mathbb{Z}_N),$$

where $e_N(x) := \exp(2\pi i x/N)$. From the inversion formula

$$f(x) = \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi) e_N(x\xi)$$

we obtain that

$$\Lambda_3(f,g,h) = \sum_{\xi_1,\xi_2,\xi_3} \widehat{f}(\xi_1)\widehat{g}(\xi_2)\widehat{h}(\xi_3)\times \\ \mathbb{E}(e_N(n\xi_1 + (n+r)\xi_2 + (n+2r)\xi_3) : n, r \in \mathbb{Z}_N).$$

Note that the expectations on the right-hand side are 1 if $\xi_1 = \xi_3$ and $\xi_2 = -2\xi_1$, and 0 otherwise. In particular,

$$\Lambda_3(f,g,h) = \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi)\widehat{g}(-2\xi)\widehat{h}(\xi).$$

From Plancherel formula $\|f\|_{L^2(\mathbb{Z}_N)} = \|\widehat{f}\|_{l^2(\mathbb{Z}_N)}$ and Hölder inequality, it follows that

$$|\Lambda_3(f,g,h)| \leq \|f\|_{L^2(\mathbb{Z}_N)}\|\widehat{g}\|_{l^4(\mathbb{Z}_N)}\|\widehat{h}\|_{l^4(\mathbb{Z}_N)}. \tag{2.1}$$

Using this estimate, we can prove:

Proposition 2.3.1. *Fix f and consider a decomposition $f = g + b$ where*

$$\|g\|_{L^\infty(\mathbb{Z}_N)}, \|b\|_{L^\infty(\mathbb{Z}_N)} = O(1)$$

and

$$\|g\|_{L^1(\mathbb{Z}_N)}, \|b\|_{L^1(\mathbb{Z}_N)} = O(\delta).$$

Then

$$\Lambda_3(f,f,f) = \Lambda_3(g,g,g) + O(\delta^{5/4}\|\widehat{b}\|_{l^4(\mathbb{Z}_N)})$$

and

$$\Lambda_3(f,f,f) = \Lambda_3(g,g,g) + O(\delta\|\widehat{b}\|_{l^\infty(\mathbb{Z}_N)}).$$

Proof. By hypothesis, $\|g\|_{L^2(\mathbb{Z}_N)}, \|b\|_{L^2(\mathbb{Z}_N)} = O(\delta^{1/2})$, so that Plancherel theorem says that

$$\|\widehat{g}\|_{l^2(\mathbb{Z}_N)}, \|\widehat{b}\|_{l^2(\mathbb{Z}_N)} = O(\delta^{1/2}).$$

Furthermore, the L^1 bounds on g and b imply

$$\|\widehat{g}\|_{l^\infty(\mathbb{Z}_N)}, \|\widehat{b}\|_{l^\infty(\mathbb{Z}_N)} = O(\delta).$$

Thus, by Hölder inequality,

$$\|\widehat{g}\|_{l^4(\mathbb{Z}_N)}, \|\widehat{b}\|_{l^4(\mathbb{Z}_N)} = O(\delta^{3/4}).$$

The proposition follows by decomposing $\Lambda_3(f,f,f)$ into eight pieces and using (2.1). □

This proposition suggests that a possible strategy to get non-trivial lower bounds on $\Lambda_3(f,f,f)$ passes by a decomposition $f = g + b$ into a good function in the sense that $\Lambda_3(g,g,g)$ is "large" and a bad function b in the sense that the l^4 norm of its Fourier transform is small.

The attentive reader noticed that we already indicated one simple possibility for the decomposition: $g = \mathbb{E}(f)$ and $b = f - \mathbb{E}(f)$. Observe that we have the relatively nice estimate $\Lambda_3(g,g,g) \geq \delta^3$, but we don't dispose of any good estimate for $\|\widehat{b}\|_{l^4(\mathbb{Z}_N)}$; our best estimates at the present moment are $O(\delta^{3/4})$, which is essentially useless since it allows the "error" term to dominate the "main" term.[4]

However, we can eliminate the case of a *linearly uniform* $b = f - \mathbb{E}(f)$, i.e., $\|\widehat{b}\|_{l^\infty} \leq \delta^2/100$. Of course, the problem is: what to do when b isn't linearly uniform? In this situation, we adopt the so-called *energy increment argument*, i.e.,

- **Energy increment argument**: once b isn't uniform, we replace g by another function whose L^2 norm is strictly greater than $\|g\|_{L^2}$. After repeating this process a *finite* number of times, we expect to reach an uniform function b (since the energy is finite).

Logically, one should work in details this idea in order to see that it leads to a proof of Roth theorem. In this direction, let us introduce the definition:

Definition 2.3.1. *Given K a positive integer, we call any function $f : \mathbb{Z}_N \to \mathbb{C}$ of the form*

$$f(x) = \sum_{j=1}^{K} c_j \exp(2\pi i \xi_j x/N),$$

where $|c_j| \leq 1$ and $\xi_j \in \mathbb{Z}_N$, a K-quasi-periodic function. *Furthermore, for a fixed $\sigma > 0$, we say that a function f is* (σ, K)-quasi-periodic *if $\|f - f_{qp}\|_{L^2(\mathbb{Z}_N)} \leq \sigma$ for some K-quasi-periodic function f_{qp}.*

[4] Indeed, $f = \chi_{[1,\delta N]}$ verifies $\Lambda_3(f,f,f) \sim \delta^2$ and $\Lambda_3(g,g,g) = \delta^3$, for instance.

Remark 2.3.1. *Notice that the product fg is $(2\sigma, K^2)$-quasi-periodic whenever f and g are (σ, K)-quasi-periodic functions.*

An important point around the concept of quasi-periodic functions f is the fact that it is possible to prove good lower bounds on $\Lambda_3(f,f,f)$ in this context:

Lemma 2.3.1 ("Multiple Recurrence" of quasi-periodic functions)**.** *Given $0 < \delta < 1$, $M \geq 1$, $0 < \sigma \leq \delta^3/100M$ and $0 \leq f \leq M$ a non-negative bounded (σ, K)-quasi-periodic function with $\mathbb{E}(f) \geq \delta$, it holds*

$$\Lambda_3(f,f,f) \geq c(K,M,\delta) - o_{K,M,\delta}(1),$$

for some $c(K,M,\delta) > 0$.

Proof. Let $f_{qp}(x) = \sum_{j=1}^{K} c_j \exp(2\pi i x \xi_j / N)$ be a K-quasi-periodic function close to f and take $\varepsilon = \varepsilon(K,\delta) > 0$ a small constant. By Dirichlet simultaneous approximation theorem (which is a consequence of Dirichlet's pigeonhole principle), we have

$$\mathbb{E}(\|r\xi_j\| \leq \varepsilon \text{ for every } j\,;\, r \in \mathbb{Z}_N) \geq c(\varepsilon, K). \tag{2.2}$$

Here, it is fundamental to stress out that the constant $c(\varepsilon, K) > 0$ *doesn't depend* on N. On the other hand, by considering the shift dynamics $T(x) := x + 1$ and by fixing r such that $\|r\xi_j\| \leq \varepsilon$ (where $1 \leq j \leq K$), we get:

$$\|f_{qp} \circ T^r - f_{qp}\|_{L^2(\mathbb{Z}_N)} \leq C(K)\varepsilon.$$

Combining this information with the triangular inequality, it follows:

$$\|f \circ T^r - f\|_{L^2(\mathbb{Z}_N)} \leq \delta^3/50M + C(K)\varepsilon.$$

Applying T^r again to this estimate, we also obtain

$$\|f \circ T^{2r} - f \circ T^r\|_{L^2(\mathbb{Z}_N)} \leq \delta^3/50M + C(K)\varepsilon.$$

Since f is bounded, from these estimates, we can infer:

$$\|f \cdot (f \circ T^r) \cdot (f \circ T^{2r}) - f^3\|_{L^1(\mathbb{Z}_N)} \leq \delta^3/2 + C(K)M\varepsilon.$$

However, because $f \geq 0$, using our hypothesis $\mathbb{E}(f) \geq \delta$ and Hölder inequality, we obtain

$$\|f^3\|_{L^1(\mathbb{Z}_N)} \geq \|f\|^3_{L^1(\mathbb{Z}_N)} \geq \delta^3.$$

Thus, $\mathbb{E}(f \cdot (f \circ T^r) \cdot (f \circ T^{2r})) \geq \delta^3/2 - C(K)M\varepsilon$. Choosing $\varepsilon > 0$ small depending on δ, K and M, it follows that

$$\mathbb{E}(f \cdot (f \circ T^r) \cdot (f \circ T^{2r})) \geq \delta^3/4.$$

Because $f \geq 0$, taking the average on r and using (2.2), we conclude

$$\mathbb{E}(f(n) \cdot (f \circ T^r)(n) \cdot (f \circ T^{2r})(n) \,|\, n, r \in \mathbb{Z}_N) \geq \delta^3 c(\varepsilon, K)/4.$$

Since $\Lambda_3(f,f,f) = \mathbb{E}(f(n) \cdot (f \circ T^r)(n) \cdot (f \circ T^{2r})(n) \,|\, n, r \in \mathbb{Z}_N)$, the proof of the lemma is complete. □

Looking for possible applications of this lemma, we will try to write arbitrary functions as a sum of a quasi-periodic function and a linearly uniform function. Keeping this goal in mind, we construct a class of *sigma-algebras* such that its measurable functions are always quasi-periodic functions:

Lemma 2.3.2. *Take $0 < \varepsilon \ll 1$ and let χ be a function of the form $\chi(x) := \exp(2\pi i x\xi/N)$. Then, there exists a sigma-algebra $\mathcal{B}_{\varepsilon,\chi}$ such that $\|\chi - \mathbb{E}(\chi|\mathcal{B}_{\varepsilon,\chi})\|_{L^\infty(\mathbb{Z}_N)} \leq C\varepsilon$ and, for every $\sigma > 0$, there is $K = K(\sigma, \varepsilon) > 0$ with the following property: every $\mathcal{B}_{\varepsilon,\chi}$-measurable function f satisfying the estimate $\|f\|_{L^\infty(\mathbb{Z}_N)} \leq 1$ is (σ, K)-quasi-periodic.*

Proof. We will use a random process to get the desired sigma-algebra: take α a complex number in the unit square and let $\mathcal{B}_{\varepsilon,\chi}$ be the sigma-algebra whose atoms have the form $\chi^{-1}(Q)$, where Q is a square such that the vertices of $Q - \varepsilon\alpha$ lies over the lattice $\varepsilon\mathbb{Z}^2$. Notice that this sigma-algebra has $O(1/\varepsilon)$ atoms and $\|\chi - \mathbb{E}(\chi|\mathcal{B}_{\varepsilon,\chi})\|_{L^\infty(\mathbb{Z}_N)} \leq C\varepsilon$. In particular, it remains only to verify the second part of the lemma. Observe that it suffices to show the desired fact for $\sigma = 2^{-n}$ (where $n \gg 1$) with a probability of $1 - O(\sigma)$ on α. Since $\mathcal{B}_{\varepsilon,\chi}$ possess $O(1/\varepsilon)$ atoms, we can restrict ourselves to the case f equals to the characteristic function of an atom A of $\mathcal{B}_{\varepsilon,\chi}$ so that our task is reduced to prove the desired property with a probability of $1 - O(c(\varepsilon)\sigma)$. Observe that, in this situation, we can rewrite f as $f(x) = 1_Q(\chi(x) - \varepsilon\alpha)$. Applying Weierstrass approximation theorem on the disc $|z| \leq O(1/\varepsilon)$, we can find a polynomial $P(z, \overline{z})$ with $C(\sigma, \varepsilon)$ terms whose coefficients are bounded by $C(\sigma, \varepsilon)$ such that $|P| \leq 1$ on the disc $|z| \leq O(1/\varepsilon)$ and $|1_Q(z) - P(z, \overline{z})| = O(c(\varepsilon)\sigma)$ for all z on this disc with the exception of a subset of measure $O(c(\varepsilon)^2\sigma^2)$. This implies that

$$\|1_Q(\chi(x) - \varepsilon\alpha) - P(\chi(x) - \varepsilon\alpha, \overline{\chi(x)} - \varepsilon\overline{\alpha})\|_{L^2(\mathbb{Z}_N)} \leq c(\varepsilon)\sigma$$

with a probability of $1 - O(c(\varepsilon)\sigma)$ on α. However $P(\chi(x) - \varepsilon\alpha, \overline{\chi(x)} - \varepsilon\overline{\alpha})$ is a linear combination of $C(\varepsilon, \sigma)$ functions of the form $\exp(2\pi i x\xi/N)$ where the coefficients are bounded by $C(\varepsilon)$. In other words, it follows that $P(\chi(x) - \varepsilon\alpha, \overline{\chi(x)} - \varepsilon\overline{\alpha})$ is a K-quasi-periodic function (where $K = C(\varepsilon, \sigma)$). In particular, f is a (σ, K)-quasi-periodic function. This concludes the proof of the lemma. □

A useful corollary of this lemma is:

Corollary 2.3.1. *Let $0 < \varepsilon_j \ll 1$ and $\chi_j(x) = \exp(2\pi i x\xi_j/N)$, where $j = 1, \dots, n$. Denote by $\mathcal{B}_{\varepsilon_j,\chi_j}$ the sigma-algebras provided by the previous lemma. Then, for all $\sigma > 0$, there is some $K = K(n, \sigma, \varepsilon_1, \dots, \varepsilon_n)$ such that any $\mathcal{B}_{\varepsilon_1,\chi_1} \vee \cdots \vee \mathcal{B}_{\varepsilon_n,\chi_n}$-measurable function f satisfying the estimate $\|f\|_{L^\infty(\mathbb{Z}_N)} \leq 1$ is (σ, K)-quasi-periodic.*

Proof. Because the number of atoms of the sigma-algebra $\mathcal{B}_{\varepsilon_1,\chi_1} \vee \cdots \vee \mathcal{B}_{\varepsilon_n,\chi_n}$ is $C(n, \varepsilon_1, \dots, \varepsilon_n)$, it suffices to show the corollary for the particular case of characteristic functions f of some atom of $\mathcal{B}_{\varepsilon_1,\chi_1} \vee \cdots \vee \mathcal{B}_{\varepsilon_n,\chi_n}$. But, in this setting, f is a product of n characteristic functions of certain atoms of the sigma-algebras $\mathcal{B}_{\varepsilon_j,\chi_j}$. Thus, the corollary follows from the previous lemma and Remark 2.3.1. □

Another interesting property of these sigma-algebras (besides the fact that they contain quasi-periodic functions) is their usefulness in the identification of *obstructions to linear uniformity*:

Lemma 2.3.3. *Let b be a bounded function with $\|\widehat{b}\|_{l^\infty(\mathbb{Z}_N)} \geq \sigma > 0$ and $0 < \varepsilon \ll \sigma$. Then, there exists a function of the form $\chi(x) = \exp(2\pi i x\xi/N)$ such that the associated sigma-algebra $\mathcal{B}_{\varepsilon,\chi}$ satisfies*

$$\|\mathbb{E}(b|\mathcal{B}_{\varepsilon,\chi})\|_{L^2(\mathbb{Z}_N)} \geq \sigma/2.$$

Proof. By definition, there exists a frequency ξ such that $|\widehat{b}(\xi)| \geq \sigma$, i.e.,

$$|\mathbb{E}(b(n)\exp(-2\pi i n\xi/N)|n \in \mathbb{Z}_N)| \geq \sigma.$$

Define $\chi(x) := \exp(2\pi i x\xi/N)$ and rewrite the previous inequality as

$$|\langle b, \chi\rangle_{L^2(\mathbb{Z}_N)}| \geq \sigma.$$

By the previous lemma, we know that there is a sigma-algebra $\mathcal{B}_{\varepsilon,\chi}$ such that

$$\|\chi - \mathbb{E}(\chi|\mathcal{B}_{\varepsilon,\chi})\|_{L^\infty(\mathbb{Z}_N)} \leq C\varepsilon.$$

On the other hand, since b is bounded and the conditional expectation is self-adjoint, we can combine these two estimates to conclude

$$\langle \mathbb{E}(b|\mathcal{B}_{\varepsilon,\chi}), \chi\rangle = \langle b, \mathbb{E}(\chi|\mathcal{B}_{\varepsilon,\chi})\rangle \geq \sigma - C\varepsilon.$$

This shows that $\|\mathbb{E}(b|\mathcal{B}_{\varepsilon,\chi})\|_{L^2(\mathbb{Z}_N)} \geq \sigma - C\varepsilon \geq \sigma/2$, so that the proof of the lemma is complete. □

The last ingredient in the proof of Roth theorem is the following *structure* proposition:

Proposition 2.3.2 ("quantitative Koopman-von Neumann theorem")**.** *Let $F : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ be an arbitrary function, $0 < \delta \leq 1$, f be a non-negative bounded function verifying $\mathbb{E}(f) \geq \delta$ and $\sigma := \delta^3/100$. Then, there are a constant $0 < K \leq C(\delta, F)$ and a decomposition $f = g + b$ such that g is non-negative, bounded, $\mathbb{E}(g) = \mathbb{E}(f)$, g is (σ, K)-quasi-periodic and b verifies*

$$\|\widehat{b}\|_{l^\infty(\mathbb{Z}_N)} \leq F(\delta, K). \tag{2.3}$$

Proof. We will use the *energy increment argument* during the construction of g and b. For this argument, we need two sigma-algebras $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ which are always of the form $\mathcal{B}_{\varepsilon_1,\chi_1} \vee \dots \vee \mathcal{B}_{\varepsilon_n,\chi_n}$ during the entire argument. Moreover, we will need some estimatives of the form

$$\|\mathbb{E}(f|\widetilde{\mathcal{B}})\|^2_{L^2(\mathbb{Z}_N)} \leq \|\mathbb{E}(f|\mathcal{B})\|^2_{L^2(\mathbb{Z}_N)} + \sigma^2/4. \tag{2.4}$$

Observe that, by Pythagoras theorem, this estimate is equivalent to

$$\|\mathbb{E}(f|\widetilde{\mathcal{B}}) - \mathbb{E}(f|\mathcal{B})\|^2_{L^2(\mathbb{Z}_N)} \leq \sigma/2.$$

For the proof of this proposition, we will employ the following algorithm:

- *Stage 0*: We start with $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ equal to the trivial sigma-algebra $\{0, \mathbb{Z}_N\}$. Note that the inequality (2.4) is automatically satisfied at this stage.
- *Stage 1*: Consider $\mathcal{B}$ a sigma-algebra of the form $\mathcal{B}_{\varepsilon_1,\chi_1} \vee \dots \vee \mathcal{B}_{\varepsilon_n,\chi_n}$, where $\chi_j(x) = \exp(2\pi i x \xi_j/N)$. Since the function $\mathbb{E}(f|\mathcal{B})$ is bounded and $\mathcal{B}$-measurable, the corollary 2.3.1 says that one can find K depending on $\delta, n, \varepsilon_1, \dots, \varepsilon_n$ such that $\mathbb{E}(f|\mathcal{B})$ is $(\sigma/2, K)$-quasi-periodic.
- *Stage 2*: Put $g = \mathbb{E}(f|\widetilde{\mathcal{B}})$ and $b = f - \mathbb{E}(f|\widetilde{\mathcal{B}})$. If $\|\widehat{b}\|_{l^\infty(\mathbb{Z}_N)} \leq F(\delta, K)$, we end the algorithm. Otherwise, we go to the third stage.
- *Stage 3*: Since we didn't end the algorithm at the stage 2, we have $\|\widehat{b}\|_{l^\infty} > F(\delta, K)$. By Lemma 2.3.3, we can find $\varepsilon \ll F(\delta, K)$ and a function χ of the form $\chi(x) = \exp(2\pi i x \xi/N)$ whose associated sigma-algebra $\mathcal{B}_{\varepsilon,\chi}$ verifies

  $$\|\mathbb{E}(b|\mathcal{B}_{\varepsilon,\chi})\|_{L^2(\mathbb{Z}_N)} \geq F(\delta, K)/2.$$

  From the identity

  $$\mathbb{E}(b|\mathcal{B}_{\varepsilon,\chi}) = \mathbb{E}(\mathbb{E}(f|\widetilde{\mathcal{B}} \vee \mathcal{B}_{\varepsilon,\chi}) - \mathbb{E}(f|\widetilde{\mathcal{B}})|\mathcal{B}_{\varepsilon,\chi})$$

  and Pythagoras theorem, we obtain also that

  $$\|\mathbb{E}(f|\widetilde{\mathcal{B}} \vee \mathcal{B}_{\varepsilon,\chi}) - \mathbb{E}(f|\widetilde{\mathcal{B}})\|_{L^2(\mathbb{Z}_N)} \geq F(\delta, K)/2.$$

Applying Pythagoras theorem again, we get an *energy increment estimate*:

$$\|\mathbb{E}(f|\widetilde{\mathcal{B}} \vee \mathcal{B}_{\varepsilon,\chi})\|^2_{L^2(\mathbb{Z}_N)} \geq \|\mathbb{E}(f|\widetilde{\mathcal{B}})\|^2_{L^2(\mathbb{Z}_N)} - F(\delta, K)^2/4.$$

- *Stage 4*: We replace $\widetilde{\mathcal{B}}$ by $\widetilde{\mathcal{B}} \vee \mathcal{B}_{\varepsilon,\chi}$. If one still don't have the estimate (2.4), we return to the stage 2; otherwise, we replace $\mathcal{B}$ by $\widetilde{\mathcal{B}}$ and we go to the stage 1.

We claim that this algorithm stops. In fact, for a fixed $\mathcal{B}$ (and consequently K), each time we pass by the stage 4, the *energy* $\|\mathbb{E}(f|\widetilde{\mathcal{B}})\|^2_{L^2(\mathbb{Z}_N)}$ increases by a factor of $F(\delta,K)^2/4$. Thus, either the algorithm ends or the estimate (2.4) is violated in $C(\delta,K,F) = C\sigma^2/F(\delta,K)^2$ steps. In the second case, we replace $\mathcal{B}$ by the sigma-algebra associated to the $C(\delta,K,F)$ functions χ and parameters $\varepsilon \geq C(\delta,F,K)^{-1}$ appearing in this process. This implies that the quantity K associated to this new sigma-algebra $\mathcal{B}$ will be changed by a quantity of the form $C(\delta,K,F)$ and the energy $\|\mathbb{E}(f|\mathcal{B})\|^2_{L^2(\mathbb{Z}_N)}$ will increase by (at least) $\sigma^2/4$ due to the violation of (2.4). On the other hand, the fact that f is bounded ensures that the energy can't be bigger than $O(1)$. Therefore, these replacements of $\mathcal{B}$ just described can only be perfomed at most $O(\sigma^{-2})$ times. Putting these informations together, we see that the whole algorithm ends in $C(\delta,F)$ steps (and the quantity K is never bigger than $C(\delta,F)$ during the entire process). This concludes the proof of the proposition. □

Finally, let us complete the proof of Roth theorem: let $F : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ be a function to be chosen in a few moments and let us apply the previous propostion to get the corresponding decomposition $f = g + b$. By the lemma 2.3.1, we know that

$$\Lambda_3(g,g,g) \geq c(\delta,K) - o_{\delta,K}(1).$$

Combining this inequality with (2.3) and the proposition 2.3.1, we have

$$\Lambda_3(f,f,f) \geq c(\delta,K) + O(\delta \cdot F(\delta,K)) - o_{\delta,K}(1).$$

Taking F "sufficiently small", we can absorb the second term of the right-hand side via the first term, so that

$$\Lambda_3(g,g,g) \geq c(\delta,K)/2 - o_{\delta,K}(1).$$

Since $K \leq C(\delta,F) = C(\delta)$, Roth theorem is proved.

Closing this section, let us review below the two main steps of the proof of Roth theorem (which are going to inspire our discussion of the Green-Tao-Szemerédi theorem):

- **1st step**: to define a class of norms (*Gowers norms* $\|.\|_{U^{k-1}}$) in order to control the expectation of a k-AP to reside in the support of f; observe that, by proposition 2.3.1, in the particular case $k=3$, the l^4 norm of the Fourier transform is a good candidate;[5]
- **2nd step**: to make an energy increment argument.

2.4 Proof of Green-Tao-Szemerédi theorem

In this last subsection of the present chapter, we will prove (along several subsections) the key results allowing us to formalize the previous ideas. However, since they are somewhat technical, the reader may get lost during the discussion. In view of this possibility, we included at the end of each subsection a "resume" of the main results demonstrated and how they are connected to the big picture of the energy increment strategy.

2.4.1 Gowers norms

Let $\{0,1\}^d$ be the d-dimensional discrete cube, and $w=(w_1,\dots,w_d)\in\{0,1\}^d$. If $h\in\mathbb{Z}_N^d$, then $w.h := w_1.h_1+\dots w_d.h_d$. If $\{f_w\}_{w\in\{0,1\}^d}$, then the *Gowers inner product* is:

$$\langle (f_w)\rangle_{U^d} := \mathbb{E}(\Pi_w f_w(n+w.h)|n\in\mathbb{Z}_N, h\in\mathbb{Z}_N^d).$$

Firstly we remark that if $f_w=f$ for every w then $\langle (f_w)\rangle_{U^d}\geq 0$. Thus, we can define the *Gowers norms* (using $f_w=f$):

$$\|f\|_{U^d} := \langle (f)\rangle_{U^d}^{\frac{1}{2^d}}.$$

A basilar tool for the analysis of the Gowers norms is the *Gowers-Cauchy-Schwarz inequality*:

$$|\langle (f_w)\rangle_{U^d}| \leq \Pi_w \|f_w\|_{U^d}.$$

The proof of this inequality follows from the fact that, when f_w doesn't depend on w_d, it holds

$$\langle (f_w)\rangle_{U^d} = \mathbb{E}(\mathbb{E}(\prod_{w'\in\{0,1\}^{d-1}} f_{w',0}(y+w'.h') : y\in\mathbb{Z}_N)\times$$
$$\mathbb{E}(\prod_{w'\in\{0,1\}^{d-1}} f_{w',1}(y+w'.h') : y\in\mathbb{Z}_N | h'\in(\mathbb{Z}_N)^{d-1})).$$

Therefore, by the Cauchy-Schwarz inequality, we have

$$|\langle (f_w)\rangle_{U^d}| \leq \langle (f_{w',0})\rangle_{U^d}^{1/2}\langle (f_{w',1})\rangle_{U^d}^{1/2}.$$

[5] Indeed, in the case $k=3$, the Gowers norm $\|.\|_{U^2}$ is the l^4 norm of the Fourier transform; see the remark 2.4.1 of the next subsection.

Since we can exchange w_d by any other digit, applying this estimate d times, we obtain the Gowers-Cauchy-Schwarz inequality.

Furthermore, the binomial formula and the multilinearity of the inner product lead us to the *Gowers triangular inequality*:

$$\|f+g\|_{U^d} \leq \|f\|_{U^d} + \|g\|_{U^d}.$$

Finally, we have the following monotonicity relation:

$$\|f\|_{U^{d-1}} \leq \|f\|_{U^d},$$

which is a direct consequence of the Gowers-Cauchy-Schwarz inequality applied to the case $f_w := 1$ when $w_d = 1$ and $f_w := f$ when $w_d = 0$.

Remark 2.4.1. *Since the norms $\|.\|_{U^d}$ are homogenous, this shows that $\|.\|_{U^d}$ are* semi-norms*. However, $\|.\|_{U^1}$ isn't a norm because $\|f\|_{U^1} = \mathbb{E}(f)$. However, one can prove (by direct calculation):*

$$\|f\|_{U^2} = (\sum \widehat{f}(\xi)^4)^{\frac{1}{4}},$$

where $\widehat{f}(\xi) = \mathbb{E}(f(x)e^{-2\pi i\xi/N}; x \in \mathbb{Z}_N)$ and the inversion formula $f(x) = \sum \widehat{f}(\xi)e^{2\pi ix\xi/N}$ holds. Consequently, the Gowers norms are genuine norms for $d \geq 2$.

Using this notation, the natural generalization of proposition 2.3.1 is:

Theorem 2.4.1 (generalized von Neumann theorem)**.** *If ν is a k-pseudorandom measure and $f_0, \ldots, f_{k-1} \in L^1(\mathbb{Z}_N)$ are some functions such that $|f_j(x)| \leq 1 + \nu(x)$, then, if $c_0, \ldots, c_{k-1} \in \mathbb{Z}_N$ are* distinct*, we have:*

$$\mathbb{E}(\Pi_j f_j(n + c_j r)|n, r \in \mathbb{Z}_N) = O(\inf \|f_j\|_{U^{k-1}}) + o(1).$$

Proof. We begin with some preliminary reductions: up to replacing ν by $(\nu+1)/2$, rearranging f_j, c_j and translating x by $c_0 r$, we can assume that

$$|f_j(x)| \leq \nu(x),\ \forall\ x \in \mathbb{Z}_N, j = 0, \ldots, k-1,$$

$$\inf_{0 \leq j \leq k-1} \|f_j\|_{U^{k-1}} = \|f_0\|_{U^{k-1}}$$

and

$$c_0 = 0.$$

This reduces our problem to prove that

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r)|x, r \in \mathbb{Z}_N\right) = O(\|f_0\|_{U^{k-1}}) + o(1).$$

We divide the proof of this inequality into two parts: in the first part we prove a Cauchy-Schwarz type inequality and we apply this inequality $k-1$ times to the left-hand side of the previous identity in order to get a control of $\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x+c_jr)|x,r\in\mathbb{Z}_N\right)$ via a *weighted sum* of f_0 over $(k-1)$-dimensional cubes; in the second part we show that the linear forms conditions implies that these weights are equal to 1 in average (so that the theorem follows).

In order to state the Cauchy-Schwarz type inequality in a reasonable way, we introduce a little bit more of notation. Given $0\le d\le k-1$, two vectors $y=(y_1,\dots,y_{k-1})\in(\mathbb{Z}_N)^{k-1}$ and $y=(y'_{k-d},\dots,y'_{k-1})\in(\mathbb{Z}_N)^d$, and a subset $S\subset\{k-d,\dots,k-1\}$, we define the vector $y^{(S)}=(y_1^{(S)},\dots,y_{k-1}^{(S)})\in(\mathbb{Z}_N)^{k-1}$ by

$$y_i^{(S)}:=\begin{cases} y_i & \text{if } i\notin S\\ y'_i & \text{if } i\in S.\end{cases}$$

In other words, S indicates the components of $y^{(S)}$ coming from y' instead of y.

Lemma 2.4.1. *Let $\nu:\mathbb{Z}_N\to\mathbb{R}^+$ be a measure and $\phi_0,\dots,\phi_{k-1}:(\mathbb{Z}_N)^{k-1}\to\mathbb{Z}_N$ some functions of the $(k-1)$ variables y_i such that ϕ_i doesn't depend on y_i for each $1\le i\le k-1$. Suppose that $f_0,\dots,f_{k-1}\in L^1(\mathbb{Z}_N)$ are some functions satisfying $|f_i(x)|\le\nu(x)$ for every $x\in\mathbb{Z}_N$ and $0\le i\le k-1$. For each $0\le d\le k-1$ and $1\le i\le k-1$, define*

$$\begin{aligned} J_d :=\mathbb{E}\Bigg(\prod_{S\subset\{k-d,\dots,k-1\}} &(\prod_{i=0}^{k-d-1} f_i(\phi_i(y^{(S)}))\times\\ &\times\prod_{i=k-d}^{k-1}\nu^{1/2}(\phi_i(y^{(S)}))\Big|y\in(\mathbb{Z}_N)^{k-1},y'\in(\mathbb{Z}_N)^d\Bigg),\end{aligned}$$

and

$$P_d:=\mathbb{E}\left(\prod_{S\subset\{k-d,\dots,k-1\}}\nu(\phi_{k-d-1}(y^{(S)}))|y\in(\mathbb{Z}_N)^{k-1},y'\in(\mathbb{Z}_N)^d\right).$$

Then, for all $0\le d\le k-2$, we have the inequality

$$|J_d|^2\le P_dJ_{d+1}.$$

Proof of Lemma 2.4.1. Consider J_d. Since ϕ_{k-d-1} doesn't depend on y_{k-d-1}, we can extract the quantities depending on ϕ_{k-d-1} from the average on y_{k-d-1}, so that we can write

$$\begin{aligned} J_d=\mathbb{E}(G(y,y')H(y,y')|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},&\\ y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N),&\end{aligned}$$

where

$$G(y,y') := \prod_{S\subset\{k-d,\dots,k-1\}} f_{k-d-1}(\phi_{k-d-1}(y^{(S)}))\nu^{-1/2}(\phi_{k-d-1}(y^{(S)}))$$

and

$$\begin{aligned}H(y,y') := \mathbb{E}(\prod_{S\subset\{k-d,\dots,k-1\}} &\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)})) \\ &\times \prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)}))|y_{k-d-1}\in\mathbb{Z}_N).\end{aligned}$$

Using the Cauchy-Schwarz inequality,

$$\begin{aligned}|J_d|^2 \le \mathbb{E}(|G(y,y')|^2|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},\\ y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N)\times\\ \times\,\mathbb{E}(|H(y,y')|^2|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},\\ y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N).\end{aligned}$$

On the other hand, since $|f_{k-d-1}(x)|\le\nu(x)$ for every x,

$$\mathbb{E}(|G(y,y')|^2|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N)\le P_d.$$

Moreover, by writing the definition of $H(y,y')$ and expanding the squares by changing the variable y_{k-d-1} by the new variables y_{k-d-1} and y'_{k-d-1}, we see that

$$\begin{aligned}\mathbb{E}(|H(y,y')|^2|y_1,\dots,y_{k-d-2},y_{k-d},\dots,y_{k-1},y'_{k-d},\dots,y'_{k-1}\in\mathbb{Z}_N)\\ = J_{d+1}.\end{aligned}$$

This completes the proof. □

Applying this lemma $(k-1)$ times, we obtain

$$|J_0|^{2^{k-1}} \le J_{k-1}\prod_{d=0}^{k-2} P_d^{2^{k-2-d}}.$$

Observe that, by definition,

$$J_0 = \mathbb{E}\left(\prod_{i=0}^{k-1} f_i(\phi_i(y))|y\in(\mathbb{Z}_N)^{k-1}\right).$$

In order to prove the desired inequality, we choose[6]

$$\phi_i(y) := \sum_{j=1}^{k-1}\left(1-\frac{c_i}{c_j}\right)y_j ,$$

so that $\phi_0(y) = y_1 + \cdots + y_{k-1}$, $\phi_i(y)$ doesn't depend on y_i and, for all y, $\phi(y) = x + c_i r$ where

$$r = -\sum_{i=1}^{k-1}\frac{y_i}{c_i}.$$

Now, the surjective map $\Phi : (\mathbb{Z}_N)^{k-1} \to (\mathbb{Z}_N)^2$ defined by

$$\Phi(y) := (y_1 + \cdots + y_{k-1}, \frac{y_1}{c_1} + \cdots + \frac{y_{k-1}}{c_{k-1}})$$

has a constant number of pre-images, so that a simple calculation shows that

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x+c_j r)|x,r\in\mathbb{Z}_N\right) = \mathbb{E}\left(\prod_{i=0}^{k-1} f_i(\phi_i(y))|y\in(\mathbb{Z}_N)^{k-1}\right) = J_0.$$

However, $P_d = 1 + o(1)$ for each $0 \le d \le k-2$, since ν satisfies the $(2^d, k-1+d, k)$-linear forms condition. In particular, from the previous estimates, we get

$$J_0^{2^{k-1}} \le (1+o(1))J_{k-1}.$$

Fix y. When S varies over the subsets of $\{1,\ldots,k-1\}$, $\phi_0(y^{(S)})$ varies on the $(k-1)$-dimensional cube $\{x + w\cdot h : w \in \{0,1\}^{k-1}\}$, where $x = y_1 + \cdots + y_{k-1}$ and $h_i = y_i' - y_i$, $i = 1,\ldots,k-1$. Hence,

$$J_{k-1} = \mathbb{E}\left(W(x,h)\prod_{w\in\{0,1\}^{k-1}} f_0(x+w\cdot h)|x\in\mathbb{Z}_N, h\in(\mathbb{Z}_N)^{k-1}\right),$$

with the weight $W(x,h)$ given by

$$\begin{aligned} W(x,h) = \mathbb{E}(&\prod_{w\in\{0,1\}}^{k-1}\prod_{i=1}^{k-2}\nu^{1/2}(\phi_i(y+wh))\times \\ &\nu^{1/2}(\phi_{k-1}(y+wh))|y_1,\ldots,y_{k-2}\in\mathbb{Z}_N) \\ = \mathbb{E}(&\prod_{i=1}^{k-2}\prod_{w\in\{0,1\}^{k-1},w_i=0}\nu(\phi_i(y+wh))\times \\ &\prod_{w\in\{0,1\}^{k-1},w_{k-1}=0}\nu(\phi_{k-1}(y+wh))|y_1,\ldots,y_{k-2}\in\mathbb{Z}_N), \end{aligned}$$

[6] Here we are using our hypothesis that c_j are distinct.

where $wh \in (\mathbb{Z}_N)^{k-1}$ is the vector of coordinates $(wh)_j := w_j h_j$ and $y \in (\mathbb{Z}_N)^{k-1}$ is the vector with components y_j for $1 \leq k-2$ and $y_{k-1} := x - y_1 - \cdots - y_{k-2}$. On the other hand, the definition of the Gowers norms says that

$$\mathbb{E}\left(\prod_{w\in\{0,1\}^{k-1}} f_0(x + w\cdot h)|x \in \mathbb{Z}_N, h \in (\mathbb{Z}_N)^{k-1}\right) = \|f_0\|_{U^{k-1}}^{2^{k-1}}.$$

Therefore, it suffices to show that

$$\mathbb{E}\left((W(x,h)-1)\prod_{w\in\{0,1\}^{k-1}} f_0(x + w\cdot h)|x \in \mathbb{Z}_N, h \in (\mathbb{Z}_N)^{k-1}\right) = o(1).$$

Because $|f_j(x)| \leq \nu(x)$, we see that our task is reduced to prove

$$\mathbb{E}\left(|W(x,h)-1|\prod_{w\in\{0,1\}^{k-1}} \nu(x + w\cdot h)|x \in \mathbb{Z}_N, h \in (\mathbb{Z}_N)^{k-1}\right) = o(1).$$

By the Cauchy-Schwarz inequality, this follows directly from the following lemma:

Lemma 2.4.2 (ν uniformly covers its own cubes)**.** *For $n = 0, 2$, it holds*

$$\mathbb{E}\left(|W(x,h)-1|^n\prod_{w\in\{0,1\}^{k-1}} \nu(x + w\cdot h)|x \in \mathbb{Z}_N, h \in (\mathbb{Z}_N)^{k-1}\right) = o(1).$$

Proof. By expanding the square, we see that it suffices to prove that, for $q = 0, 1, 2$, it holds

$$\mathbb{E}\left(W(x,h)^q\prod_{w\in\{0,1\}^{k-1}} \nu(x + w\cdot h)|x \in \mathbb{Z}_N, h \in (\mathbb{Z}_N)^{k-1}\right) = o(1).$$

However, this is a immediate consequence of the linear forms condition:

- for $q = 0$, we apply the $(2^{k-1}, k, 1)$-linear forms condition with variables $x, h_1, \ldots, h_{k-1}$ and linear forms $x + w\cdot h$, $w \in \{0,1\}^{k-1}$;
- for $q = 1$, we apply the $(2^{k-2}(k+1), 2k-2, k)$-linear forms condition with variables $x, h_1, \ldots, h_{k-1}, y_1, \ldots, y_{k-2}$ and linear forms

$$\begin{cases} \phi_i(y + w\cdot h), & w \in \{0,1\}^{k-1}, w_i = 0 \text{ for } 1 \leq i \leq k-1 \\ x + w\cdot h, & w \in \{0,1\}^{k-1}; \end{cases}$$

- for $q = 2$, we apply the $(k2^{k-1}, 3k-4, k)$-linear forms condition with variables
$$x, h_1, \dots, h_{k-1}, y_1, \dots, y_{k-2}, y'_1, \dots, y'_{k-2}$$
and linear forms
$$\begin{cases} \phi_i(y + w \cdot h), \quad w \in \{0,1\}^{k-1}, w_i = 0 \text{ for } 1 \le i \le k-1 \\ \phi_i(y' + w \cdot h), \quad w \in \{0,1\}^{k-1}, w_i = 0 \text{ for } 1 \le i \le k-1 \\ x + w \cdot h, \quad w \in \{0,1\}^{k-1}; \end{cases}$$

Here we are adopting the conventions $y_{k-1} = x - y_1 - \cdots - y_{k-2}$ and $y'_{k-1} = x - y'_1 - \cdots - y'_{k-2}$. Clearly, this completes the proof of the lemma. □

As we told before, this ends the proof of the generalized von Neumann theorem (Theorem 2.4.1). □

Remark 2.4.2. *Note that we used only the linear forms condition during the proof of Theorem 2.4.1.*

Closing the study of the Gowers norms, we state the following simple and useful lemma about the Gowers distance $\|.\|_{U^{k-1}}$ between an arbitrary k-pseudorandom measure ν and the k-pseudorandom measure $\nu_{const} \equiv 1$:

Lemma 2.4.3. *Suppose that ν is a k-pseudorandom measure. Then,*
$$\|\nu - 1\|_{U^d} = o(1),$$
for every $d \le k-1$.

Proof. Observe that the linear forms condition for ν easily implies that $\|\nu\|_{U^{k-1}} = 1 + o(1)$. But, we can refine this reasoning: indeed, let us note that the monotonicity property of the Gowers norms means that it suffices to prove $\|\nu - 1\|_{U^{k-1}} = o(1)$. Multiplying by the factor 2^{k-1}, this reduces our task to prove
$$\mathbb{E}\left(\prod_{w \in \{0,1\}^{k-1}} \nu(x + w \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\right) = o(1).$$
The left-hand side of this identity can be expanded as
$$\sum_{A \subset \{0,1\}^{k-1}} (-1)^{|A|} \mathbb{E}\left(\prod_{w \in A} \nu(x + w \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\right).$$
Looking at the expression
$$\mathbb{E}\left(\prod_{w \in A} \nu(x + w \cdot h) \middle| x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\right)$$

for some fixed $A \subset \{0,1\}^{k-1}$, we see that it has the form

$$\mathbb{E}\left(\nu(\phi_1(\mathbf{x}))\dots\nu(\phi_{|A|}(\mathbf{x}))\middle|\mathbf{x}\in\mathbb{Z}_N^k\right),$$

where $\mathbf{x} := (x, h_1, \dots, h_{k-1})$ and $\phi_1, \dots, \phi_{|A|}$ are an ordering of the $|A|$ linear forms $x \mapsto x + w \cdot h$ with $w \in A$. Obviously, each of these linear forms isn't a rational multiple of any other, so that the $(2^{k-1}, k, 1)$-linear forms condition can be used to conclude the proof of the lemma. □

Let us resume the discussion of this subsection.

Resume of the subsection "Gowers norms":

In this subsection we identified a class of norms (namely, Gowers norms) naturally associated to the problem of counting arithmetic progressions whose elements belong to the support of a given family of functions and we proved Theorem 2.4.1 saying that the Gowers norms can effectively give upper bounds on the number of such progressions up to a negligible error. As we saw during the proof of Roth theorem, this good upper bound is useful during the task of getting lower bounds of certain expectations (our primary goal). The next step will be to introduce the concept of *anti-uniformity*, which plays a fundamental role during the decomposition of arbitrary functions into good and bad parts.

2.4.2 Anti-Uniformity

Since the Gowers norms (for $d \geq 2$) are genuine norms, we can consider the dual norms:

$$\|g\|_{(U^{k-1})^*} := \sup_{\|f\|_{U^{k-1}}\leq 1} |\langle f, g\rangle|,$$

where $\langle f, g\rangle$ denotes the usual L^2 product. We say that g is *anti-uniform* if $\|g\|_{(U^{k-1})^*} = O(1)$ and $\|g\|_{L^\infty} = O(1)$.

Remark 2.4.3. *Although we aren't going to use this fact, observe that, for $k = 3$, the remark 2.4.1 gives the formula:*

$$\|g\|_{(U^2)^*} = \left(\sum_{\xi\in\mathbb{Z}_N} |\widehat{g}(\xi)|^{4/3}\right)^{3/4}.$$

Observe that, if g is anti-uniform and $|\langle f, g\rangle|$ is large then f can't be uniform (i.e., $\|f\|_{L^\infty} = O(1)$ and $\|f\|_{U^{k-1}} = O(1)$) since $|\langle f, g\rangle| \leq \|f\|_{U^{k-1}}\|g\|_{(U^{k-1})^*}$. Thus, this gives us an obstruction to the uniformity.

Moreover, we have a canonical way to construct anti-uniform functions: given $F \in L^1(\mathbb{Z}_N)$, we define the dual of F by:

$$DF(x) := \mathbb{E}(\Pi_{w\neq 0}F(x + w.h)|h \in \mathbb{Z}_N^{k-1}).$$

Among the several elementary properties of these functions, we can quote:

Lemma 2.4.4. *Let ν be a k-pseudorandom measure and $F \in L^1(\mathbb{Z}_{\mathbb{N}})$ be an arbitrary function. It holds:*

- $\langle F, DF\rangle = \|F\|_{U^{k-1}}^{2^{k-1}}$,
- $\|DF\|_{(U^{k-1})^*} = \|F\|_{U^{k-1}}^{2^{k-1}-1}$ *and*
- *if* $|F| \leq 1+\nu$, *then* $\|DF\|_{L^\infty} \leq 2^{2^{k-1}-1} + o(1)$.

Proof. The identity $\langle F, DF\rangle = \|F\|_{U^{k-1}}^{2^{k-1}}$ follows directly from the definitions of Gowers norms and DF, and we left its verification as an exercise. Concerning the second identity, consider $F \neq 0$ (since the case $F = 0$ is trivial) and note that the definition of the dual norms combined with the identity $\langle F, DF\rangle = \|F\|_{U^{k-1}}^{2^{k-1}}$ say that it is sufficient to show that, for any function f, we have

$$|\langle f, DF\rangle| \leq \|f\|_{U^{k-1}} \|F\|_{U^{k-1}}^{2^{k-1}-1}.$$

However, the definition of DF shows that $\langle f, DF\rangle$ is the Gowers inner product of $\langle (f_w)_{w\in\{0,1\}^{k-1}}\rangle_{U^{k-1}}$ where $f_w := f$ when $w = 0$ and $f_w := F$ otherwise, so that the desired inequality above follows from the Gowers-Cauchy-Schwarz inequality.

Finally, the last item follows from the linear forms condition. In fact, since F is bounded by $2(1+\nu)/2 := 2\nu_{1/2}$, we see our task is reduced to prove

$$D\nu_{1/2}(x) \leq 1 + o(1)$$

uniformly for every $x \in \mathbb{Z}_N$. On the other hand, the definition of the dual function says that $D\nu_{1/2}$ can be expanded as

$$\mathbb{E}\left(\prod_{w\in\{0,1\}^{k-1}-\{0\}} \nu_{1/2}(x + w\cdot h) \,\middle|\, h \in \mathbb{Z}_N^{k-1} \right).$$

Since $\nu_{1/2}$ is a k-pseudorandom measure, the linear forms condition implies that this term is $1 + o(1)$. □

Remark 2.4.4. *The proof of this lemma is the* unique *place where we use the linear forms condition with* inhomogeneuos *terms $b_i \neq 0$; indeed, during the previous argument, all of the b_i's are equal to x.*

We call *basic anti-uniform functions* any dual function DF where F is pointwise bounded by $1+\nu$; a relevant property of these functions is its good distribution with respect to ν:

Proposition 2.4.1. *If ν is k-pseudorandom, $\Phi : I^K \to \mathbb{R}$ is continuous and $DF_1, \dots, DF_K$ are basic anti-uniform functions, we define*

$$\Psi(x) = \Phi(DF_1(x), \dots, DF_K(x)).$$

Then, $\langle \nu - 1, \Psi\rangle = o_{k,\Phi}(1)$. Furthermore, the right-hand side quantity can be taken uniform on any compact set of Φ's.

Proof. The basic idea is to use Weierstrass approximation theorem and the fact that ν is a measure to reduce our task to the proof of this proposition to the "simpler" case of a polynomial Φ.

Observe that, by replacing ν by $(\nu+1)/2$, we can assume $|F_j(x)| \leq \nu(x)$ for all $x \in \mathbb{Z}_N$, $1 \leq j \leq K$.

Lemma 2.4.5. *Let $d \geq 1$. For any polynomial function P of degree d with real-valued coefficients (independents of N) it holds*

$$\|P(DF_1, \dots, DF_K)\|_{(U^{k-1})^*} = O_{K,d,P}(1).$$

Proof. By linearity (and increasing K to dK if necessary), it suffices to show the lemma when $P(x_1, \dots, x_K) = x_1 \dots x_K$. In other words, we want to verify

$$\langle f, \prod_{j=1}^{K} DF_j\rangle = O_K(1)$$

for any $f : \mathbb{Z}_N \to \mathbb{R}$ with $\|f\|_{U^{k-1}} \leq 1$. Expanding the left-hand side as

$$\mathbb{E}\left(f(x) \prod_{j=1}^{K} \mathbb{E}(\prod_{w\in\{0,1\}^{k-1}:w\neq 0} F_j(x + w\cdot h^{(j)}) | h^{(j)} \in (\mathbb{Z}_N)^{k-1}) \middle| x \in \mathbb{Z}_N \right)$$

Making the change of variables $h^{(j)} = h + H^{(j)}$ for any $h \in (\mathbb{Z}_N)^{k-1}$, taking the averages on h, we expand the products on j and interchanging the expectations, we can rewrite this expression in terms of Gowers inner product:

$$\mathbb{E}(\langle (f_{w,H})_{w\in\{0,1\}^{k-1}}\rangle_{U^{k-1}} | H \in ((\mathbb{Z}_N)^{k-1})^K),$$

where $H = (H^{(1)}, \dots, H^{(K)})$, $f_{0,H} := f$, $f_{w,H} := g_{w\cdot H}$ for $w \neq 0$ with $w \cdot H = (w \cdot H^{(1)}, \dots, w \cdot H^{(K)})$ and

$$g_{u^{(1)},\dots,u^{(K)}}(x) := \prod_{j=1}^{K} F_j(x + u^{(j)}) \text{ for all } u^{(1)}, \dots, u^{(K)} \in \mathbb{Z}_N.$$

In particular, the Gowers-Cauchy-Schwarz inequality and the fact $\|f\|_{U^{k-1}} \leq 1$ reduce the proof of the lemma to prove the estimate

$$\mathbb{E}\left(\prod_{w\in\{0,1\}^{k-1}:w\neq 0} \|g_{w\cdot H}\|_{U^{k-1}} | H \in ((\mathbb{Z}_N)^{k-1})^K \right) = O_K(1).$$

By Hölder inequality, it suffices to show

$$\mathbb{E}(\|g_{w\cdot H}\|_{U^{k-1}}^{2^{k-1}-1}|H \in ((\mathbb{Z}_N)^{k-1})^K)) = O_K(1),$$

for each $w \in \{0,1\}^{k-1} - 0$.

Fix w. Since $2^{k-1} - 1 \leq 2^{k-1}$ and we have dealing with probability spaces, it suffices to prove

$$\mathbb{E}(\|g_{w\cdot H}\|_{U^{k-1}}^{2^{k-1}}|H \in ((\mathbb{Z}_N)^{k-1})^K)) = O_K(1).$$

This last estimate is true by the following argument: $w \neq 0$ implies that $w \to w \cdot H$ is a covering map; this allows us to use it to perform a change of variables so that the left-hand side of the previous identity is

$$\mathbb{E}(\|g_{u^{(1)},\dots,u^{(K)}}\|_{U^{k-1}}^{2^{k-1}}|u^{(1)},\dots,u^{(K)} \in \mathbb{Z}_N).$$

Using the definitions of the Gowers norms and $g_{u^{(1)},\dots,u^{(K)}}$, we can expand this term as

$$\mathbb{E}\left(\prod_{\widetilde{w}\in\{0,1\}^{k-1}}\prod_{j=1}^{K} F_j(x+u^{(j)}+h\cdot\widetilde{w})\,\middle|\, x,u^{(1)},\dots,u^{(K)} \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\right).$$

By factorizing, we obtain

$$\mathbb{E}\left(\prod_{j=1}^{K}\mathbb{E}(F_j(x+u^{(j)}+h\cdot\widetilde{w})\,|\,u^{(j)} \in \mathbb{Z}_N)\,\middle|\, x \in \mathbb{Z}_N,\, h \in \mathbb{Z}_N^{k-1}\right).$$

Using the assumption $|F_j(x)| \leq \nu(x)$, our task is reduced to the verification of the estimate

$$\mathbb{E}\left(\mathbb{E}(\nu(x+u+h\cdot\widetilde{w})\,|\,u \in \mathbb{Z}_N)^K\,\middle|\, x \in \mathbb{Z}_N,\, h \in \mathbb{Z}_N^{k-1}\right).$$

Performing the change of variables $y = x + u$ and taking the average over x, our goal is to prove

$$\mathbb{E}\left(\mathbb{E}(\nu(y+h\cdot\widetilde{w})\,|\,y \in \mathbb{Z}_N)^K\,\middle|\, h \in \mathbb{Z}_N^{k-1}\right) = O_K(1).$$

At this point, we are ready to use the correlation condition, which says that

$$\mathbb{E}\left(\nu(y+h\cdot\widetilde{w})\,\middle|\, y \in \mathbb{Z}_N\right) \leq \sum_{\widetilde{w},\widetilde{w}'\in\{0,1\}^{k-1},\widetilde{w}\neq\widetilde{w}'} \tau(h\cdot(\widetilde{w}-\widetilde{w}')),$$

where τ is a weight function satisfying $\mathbb{E}(\tau^q) = O_q(1)$. Using the triangular inequality in $L^K(\mathbb{Z}_N^{k-1})$, we see that it suffices to prove

$$\mathbb{E}\left(\tau(h \cdot (\widetilde{w} - \widetilde{w}'))^K \,\middle|\, h \in \mathbb{Z}_N^{k-1}\right) = O_K(1),$$

for all $\widetilde{w}, \widetilde{w}' \in \{0,1\}^{k-1}$ distinct. But, since $h \mapsto h \cdot (\widetilde{w} - \widetilde{w}')$ is a covering map, the left-hand side is $\mathbb{E}(\tau^K) = O_K(1)$. □

Now let's go back to the proof of the proposition. Recall that the lemma 2.4.4 says that a basic anti-uniform function takes its values in the interval $I = [-2^{2^{k-1}}, 2^{2^{k-1}}]$. By Weierstrass approximation theorem, given $\varepsilon > 0$, there exists a polynomial P close to the continuous function Φ in the sense that

$$\|\Phi(DF_1, \dots, DF_K) - P(DF_1, \dots, DF_K)\|_{L^\infty} \leq \varepsilon.$$

Since ν is a measure (i.e., $\mathbb{E}(\nu) = 1 + o(1)$), we have

$$|\langle \nu - 1, \Phi(DF_1, \dots, DF_K) - P(DF_1, \dots, DF_K)\rangle| \leq (2 + o(1))\varepsilon.$$

On the other hand, combining the lemmas 2.4.3, 2.4.5, we obtain

$$|\langle \nu - 1, P(DF_1, \dots, DF_K)\rangle| = o_{K,\varepsilon}(1)$$

because P depends only on K and ε. Making N large (depending on K, ε), we see that

$$|\langle \nu - 1, \Phi(DF_1, \dots, DF_K)\rangle| \leq 4\varepsilon.$$

This ends the proof of the proposition 2.4.1. □

Remark 2.4.5. *The* unique *place in this book where we used the correlation condition was during the final part of the proof of Lemma 2.4.5.*

Resume of the subsection "Anti-Uniformity":

In this subsection we introduced the notion of anti-uniformity, which is useful for the identification of non-uniform functions (that is, they give obstructions to uniformity). In fact, we saw that any function F naturally generates a (basic) anti-uniform function DF such that the correlation $\langle F, DF\rangle$ is large whenever F isn't uniform; moreover, we showed a result saying that a pseudorandom measure is well-distibuted with respect to the *algebra* generated by the basic anti-uniform functions.

In the sequel, we will study the sigma-algebras generated by the level sets of anti-uniform functions, which are the basic pieces of the sigma-algebra whose conditional expectations provides good (=anti-uniform) functions.

2.4.3 Sigma-algebras generated by basic anti-uniform functions

The following proposition says that basic anti-unform functions naturally generate sigma-algebras where they are well-behaved (i.e., one can use Szemerédi theorem in its original form).

Proposition 2.4.2. *If ν is k-pseudorandom measure and $DF_1, \dots, DF_K$ are basic anti-uniform functions, then, for every $\epsilon < 1$ and $\sigma < 1/2$, there exists a sigma-algebra $\mathcal{B}$ such that, if N is a large prime, it holds:*

- $\|DF_j - \mathbb{E}(DF_j|\mathcal{B})\|_{L^\infty} \leq \epsilon \ \forall j$.
- *There is an exceptional subset $\Omega \subset \mathcal{B}$ such that $\mathbb{E}((\nu+1)1_\Omega) = O_{K,\epsilon}(\sigma^{1/2})$.*
- $\|(1-1_\Omega)\mathbb{E}(\nu-1|\mathcal{B})\|_{L^\infty} = O_{K,\epsilon}(\sigma^{1/2})$.

Proof. The starting point of the argument is the following lemma ensuring that each function generates a sigma-algebra:

Lemma 2.4.6. *Let ν be a k-pseudorandom measure, $0 < \epsilon < 1$ and $0 < \sigma < 1/2$ be parameter, and $G \in L^\infty(\mathbb{Z}_N)$ be a function taking its values on the interval $I := [-2^{2^{k-1}}, 2^{2^{k-1}}]$. Then, there exists $\mathcal{B}_{\epsilon,\sigma}(G)$ a sigma-algebra such that*

- *(G belongs to its own σ-algebra) For every σ-algebra $\mathcal{B}$, we have*

$$\|G - \mathbb{E}(G|\mathcal{B} \vee \mathcal{B}_{\epsilon,\sigma}(G))\|_{L^\infty(\mathbb{Z}_N)} \leq \epsilon.$$

- *(Bounded complexity) $\mathcal{B}_{\epsilon,\sigma}(G)$ has $O(1/\epsilon)$ atoms.*
- *(Nice approximation by continuous functions of G) If A is an atom of $\mathcal{B}_{\epsilon,\sigma}(G)$, then there is a $\Psi_A : I \to [0,1]$ such that*

$$\|(1_A - \Psi_A(G))(\nu+1)\|_{L^1(\mathbb{Z}_N)} = O(\sigma).$$

Moreover, Ψ_A belongs to a compact subset $E \subset C^0(I)$ independent of G, ν, N and A.

Proof of lemma 2.4.6. Putting together Fubini's theorem with the fact that ν is a measure, we have

$$\int_0^1 \sum_{n\in\mathbb{Z}} \mathbb{E}(1_{G(x)\in[\epsilon(n-\sigma+\alpha),\epsilon(n+\sigma+\alpha)]}(\nu(x)+1)|x \in \mathbb{Z}_N)d\alpha$$
$$= 2\sigma\mathbb{E}(1+\nu(x)|x \in \mathbb{Z}_N) = O(\sigma).$$

Hence, we can fix α such that

$$\sum_{n\in\mathbb{Z}} \mathbb{E}(1_{G(x)\in[\epsilon(n-\sigma+\alpha),\epsilon(n+\sigma+\alpha)]}(\nu(x)+1)|x\in\mathbb{Z}_N) = O(\sigma). \qquad (2.5)$$

Define $\mathcal{B}_{\epsilon,\sigma}(G)$ as the σ-algebra whose atoms are $G^{-1}([\epsilon(n+\alpha),\epsilon(n+\alpha+1)])$ for $n\in\mathbb{Z}$. Note that $\mathcal{B}_{\epsilon,\sigma}(G)$ is well-defined because the intervals $[\epsilon(n+\alpha),\epsilon(n+\alpha+1)]$ are a partition of the real line.

Clearly, if $\mathcal{B}$ is a σ-algebra, then the function G restricted to a atom of $\mathcal{B}\vee\mathcal{B}_{\epsilon,\sigma}(G)$ takes its values on an interval of diameter ϵ, which gives the first item of the lemma (G belongs to its own σ-algebra). Now, let $A := G^{-1}([\epsilon(n+\alpha),\epsilon(n+\alpha+1)])$ be a atom of $\mathcal{B}_{\epsilon,\sigma}(G)$. Since G takes its values on I, we have $n = O(1/\epsilon)$ (otherwise, $A=\emptyset$). This proves the second item of the lemma (bounded complexity). Finally, let $\psi_\sigma:\mathbb{R}\to[0,1]$ be a fixed bump function such that $\psi_\sigma = 1$ on $[\sigma,1-\sigma]$ and $\psi_\sigma=0$ on $[-\sigma,1+\sigma]$, and define $\Psi_A(x) := \psi_\sigma(\frac{x}{\epsilon}-n-\alpha)$. Obviously, Ψ_A varies on a compact subset $E_{\epsilon,\sigma}$ of $C^0(I)$ (since n and α are bounded) and the identity (2.5) implies the third item of the lemma (nice approximation by continuous functions of G). □

At this point, we come back to the proof of Proposition 2.4.2. We take $B := B_{\epsilon,\sigma}(DF_1)\vee\cdots\vee B_{\epsilon,\sigma}(DF_K)$, where $B_{\epsilon,\sigma}(DF_j)$ is the sigma-algebra provided by Lemma 2.4.6. Clearly the first item of Proposition 2.4.2 follows from the first item of Lemma 2.4.6. On the other hand, since each $B_{\epsilon,\sigma}(DF_j)$ has $O(1/\epsilon)$ atoms, B is generated by $O_{K,\epsilon}(1)$ atoms. We say that an atom A of B is *small* if $\mathbb{E}((\nu+1)1_A)\leq\sigma^{1/2}$. Denote by Ω the union of all small atoms. Then, $\Omega\in B$ and the second item of Proposition 2.4.2 holds. In order to prove the last item of this proposition, it suffices to show

$$\frac{\mathbb{E}((\nu-1)1_A)}{\mathbb{E}(1_A)} = \mathbb{E}(\nu-1|A) = o_{K,\epsilon,\sigma}(1) + O_{K,\epsilon}(\sigma^{1/2})$$

for all non-small atom A. From the smallness definition, we have

$$\mathbb{E}((\nu-1)1_A) + 2\mathbb{E}(1_A) = \mathbb{E}((\nu+1)1_A) \geq \sigma^{1/2}$$

for a non-small A. Thus, since σ is small and N is large, our task is reduced to the verification of

$$\mathbb{E}((\nu-1)1_A) = o_{K,\epsilon,\sigma}(1) + O_{K,\epsilon}(\sigma^{1/2}).$$

However, since A is the intersection of K atoms $A_j\in B_{\epsilon,\sigma}(DF_j)$, $j=1,\ldots,K$, Lemma 2.4.6 and Hölder inequality show that there exists $\Psi_A : I^K\to[0,1]$ such that

$$\|(\nu+1)(1_A-\Psi_A(DF_1,\ldots,DF_K))\|_{L^1(\mathbb{Z}_N)} = O_K(\sigma),$$

so that

$$\|(\nu-1)(1_A-\Psi_A(DF_1,\ldots,DF_K))\|_{L^1(\mathbb{Z}_N)}=O_K(\sigma).$$

Furthermore, Ψ_A belongs to a compact set $E_{\epsilon,K,\sigma}$ of $C^0(I^K)$. This fact and Proposition 2.4.1 (of uniform distribution with respect to basic anti-uniform functions) imply

$$\mathbb{E}((\nu-1)\Psi_A(DF_1,\ldots,DF_K))=o_{K,\epsilon,\sigma}(1)=O_{K,\epsilon}(\sigma^{1/2}),$$

because N is large (depending on K,ϵ,σ). Of course, this estimate and the triangular inequality conclude the argument. □

Resume of the subsection "Sigma-Algebras generated by basic anti-uniform functions":

In this subsection we associated to each basic anti-uniform function DF a *sigma-algebra* $\mathcal{B}$ so that the conditional expectation $\mathbb{E}(DF|\mathcal{B})$ approximates DF (i.e., DF is almost constant on each atom of $\mathcal{B}$) and the pseudorandom measure ν takes values close to 1 on average (with respect to $\mathcal{B}$).

The next stage consists into the utilization of the machinery of basic anti-uniform functions and its associated sigma-algebras to formalize the decomposition process of an arbitrary function into good (anti-uniform) and bad (uniform) parts via an inductive scheme. The key point will be to guarantee that this inductive procedure *stops* in a finite number of steps. This fact will be a consequence of an *energy increment argument.*

2.4.4 The energy increment argument

Using the sigma-algebras generated by basic anti-uniform functions, we can obtain the desired decomposition into uniform and anti-uniform parts:

Theorem 2.4.2 (generalized Koopman-von Neumann theorem)**.** *Consider ν a k-pseudorandom measure, $f\in L^1$ such that $0\le f\le\nu$, $\epsilon<<1$ and N a large prime number. Then, there exists a sigma-algebra $\mathcal{B}$ and an exceptional subset $\Omega\in\mathcal{B}$ such that:*

- $\mathbb{E}(\nu\cdot 1_\Omega)=o_\epsilon(1)$ *(the exceptional set is small).*
- $\|(1-1_\Omega)\mathbb{E}(\nu-1|\mathcal{B})\|_{L^\infty}=o_\epsilon(1)$ *(good distribution of functions outside the exceptional subset).*
- $\|(1-1_\Omega)(f-\mathbb{E}(f|\mathcal{B})\|_{U^{k-1}}\le\epsilon^{1/2^k}$ *(uniformity on $\mathcal{B}$)*

Proof. The basic strategy is the same of the structure theorem of Furstenberg[7]: we start with the trivial sigma-algebra $\mathcal{B} = \{\emptyset, \mathbb{Z}_N\}$ and we look at the function $f - \mathbb{E}(f|\mathcal{B})$. If it is uniform (i.e., the third item above holds), we are done. Otherwise, we use the results about anti-uniformity to find an anti-uniform function G_1 with non-trivial correlation with f and we add the level sets of G_1 to the sigma-algebra $\mathcal{B}$. The non-trivial correlation property will tell us that the L^2 norm of $\mathbb{E}(f|\mathcal{B})$ increases by a non-trivial amount[8], while the pseudorandomness of ν shows that $\mathbb{E}(f|\mathcal{B})$ stays uniformly bounded. At this point, we repeat this procedure until $f - \mathbb{E}(f|\mathcal{B})$ becomes sufficiently uniform; note that the algorithm stops in a finite number of steps (of order $2^{2^k}/\epsilon$) due to the definite energy increment at each step.

Now let us write this strategy in a more organized manner. Fix ϵ and let K_0 be the smallest integer greater than $1+2^{2^k}/\epsilon$. We will need a parameter $0 < \sigma \ll \epsilon$ and we will take N large depending on ϵ and σ. We construct $\mathcal{B}$ and Ω via a sequence of basic anti-uniform functions $DF_1, \ldots, DF_K$ on $\mathbb{Z}_N$, exceptional subsets $\Omega_0 \subset \cdots \subset \Omega_K \subset \mathbb{Z}_N$ and sigma-algebras $\mathcal{B}_0 \subset \cdots \subset \mathcal{B}_K$ for some $0 \leq K \leq K_0$ as follows:

- Stage 0: We begin with K=0, $\mathcal{B}_0 := \{\emptyset, \mathbb{Z}_N\}$ and $\Omega_0 := \emptyset$.

- Stage 1: We put $F_{K+1} := (1 - 1_{\Omega_K})(f - \mathbb{E}(f|\mathcal{B}_K))$. If
$$\|F_{K+1}\|_{U^{k-1}} \leq \epsilon^{1/2^k}$$
we define $\Omega := \Omega_K$, $\mathcal{B} = \mathcal{B}_K$ and we end the algorithm successfully. Otherwise, we go to the stage 2.

- Stage 2: If
$$\|F_{K+1}\|_{U^{k-1}} > \epsilon^{1/2^k}$$
we define $\mathcal{B}_{K+1} := \mathcal{B}_K \vee \mathcal{B}_{\epsilon,\sigma}(DF_{K+1})$ and we go to the stage 3.

- Stage 3: We look for an exceptional subset $\Omega_{K+1} \supset \Omega_K$ in $\mathcal{B}_{K+1}$ with
$$\mathbb{E}((\nu+1)1_{\Omega_{K+1}}) = O_{K,\epsilon}(\sigma^{1/2}) \tag{2.6}$$
and
$$\|(1 - 1_{\Omega_{K+1}})\mathbb{E}(\nu - 1|\mathcal{B}_{K+1})\|_{L^\infty} = O_{K,\epsilon}(\sigma^{1/2}).$$
If such an exceptional subset can't be located, we end the algorithm with an error. Otherwise, we proceed to the stage 4.

[7] This result says that we can decompose any dynamical system in a weak-mixing extension of a tower of compact extensions.

[8] The idea that non-trivial correlation implies an augmentation of the L^2 norm is precisely the *energy increment argument*.

- Stage 4: We replace K by $K+1$. If $K > K_0$, we end the algorithm with an error. Otherwise, we go back to stage 1.

Assume momentarily that the algorithm ends without any errors in the stages 3 or 4. Then, it is clear that after K_0+1 interactions (at most), we constructed a σ-algebra $\mathcal{B}$ and an exceptional subset Ω with the desired properties, except by the fact that the error terms are $O_{K,\epsilon}(\sigma^{1/2})$ instead of $o_\epsilon(1)$, for a large N depending on σ, K, ϵ. However, this little difficulty is easily overcome by making σ converging to zero very slowly.

In other words, we reduced the proof of this theorem to show that the algorithm ends without any errors. The argument is inductive: as an induction hypothesis on $0 \leq K_1 \leq K_0$, suppose that the algorithm ends without any errors or it attains the stage 2 at the K_1-th interaction without any errors. Note that this is obvious for $K_1 = 0$. Assuming this is valid for some $K_1 < K_0$, we would like to show that this is true for K_1+1. Observe that we can assume that the algorithm doesn't end before the stage 2 of the K_1-th interaction. At this stage, we have σ-algebras $\mathcal{B}_0, \dots, \mathcal{B}_{K_1+1}$, basic anti-uniform functions $DF_1, \dots, DF_{K_1+1}$ and exceptional subsets $\Omega_0, \dots, \Omega_{K_1}$ already constructed. We claim that

$$\|DF_j\|_{L^\infty} \leq 2^{2^{k-1}} + O_{j,\epsilon}(\sigma^{1/2}), \tag{2.7}$$

for any $1 \leq j \leq K_1+1$. This follows from the stage 3 of the previous interactions (or the stage 0 when $j=1$) because they say that

$$\|(1 - 1_{\Omega_{j-1}})\mathbb{E}(\nu - 1|\mathcal{B}_{j-1})\|_{L^\infty} = O_{j,\epsilon}(\sigma^{1/2}),$$

so

$$\mathbb{E}(\nu|\mathcal{B}_{j-1})(x) = 1 + O_{j-1,\epsilon}(\sigma^{1/2}),$$

for all $x \notin \Omega_{j-1}$. Since $0 \leq f(x) \leq \nu(x)$, we conclude the pointwise estimates

$$0 \leq (1 - 1_{\Omega_{j-1}}(x))\mathbb{E}(f|\mathcal{B}_{j-1})(x) \leq 1 + O_{j,\epsilon}(\sigma^{1/2}), \tag{2.8}$$

so that, we can see that, by the definition of F_j,

$$|F_j(x)| \leq (1 + O_{j,\epsilon}(\sigma^{1/2}))(\nu(x)+1). \tag{2.9}$$

In particular, a simple application of Lemma 2.4.4 proves our claimed estimate (2.7).

On the other hand, since $\mathcal{B}_{K_1+1}$ is the σ-algebra generated by

$$\mathcal{B}_{\epsilon,\alpha_1}(DF_1), \dots, \mathcal{B}_{\epsilon,\alpha_{K_1+1}}(DF_{K_1+1}),$$

Proposition 2.4.2 permits to find some subset Ω such that

$$\mathbb{E}((\nu+1)1_\Omega) = O_{K_1,\epsilon}(\sigma^{1/2})$$

and

$$\|(1-1_\Omega)\mathbb{E}(\nu-1|\mathcal{B}_{K_1+1})\|_{L^\infty} = O_{K_1,\epsilon}(\sigma^{1/2}).$$

Define $\Omega_{K_1+1} := \Omega \cup \Omega_{K_1}$. Obviously, Ω_{K_1+1} has the required properties to execute the stage 3 without errors, so that we can go to the stage 2 of the K_1+1-th interaction (or we end the algorithm without any errors), so that the inductive argument is complete.

In other words, at this moment, we proved that there are only two possibilities for the algorithm: either it ends without errors or it goes all the way to the K_0-th interaction. In order to conclude the proof of the theorem, we affirm that, if the algorithm reaches the stage 3 of the K_0-th iterate, then the following key-property (namely, the *energy increment* estimate) holds:

$$\begin{aligned}&\|(1-1_{\Omega_j})\mathbb{E}(f|\mathcal{B}_j)\|_{L^2}^2 \\ &\geq \|(1-1_{\Omega_{j-1}})\mathbb{E}(f|\mathcal{B}_{j-1})\|_{L^2}^2 \\ &\quad + 2^{2^k-2}\epsilon - O_{j,\epsilon}(\sigma^{1/2}) - O(\epsilon^2),\end{aligned} \tag{2.10}$$

for every $1 \leq j \leq K_0$ (if N is large depending on K_0 and ϵ). Of course, this key property is sufficient to conclude the proof of the theorem because the estimate (2.8) gives us

$$0 \leq \|(1-1_{\Omega_j})\mathbb{E}(f|\mathcal{B}_j)\|_{L^2}^2 \leq 1 + O_{j,\epsilon}(\sigma^{1/2}), \tag{2.11}$$

for any $0 \leq j \leq K_0$. Since K_0 is the smallest integer greater than $2^{2^k}/\epsilon+1$, the pigeonhole principle leads us to a contradiction for sufficiently small ϵ and σ, and for a sufficiently large N depending on ϵ, σ.

Finally, it remains only to prove the energy increment estimate. The idea is to exploit the fact that the algorithm doesn't stop at the stage 2 of the $(j-1)$-th iteraction: indeed, observe that this fact implies

$$\|F_j\|_{U^{k-1}} \geq \epsilon^{1/2^k}.$$

Combining this estimate with the definition of F_j and Lemma 2.4.4, we obtain

$$|\langle (1-1_{\Omega_{j-1}})(f-\mathbb{E}(f|\mathcal{B}_{j-1})), DF_j\rangle| = \|F_j\|_{U^{k-1}}^{2^{k-1}} \geq \epsilon^{1/2}.$$

On the other hand, the pointwise estimates (2.7), (2.9), (2.6) above show that

$$\langle (1_{\Omega_j} - 1_{\Omega_{j-1}})(f-\mathbb{E}(f|\mathcal{B}_{j-1}), DF_j)\rangle = O_{j,\epsilon}(\sigma^{1/2}),$$

while Lemma 2.4.6 and the estimate (2.9) tell us

$$\langle (1-1_{\Omega_j})(f-\mathbb{E}(f|\mathcal{B}_{j-1}), DF_j - \mathbb{E}(DF_j|\mathcal{B}_j))\rangle = O(\epsilon).$$

In particular, by the triangular inequality, we obtain the lower bound:

$$|\langle(1-1_{\Omega_j})(f-\mathbb{E}(f|\mathcal{B}_{j-1})),\mathbb{E}(DF_j|\mathcal{B}_j)\rangle| \geq \epsilon^{1/2}-O_{j,\epsilon}(\sigma^{1/2})-O(\epsilon).$$

Since the functions $(1-1_{\Omega_j})$, $\mathbb{E}(DF_j|\mathcal{B}_j)$ and $\mathbb{E}(f|\mathcal{B}_{j-1})$ are $\mathcal{B}_j$-measurable, we can replace f by $\mathbb{E}(f|\mathcal{B}_j)$, so that we get

$$|\langle(1-1_{\Omega_j})(\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1})),\mathbb{E}(DF_j|\mathcal{B}_j)\rangle| \geq \epsilon^{1/2}-O_{j,\epsilon}(\sigma^{1/2})-O(\epsilon).$$

Using the Cauchy-Schwarz inequality and the estimate (2.7), we conclude:

$$\|(1-1_{\Omega_j})(\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1}))\|_{L^2} \geq 2^{-2^{k-1}+1}\epsilon^{1/2}-O_{j,\epsilon}(\sigma^{1/2})-O(\epsilon). \tag{2.12}$$

This inequality morally implies, by the Pythagorean theorem, the energy increment estimate, although there is a little problem due to the presence of the exceptional subsets Ω_{j-1},Ω_j, which are not *a priori* trivial to control because we don't dispose of good L^2 bounds for ν. In order to treat this little technicality, we notice that (2.6) and (2.8) imply

$$\|(1_{\Omega_j}-1_{\Omega_{j-1}})\mathbb{E}(f|\mathcal{B}_{j-1})\|_{L^2}=O_{j,\epsilon}(\sigma^{1/2}).$$

Thus, the triangular inequality and (2.11) show that the energy increment estimate (2.10) follows from the following estimate

$$\begin{aligned}&\|(1-1_{\Omega_j})\mathbb{E}(f|\mathcal{B}_j)\|^2_{L^2}\\ &\geq \|(1-1_{\Omega_{j-1}})\mathbb{E}(f|\mathcal{B}_{j-1})\|^2_{L^2}+\epsilon^{1/2}-O_{j,\epsilon}(\sigma^{1/2})-O(\epsilon).\end{aligned}$$

However, the left-hand side above can be expanded (by the cosine law) as

$$\begin{aligned}&\|(1-1_{\Omega_j})\mathbb{E}(f|\mathcal{B}_{j-1})\|^2_{L^2}+\|(1-1_{\Omega_j})(\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1}))\|^2_{L^2}\\ &+2\langle(1-1_{\Omega_j})\mathbb{E}(f|\mathcal{B}_{j-1}),(1-1_{\Omega_j})(\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1}))\rangle.\end{aligned}$$

Therefore, by (2.12), we see that it suffices to prove the following quasi-orthogonality relation:

$$\langle(1-1_{\Omega_j})\mathbb{E}(f|\mathcal{B}_{j-1}),(1-1_{\Omega_j})(\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1}))\rangle=O_{j,\epsilon}(\sigma^{1/2}).$$

Since $(1-1_{\Omega_j})^2=(1-1_{\Omega_j})$, we can rewrite this identity as

$$\langle(1-1_{\Omega_j})\mathbb{E}(f|\mathcal{B}_{j-1}),\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1})\rangle=O_{j,\epsilon}(\sigma^{1/2}).$$

Now we observe that, since the function $(1-1_{\Omega_{j-1}})\mathbb{E}(f|\mathcal{B}_{j-1})$ is measurable with respect to $\mathcal{B}_{j-1}$, it is orthogonal to the function $\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1})$ because $\mathcal{B}_{j-1}$ is a sub-sigma-algebra of $\mathcal{B}_j$ (by construction). In particular, we can again rewrite the previous expression as

$$\langle(1_{\Omega_j}-1_{\Omega_{j-1}})\mathbb{E}(f|\mathcal{B}_{j-1}),\mathbb{E}(f|\mathcal{B}_j)-\mathbb{E}(f|\mathcal{B}_{j-1})\rangle=O_{j,\epsilon}(\sigma^{1/2}).$$

Using again the fact that $(1_{\Omega_j} - 1_{\Omega_{j-1}})\mathbb{E}(f|\mathcal{B}_{j-1})$ is a $\mathcal{B}_j$-measurable (so that it must be orthogonal to $f - \mathbb{E}(f|\mathcal{B}_j)$), we see that the previous identity is equivalent to

$$\langle (1_{\Omega_j} - 1_{\Omega_{j-1}})\mathbb{E}(f|\mathcal{B}_{j-1}), f - \mathbb{E}(f|\mathcal{B}_{j-1})\rangle = O_{j,\epsilon}(\sigma^{1/2}).$$

However, this equality is certainly true in view of $0 \leq f \leq \nu$ and the estimates (2.6), (2.8). This proves the energy increment estimate (2.10) and, consequently, the proof of Theorem 2.4.2. □

Resume of the subsection "The energy increment argument":

In this subsection, we used the machinery of sigma-algebras associated to the anti-uniform functions to exhibit, for a density f bounded by a pseudo-random measure, an *algorithmic* construction of *small* exceptional subsets and a sigma-algebra such that the function f has an *uniform* behavior outside the exceptional subset. In particular, this shows us how to decompose f into an uniform (bad) part and an anti-uniform (good) part. This was the content of the generalized Koopman-von Neumann theorem (Theorem 2.4.2). Moreover, the algorithm related to the proof of the generalized Koopman-von Neumann theorem was finite (i.e., it stops after a finite steps of interaction) due to the energy increment argument at each step (indeed, since the energy was always increasing by a definite amount and it is *a priori* bounded during the whole process, this gives immediately the desired conclusion).

The last step of the proof of Green-Tao-Szemerédi theorem is the application of this decomposition (provided by Theorem 2.4.2) in order to conclude the argument.

2.4.5 End of the proof of Green-Tao-Szemerédi theorem

Once we formalized (and quantified) the machinery related to the uniformity and anti-uniformity issues, and the decomposition of arbitrary functions into uniform and non-uniform pieces, it is fairly easy to imitate the scheme proposed in the section 2.3 during the proof of Roth theorem in order to get a proof of Green-Tao-Szemerédi theorem (Theorem 2.2.1):

Let f, ν and δ be the objects appearing in the statement of Theorem 2.2.1. Take $\epsilon << \delta$ and consider $\mathcal{B}$ the sigma-algebra provided by the generalized Koopman-von Neumann theorem (Theorem 2.4.2). Define the functions:

- $f_U = (1 - 1_\Omega)(f - \mathbb{E}(f|\mathcal{B}))$
- $f_{AU} = (1 - 1_\Omega)\mathbb{E}(f|\mathcal{B})$.

Recall that, by hypothesis, $0 \leq f \leq \nu$ (pointwise) and $\mathbb{E}(f) \geq \delta$. Hence, Theorem 2.4.2 ensures that

$$\mathbb{E}(f_{AU}) = \mathbb{E}((1-1_\Omega)f) \geq \mathbb{E}(f) - \mathbb{E}(\nu 1_\Omega) \geq \delta - o_\epsilon(1).$$

Moreover, we have $f_{AU} \leq 1 + o_\epsilon(1)$, so that we can use Szemerédi theorem[9]. In particular:

$$\mathbb{E}(f_{AU}(n)\dots f_{AU}(n+(k-1)r)|n,r\in\mathbb{Z}_N) \geq c(k,\delta) - o_\epsilon(1).$$

On the other hand, we know that $\|f_U\|_{U^{k-1}} \leq \epsilon^{1/2^k}$. Thus, by the generalized von Neumann theorem 2.4.1, we have:

$$\mathbb{E}(f_{*_1}(n)\dots f_{*_k}(n+(k-1)r)|n,r\in\mathbb{Z}_N) = O(\epsilon^{1/2^k})$$

where $*_j = U$ or AU, and $*_j = U$ for at least one index j.

Therefore, we obtain:

$$\mathbb{E}(f(n)\dots f(n+(k-1)r)|n,r\in\mathbb{Z}_N) \geq c(k,\delta) - O(\epsilon^{1/2^k}) - o_\epsilon(1).$$

Since ϵ is arbitrary, Green-Tao-Szemerédi theorem is proved!

Remark 2.4.6. *The careful reader noticed the strong analogy between the previous estimates and the estimates of the proposition 2.3.1. In fact, as it is natural to expect, we separated, in both arguments, a "good" term (it was $\Lambda_3(g,g,g)$ in Roth case and $\mathbb{E}(f_{AU}(n)\dots f_{AU}(n+(k-1)r)|n,r\in\mathbb{Z}_N)$ in Green-Tao case) which is relatively big (its order was δ^3 in Roth case and $c(k,\delta)-o_\epsilon(1)$ in Green-Tao case) and our task was to control the "bad" terms ($\Lambda_3(.,.,.)$ where the bad function b must appear in some entry in Roth case and $\mathbb{E}(f_{*_1}(n)\dots f_{*_k}(n+(k-1)r)|n,r\in\mathbb{Z}_N)$ where the bad function f_U must appear in some entry in Green-Tao case). In order to accomplish this goal, we used Hölder inequality in Roth case and the generalized von Neumann theorem in Green-Tao case to reduce the problem to the (non-trivial) "fact" that b in Roth case and f_U in Green-Tao can be taken* uniform. *Logically, this fact was obtained from the generalized Koopman-von Neumann theorem in Green-Tao case, which uses (in its proof) the energy increment argument, as we annouced in the beginning to the section.*

[9] In fact we are omitting a little detail here: since f_{AU} is not exactly bounded from above by 1 and $\mathbb{E}(f_{AU})$ is not exactly bounded from below by δ, Szemerédi theorem doesn't apply directly. However, this is easily overcome by using Szemerédi theorem for a function equal to f modulo a term of the form $o_\epsilon(1)$ which can be trivially controlled in this case.

Part II

Elkies-McMullen theorem

Chapter 3

Elkies-McMullen theorem

3.1 Distribution of sequences on the circle

A basic problem in Number Theory consists into the study of the distribution of a given sequence of real numbers on the circle $\mathbb{R}/\mathbb{Z}$.

More precisely, given a real number $x \in \mathbb{R}$, let us denote by

$$\{x\} = x(\text{mod } 1) \in S^1 := \mathbb{R}/\mathbb{Z}$$

its fractionary part. In this setting, the problem quoted above can be described as follows: given a sequence of real numbers x_n, $n = 1, 2, \dots$, what can we say about the behavior of the sequence $\{x_n\}$ on the circle S^1?

The distribution problem of given sequences (mod 1) is very old, so that we will make just a few comments. Firstly, a classical result (due to Kronecker) says that the sequence $\{n\theta\}$ is dense in S^1 for all *irrational* $\theta \in \mathbb{R}$. Another classical result (due to Weyl) says that the same sequence $\{n\theta\}$ (for irrational θ) is *equidistributed*, i.e., for all interval $I \subset S^1$,

$$\frac{\#\{1 \leq n \leq N : \{n\theta\} \in I\}}{N} \to \frac{|I|}{|S^1|} \quad \text{when} \quad N \to \infty,$$

where $|J|$ denotes the length of J. Furthermore, it is known that the sequence $\{\kappa^n\}$ is equidistributed *for almost every* $\kappa > 1$ (although certain specific cases such as $\{(3/2)^n\}$ being still open). More generally, we have Weyl's criterion saying that a given sequence is equidistributed (mod 1) if and only if certain associated *exponential sums* converge to zero.

A particularly interesting example for our future purposes is the sequence $\{n^\alpha\}$ where $0 < \alpha < 1$. A preliminary simple result about this sequence is the fact that it is equidistributed:

Exercise 3.1.1. *Show that* $\{n^\alpha\}$ *is equidistributed on* S^1 *for each* $0 < \alpha < 1$. Hint*: By making a suitable re-interpretation of the problem on the real line* $\mathbb{R}$*, use the fact that* $(n+1)^\alpha - n^\alpha \to 0$ *and* $n^\alpha \to \infty$ *when* $n \to \infty$.

A quite popular fashion of studying the distribution of a sequence x_n consists into the analysis of the gaps $J_1, \dots, J_N$ determined by these points on S^1, i.e., $\mathcal{J}(N) := \{J_1, \dots, J_N\}$ are the connected components of $S^1 - \{x_1, \dots, x_N\}$. Observe that the sum of the lengths $|J_i|$ of the gaps J_i equals to 1, so that the average length of these gaps is:

$$\frac{1}{N}\sum_{i=1}^{N}|J_i| = 1/N.$$

In other words, the "natural scale" for the study of the lengths $|J_i|$ of the gaps is $1/N$. Taking this "natural scale" into account, we introduce the following definition:

Definition 3.1.1. *We say that a sequence x_n is* exponentially distributed *whenever, for any $0 \leq a \leq b$, we have*

$$\frac{1}{N}\#\{J \in \mathcal{J}(N) : |J| \in [a/N, b/N]\} \to \int_a^b e^{-t}dt$$

when $N \to \infty$.

Example 3.1.1. *A random choice of points on the circle gives an exponentially distributed sequence (for further details see the page 158 of Feller's book [5]).*

Coming back to the sequence $\{n^\alpha\}$ with $0 < \alpha < 1$, some numerical experiments suggest that it is exponentially distributed for every $\alpha \neq 1/2$. Furthermore, M. Boshernitzan observed (numerically) in 1993 a special distribution for the specific case $\alpha = 1/2$. However, a rigorous confirmation of this numerical observation of Boshernitzan was only recently obtained by N. Elkies and C. McMullen.

More precisely, N. Elkies and C. McMullen [6] showed the following theorem about the distribution of the gaps of the sequence $\{\sqrt{n}\}$ (mod 1):

Theorem 3.1.1 (Elkies and McMullen (2004))**.** *The distribution of the gaps of $\{\sqrt{n}\}$ is given by the continuous function*

$$F(t) := \begin{cases} 6/\pi^2, & t \in [0, 1/2], \\ F_2(t), & t \in [1/2, 2], \\ F_3(t), & t \in [2, \infty), \end{cases}$$

where $F_2(t)$ and $F_3(t)$ are explicit *real-analytic functions. Moreover, $F(t)$* isn't *analytic (and it isn't even C^3) at the points $t = 1/2$ and $t = 2$. In other words, $F(t)$ exhibits a* genuine phase transition *at the points $t = 1/2$ and $t = 2$.*

For the reader's convenience, we recall that $G(t)$ is the distribution of the gaps of a given sequence x_n whenever, for any $0 \leq a \leq b$, we have

$$\frac{1}{N}\#\{J \in \mathcal{J}(N) : |J| \in [a/N, b/N]\} \to \int_a^b G(t)dt \quad \text{when} \quad N \to \infty.$$

In the statement of Elkies and McMullen theorem, we avoided the introduction of the explicit expressions of the functions $F_2(t)$ and $F_3(t)$ in order to get a short statement. However, it is not hard to write the formulas of these functions. To do so, we denote by $r := 1/2x$ and define

$$F_2(x) = \frac{6}{\pi^2}(\frac{2}{3}(4r-1)^{\frac{3}{2}}\psi(r) + (1-6r)\log r + 2r - 1) \quad \text{if} \quad \frac{1}{2} \leq x \leq 2,$$

$$F_3(x) = \frac{6}{\pi^2}(f(\alpha) - g(\alpha) - h(\alpha)) \quad \text{if} \quad x \geq 2.$$

Here $\psi(r) = \tan^{-1}[(2r-1)/\sqrt{4r-1}] - \tan^{-1}[1/\sqrt{4r-1}]$, $2\alpha = 1-\sqrt{1-4r}$, $f(\alpha) = 4(1-4\alpha)(1-\alpha)^2\log(1-\alpha)$, $g(\alpha) = 2(1-2\alpha)^3\log(1-2\alpha)$ and $h(\alpha) = 2\alpha^2$.

In particular, we see that $F(t)$ is continuous at $t = 1/2$ and $t = 2$ (with values $F(1/2) = 6/\pi^2$ and $F(2) = 6(\log 2 - 1/2)/\pi^2$). Also, $F(t)$ is C^1 but $F(t)$ is not C^2 at $t = 2$ and $F(t)$ is not C^3 at $t = 1/2$ (as a simple argument with Taylor series around these points shows). Another direct consequence of this explicit formula of $F(t)$ is the fact that the "tail" of the distribution of $\sqrt{n}$ (mod 1) *is not* exponential: $F(t) \sim 3t^{-3}/\pi^2$ when $t \to \infty$. Comparing this fact with the example 3.1.1, we see that the appearance of large gaps is more frequent for the sequence $\sqrt{n}$ (mod 1) than it is for a random sequence of points.

This being said, we will concentrate our efforts on the discussion of the beautiful proof of Elkies and McMullen theorem. Since the proof of this result involves a number of technical steps, we will describe in the next paragraph a very rough sketch of the main arguments, leaving the details to the subsequent sections.

Roughly speaking, Elkies and McMullen idea consists into a translation of the problem of the computation of the distribution $F(t)$ of the gaps of $\sqrt{n}$ (mod 1) to the calculation of the probability of a *random affine lattice* of $\mathbb{R}^2$ to intersect a certain fixed triangle. The advantage of this apparently artificial approach resides in the fact that the Ergodic Theory of random (affine) lattices is well-understood due to the so-called Ratner theory, which allows us to compute precisely the desired probability (of a random lattice to meet a fixed triangle).

At this point, we are ready to discuss this scheme with a little bit more of details.

3.2 Ergodic version of Elkies-McMullen theorem

Before starting the translation of the calculation of the distribution of $\sqrt{n}$ (mod 1) to an ergodic-theoretical problem, let us introduce some notation. We remember that $\Lambda_0 \subset \mathbb{R}^2$ is a *lattice* if Λ_0 is a discrete subgroup isomorphic to $\mathbb{Z}^2$. We say that a lattice Λ_0 is *unimodular* if the torus $\mathbb{R}^2/\Lambda_0$ has unit area. Furthermore, an *affine lattice* $\Lambda \subset \mathbb{R}^2$ is a subset of the following form: $\Lambda = \Lambda_0 + v$ where $v \in \mathbb{R}^2$ and Λ_0 is a lattice.

We denote by E the space of unimodular affine lattices. As the reader can easily check, E is naturally identified with the space

$$E = ASL(2,\mathbb{R})/ASL(2,\mathbb{Z}),$$

where $ASL(2,\mathbb{R})$ is the group of affine transformations $g : \mathbb{R}^2 \to \mathbb{R}^2$ of the form $g(v) = Av + b$ with $\det A = 1$ and $ASL(2,\mathbb{Z})$ is the discrete subgroup of $ASL(2,\mathbb{R})$ whose elements $g(v) = Av + b$ verify $A \in SL(2,\mathbb{Z})$ and $b \in \mathbb{Z}^2$.

An immediate consequence of this identification is the fact that E comes equipped with an *unique* probability measure μ_E which is invariant by the left action of $ASL(2,\mathbb{R})$ (namely, μ_E is the so-called Haar measure of E). In particular, μ_E gives a sense to the notion of a *random lattice* of E: a random lattice is a lattice with generic properties with respect to μ_E, i.e., a lattice belonging to a μ_E-full measure subset of μ_E.

Keeping these notations in mind, we are ready to give an informal scheme of the proof of Elkies and McMullen theorem.

3.2.1 Scheme of the proof of Elkies and McMullen theorem

Let us fix N a large integer, $t > 0$ and $I = [x, x + t/N] \subset \mathbb{R}/\mathbb{Z}$, where $x \in [0,1]$ is randomly choosen (with respect to Lebesgue measure). In these terms, the Elkies and McMullen program is:

1. computing the gap distribution of $\sqrt{n}$ (mod 1) is equivalent to calculate the probability $P_N(t)$ of some point $\{\sqrt{n}\}$ with $0 \leq n \leq N$ to belong to I;

2. on the other hand, $\{\sqrt{n}\} \in I \iff \sqrt{n} \in I + a$ for some $a \in \mathbb{Z} \iff n \in (I + a)^2$ for some $a \in \mathbb{Z}$;

3. now, for the purpose of the computation of the gap distribution of $\{\sqrt{n}\}$, it is possible to show that one can harmlessly replace $(I + a)^2$ by its *linear approximation*

$$(I + a)^2 \sim (a + x)^2 + 2(a + x)(I - x) = a^2 - x^2 + 2(a + x)I;$$

4. also, one can show that we can assume that N is a square (without loss of generality);

5. under these assumptions, we look at the linear approximation $a^2 - x^2 + 2(a+x)I$ of $(I+a)^2$ and note that
$$n \in a^2 - x^2 + 2(a+x)I \quad (\text{for } 0 \leq n \leq N)$$
$$\Updownarrow$$
$$(\mathbb{Z} + x^2) \cap 2(a+x)I \neq \emptyset \quad (\text{for } 0 \leq a+x \leq \sqrt{N});$$

6. this last condition can be rewritten as $T \cap \mathbb{Z}^2 \neq \emptyset$ where $T \subset \mathbb{R}^2$ is the triangle of area t given by
$$T := \{(a,b) : b + x^2 \in 2(a+x)I \text{ and } a+x \in [0,\sqrt{N}]\};$$

7. denoting by S_t the "standard" triangle of area t with vertices $(0,0)$, $(1,0)$, $(0,2t)$ and considering $g \in ASL(2,\mathbb{R})$ the unique afinne transformation with $g(T) = S_t$ and $g(-x,-x^2) = (0,0)$, we can rephrase the above condition as
$$g(\mathbb{Z}^2) \cap S_t \neq \emptyset;$$

8. in resume, this (sketchy) discussion allows us to translate the computation of the probability $P_N(t)$ of some element $\{\sqrt{n}\}$ with $0 \leq n \leq N$ to belong to $I = [x, x+t/N]$ into the problem of calculating the probability of the affine lattice $\Lambda_N(x) := g(\mathbb{Z}^2)$ to intersect the standard triangle S_t;

9. on the other hand, for $N \gg 1$ large, we expect that $\Lambda_N(x) \in E$ behaves like a random affine lattice: in fact, we will see that, as any good random affine lattice, the sequence $\Lambda_N(x)$ is uniformly distributed - for every compactly supported continuous function f of E it holds
$$\int_0^1 f(\Lambda_N(x))dx \to \int_E f(\Lambda) d\mu_E(\Lambda) \quad \text{when} \quad N \to \infty;$$

10. using this uniform distribution result, we get that $P_N(t) \to p(t)$ when $N \to \infty$, where $p(t)$ is the probability of a random affine lattice to intersect the standard triangle S_t;

11. finally, by reversing this translation, one can show that $p''(t) = -F(t)$, where $F(t)$ is the gap distribution of$\{\sqrt{n}\}$, which ends the proof of Elkies and McMullen theorem because $p''(t)$ can be explicitly calculated from natural expressions for the Haar probability μ_E.

After this informal description of Elkies and McMullen arguments, we will start a more or less detailed discussion of all of these topics.

3.2.2 Some preliminary reductions

For each $N \geq 1$ integer, we define a function $\lambda_N : [0, \infty) \to [0, 1]$ as follows. We consider the N points $\{\sqrt{n}\}$, $n = 1, \dots, N$, of the circle $\mathbb{R}/\mathbb{Z}$. This gives us a partition of the circle into N intervals $J_1, \dots, J_N$ (among them, $\lfloor \sqrt{N} \rfloor - 1$ have length zero). In this setting,

$$\lambda_N(x) := \frac{1}{N} \#\{1 \leq i \leq N : |J_i| < x/N\}.$$

We left the verification of the following elementary properties of $\lambda_N(x)$ as an exercise to the reader:

Exercise 3.2.1. *Show that $\lambda_N(x)$ has the following properties:*

- *λ_N is a non-decreasing left-continuous function (indeed, λ_N is constant except for a finite number of jump discontinuities); moreover, $\lambda_N(0) = 0$ and $\lambda_N(\infty) = 1$;*
- *$\int_0^\infty (1 - \lambda_N(x))dx = \int_0^\infty x d\lambda_N(x) = 1$ (because $\int_0^\infty x d\lambda_N(x)$ is the sum of the lengths of the gaps).*

In these terms, Elkies and McMullen theorem is equivalent to:

Theorem 3.2.1. *There exists a continuous function $\lambda_\infty : [0, \infty) \to [0, 1]$ such that $\lambda_N \to \lambda_\infty$ uniformly in compact sets when $N \to \infty$. Moreover,*

$$\lambda_\infty(x) = \int_0^x F(\xi) d\xi,$$

where F is an explicit function to be computed later.

Observe that F is the desired gap distribution of $\{\sqrt{n}\}$: in fact, for all $0 \leq x_1 < x_2 < \infty$, the quantity of gaps of $\{\sqrt{n}\}$ $(1 \leq n \leq N)$ with length between x_1/N and x_2/N is asymtotic to $\int_{x_1}^{x_2} F(\xi) d\xi$.

In order to study λ_N, we introduce $L_N : \mathbb{R}/\mathbb{Z} \to [0, \infty)$ defined by

$$L_N(t) = \begin{cases} 0, & \text{if } t = \{\sqrt{n}\} \text{ for some } 0 \leq n \leq N \\ N \times \text{length of the gap containing } t, & \text{otherwise.} \end{cases}$$

Using L_N we can write the union of the gaps of length $< x/N$ as

$$I_N(x) := \{t \in \mathbb{R}/\mathbb{Z} : L_N(t) < x\}.$$

In particular, $|I_N(x)| = \int_0^x \xi d\lambda_N(\xi)$, so that the theorem 3.2.1 is equivalent to:

Theorem 3.2.2. *$|I_N(x)| \to \int_0^x \xi F(\xi) d\xi$ uniformly on x varying on compact sets when $N \to \infty$.*

Now let us talk about the preliminary reductions of these theorems. The first simplification of the statement of Theorem 3.2.1 was annouced at the fourth item of Elkies and McMullen scheme discussed above: namely, in the statement of this theorem, we can assume that N is a square, i.e., $N = s^2$ with $s \in \mathbb{Z}$. More precisely, we have the following result:

Lemma 3.2.1 (Lemma 3.1 of Elkies and McMullen)**.** *Suppose that λ_{s^2} converges (uniformly on compact sets) to a continuous function λ_∞ when $s \to \infty$. Then, λ_N also converges (uniformly on compact sets) to λ_∞ when $N \to \infty$.*

Proof. Observe that any integer N is far from some square s^2 by a distance of $O(\sqrt{N})$. On the other hand, by replacing N by s^2, we change $3|N-s^2| \lesssim \sqrt{N}$ of the lengths of the gaps (at most) and we multiply the normalizing factor $1/N$ by $N/s^2 = 1 + O(\sqrt{1/N})$. In particular, the desired lemma follows. □

The second simplification was described at the second and third items of Elkies and McMullen program: during the study of $\lambda_N(x)$ we can replace the quadratic expressions by its linear approximations without changing the asymptotic results. In order to formalize this fact, we will need more notation. Recall that each integer n can be uniquely written as $a^2 + b$ where $a = \lfloor \sqrt{n} \rfloor = \sqrt{n} - \{\sqrt{n}\}$. Using Lemma 3.2.1, we can assume that N is a perfect square, say $N = s^2$. In this situation, we see that

$$L_N(t) = N(t_2 - t_1)$$

where t_2 is the smallest real number $\geq t$ with $(a_2 + t_2)^2 \in \mathbb{Z}$ for some integer $a_2 < s$ and t_1 is the biggest real number $\leq t$ with $(a_1 + t_1)^2 \in \mathbb{Z}$ for some integer $a_1 < s$. In order to understand the properties of the function L_N we make the following arithmetical remark:

Remark 3.2.1. *For $a_j \in \mathbb{Z}$ ($j = 1, 2$), we have $(a_j + t_j)^2 = a_j^2 + 2a_jt_j + t_j^2 \in \mathbb{Z}$ if and only if $b_j := 2a_jt_j + t_j^2 \in \mathbb{Z}$. Furthermore, we have $0 \leq b_j \leq (a_j + 1)^2 - a_j^2 = 2a_j + 1$ when $0 \leq t_j \leq 1$.*

Using this remark, we can rewrite the identity $L_N(t) = N(t_2 - t_1)$ as

$$L_N(t) = N((t_2 - t) - (t_1 - t)) = N\left(\min_{r_t(a,b)\geq 0} r_t(a,b) - \max_{r_t(a,b)\leq 0} r_t(a,b)\right)$$

where a, b vary over the set of integers satisfying

$$0 < a < s \quad \text{and} \quad 0 \leq b \leq 2a + 1$$

and $r_t(a,b) := \sqrt{a^2 + b} - a - t$.

Remark 3.2.2. *In fact, this last condition on b is superfluous: on one hand, $0 \leq b \leq 2a+1$ is equivalent to $0 \leq r_t(a,b)+t \leq 1$ and, on the other hand, $\min\limits_{r_t(a,b)\geq 0} r_t(a,b)+t$ and $\max\limits_{r_t(a,b)\leq 0} r_t(a,b)$ belong to the interval $[0,1]$ since $r_t(1,0)+t=0$ and $r_t(1,3)+t=1$.*

Going back to the analysis of L_N, we will apply the idea discussed in the third item of the Elkies and McMullen program: in the definition of b_j (see remark 3.2.1), we replaced t_j^2 by its linear approximation $t^2+2t(t_j-t) = 2t_jt-t^2$ nearby t. Consequently, b_j is substituted by $2a_jt_j+(2t_jt-t^2) = 2(a+t)t_j-t^2$. By this reason, we will consider $\tau_j = (b_j+t^2)/2(a+t)$ the solution of the equation

$$2(a+t)\tau_j - t^2 = b_j$$

and we will replace t_j by τ_j in the definition of L_N, so that we get the function

$$\widetilde{L}_N(t) := N\left(\min_{\rho_t(a,b)\geq 0} r_t(a,b) - \max_{\rho_t(a,b)\leq 0} r_t(a,b)\right)$$

where $\rho_t(a,b) := \frac{b+t^2}{2(a+t)} - t = \frac{a^2+b-(a+t)^2}{2(a+t)}$ and a, b vary over the set of integers satisfying

$$0 < a < s \quad \text{and} \quad 0 \leq b \leq 2a+1.$$

Remark 3.2.3. *Similarly to Remark 3.2.2, this last condition on b is superfluous.*

Analogously to the definition of the set $I_N(x)$, we introduce

$$\widetilde{I}_N(x) := \{t \in \mathbb{R}/\mathbb{Z} : \widetilde{L}_N(t) < x\}.$$

As we told in the third item of the Elkies and McMullen program, replacing t_j^2 by its linear approximation (or equivalently the replacement of t_j by τ_j) in the definition of L_N doesn't affect the asymptotics. More precisely, we have the following theorem:

Theorem 3.2.3 (Proposition 3.2 of Elkies and McMullen)**.** *Suppose that $|\widetilde{I}_N(x)|$ converges (uniformly on compact sets) to $\int_0^x \xi F(\xi)d\xi$ with F continuous. Then, the same happens to I_N:*

$$|I_N(x)| \to \int_0^x \xi F(\xi)d\xi$$

when $N \to \infty$ (uniformly on compact sets).

A "conceptual" explanation for the validity of this theorem goes as follows: typically, we expect that $t_j = t + O(1/N)$, so that $|t_j - \tau_j| = O(1/a_j N^2)$ (since τ_j is the solution of the equation $2(a_j + t)\tau_j - t^2 = b_j = 2a_j t_j + t_j^2$). Moreover, given $\varepsilon > 0$, we get $1/a_j < \varepsilon$ for the "majority" of the pairs (a_j, b_j). Hence, the natural expectations is the replacement of t_j by τ_j changes the lengths of most of the gaps by $O(\varepsilon/N^2)$, which certainly doesn't affect the asymptotic behavior of $|I_N|$.

In order to formalize the previous heuristic, we consider the quocient

$$\frac{\rho_t(a,b)}{r_t(a,b)} = \frac{\sqrt{a^2+b}+a+t}{2(a+t)} \in \left[\frac{2a+1}{2a+2}, \frac{2a+1}{2a}\right].$$

Manipulating this information (see Lemma 3.3 of Elkies and McMullen paper), it follows that

Lemma 3.2.2. *For all $t \in [0,1]$, it holds $\frac{3}{4}L_N(t) \leq \widetilde{L}_N(t) \leq \frac{3}{2}L_N(t)$. Furthermore, for all $A \in \mathbb{N}$, we have the estimate*

$$\frac{2A+1}{2A+2}\widetilde{L}_N(t) \leq \frac{2A+1}{2A}L_N(t)$$

for all $t \in [0,1]$ except for a subset of size $\leq (A+2)(A-1)/(s-1)$.

As the reader can check, this lemma easily implies the following estimates

$$|\widetilde{I}_N(3x/4)| \leq |I_N(x)| \leq |\widetilde{I}_N(3x/2)|$$

and

$$|\widetilde{I}_N(\frac{2A+1}{2A+2}x)| - O(A^2/s) \leq |I_N(x)| \leq |\widetilde{I}_N(\frac{2A+1}{2A}x)| + O(A^2/s)$$

for all $x \in [0,\infty)$ and $A = 1, 2, \dots$. Using these estimates with $A = 1 + \lfloor s^{1/3} \rfloor$ (or any other increasing function of s with $A^2/s \to 0$), we get the proof of Theorem 3.2.3.

Once we have the theorem 3.2.3 in our toolbox, our goal is reduced to the asymptotic study of $\widetilde{L_N}$. In this direction, we are going to interpret (in the next subsection) this quantity in terms of the Ergodic Theory of affine lattices (as proposed in the items 5, 6, 7 and 8 of Elkies and McMullen scheme.

3.2.3 Geometrical interpretation of $\widetilde{L}_N$

As we already told in the item 6 of the program, we are going to use a convenient triangle T (whose specific form was described above) in our arguments. In our current notation, T is the triangle of the (a,b)-plane whose interior is determined by the inequalities

$$0 < a+t < s, \ 2c_-(a+t)/s^2 < b - 2ta - t^2 < 2c_+(a+t)/s^2,$$

for $c_- < 0 < c_+$. Making a direct translation between our notations, the reader can easily check that $\widetilde{L}_N$ is interpreted in terms of T as described in the following lemma:

Lemma 3.2.3. *For each $N = s^2$ and $t \in [0,1]$, we have the following possibilities:*

- *if $\widetilde{L}_N(t) \neq 0$, then $\widetilde{L}_N(t)$ is the area $c_+ - c_-$ of the biggest triangle T as above whose interior doesn't intersect $\mathbb{Z}^2 - \{(0,0),(0,1)\}$;*
- *if $\widetilde{L}_N(t) = 0$, then all triangles T as above contains the point (a,b) whose coordinates satisfy $b - 2ta - t^2 = 0$ with $0 < a < s$.*

Now, we will "clean" the statement of this lemma by observing that the possibilities $(a,b) = (0,0),(0,1)$ don't affect $\widetilde{L}_N(t)$ except for $t \in [0,1]$ in a subset of size $O(1/s)$:

Lemma 3.2.4. *The characterization of the quantity $\widetilde{L}_N(t)$ in lemma 3.2.3 isn't affected by the inclusion of the extra cases $(a,b) = (0,0),(0,1)$ except when $0 \leq t < 1/(s-1)$ or $t^{-1} - t < 1/(s-1)$.*

The proof of this lemma is a simple case-by-case verification (for further details, see Lemma 3.7 of Elkies and McMullen). Using this lemma, it follows that the inclusion of the extra cases $(a,b) = (0,0),(0,1)$ doesn't affect the asymptotic of $|\widetilde{I}_N(t)|$ (because this minor modification alters the values of $\widetilde{L}_N(t)$ only on a subset of size $O(1/s)$).

Now we will apply the idea discussed in the item 7 of the Elkies and McMullen scheme: consider the affine transformation g of $\mathbb{R}^2$ defined by

$$g(a,b) = (w_1, w_2) = (s(b - 2ta - t^2), (a+t)/s).$$

Observe that g sends the vertex $(-t, -t^2)$ to the origin $(0,0)$, the triangle T into the "standard" triangle

$$\Delta_{c_-,c_+} := \{(w_1,w_2) \in \mathbb{R}^2 : 0 < w_2 < 1,\, 2c_-w_2 < w_1 < 2c_+w_2\}$$

and the lattice $\mathbb{Z}^2$ to the lattice $\Lambda_{s^2}(t) = g(\mathbb{Z}^2)$ given by

$$\Lambda_{s^2}(t) := \{(s(b - 2ta - t^2), (a+t)/s) : (a,b) \in \mathbb{Z}^2)\}.$$

Note that the "standard" triangle Δ_{c_-,c_+} depends on c_-, c_+ but it doesn't depend on s,t (for the effect of comparison, this standard triangle corresponds to the triangle S_t of item 7 of Elkies and McMullen program).

Closing our series of translations, we introduce the following definition:

Definition 3.2.1. *Given an affine lattice Λ in the (w_1,w_2)-plane, we denote by $L(\Lambda)$ the area $c_+ - c_-$ of the biggest triangle of the form Δ_{c_-,c_+} which is disjoint from Λ with the conventions $L(\Lambda) = 0$ when such a triangle doesn't exist and $L(\Lambda) = \infty$ when all of these triangles are disjoint from Λ.*

Using this notation, we can apply Lemma 3.2.4 to resume the discussion of this subsection into the following proposition:

Proposition 3.2.1 (Proposition 3.8 of Elkies and McMullen)**.** *For all $s \in \mathbb{N}, x \in \mathbb{R}$, the set of $t \in [0,1]$ with $L(\Lambda_{s^2}(t)) \leq x$ has size*

$$|\widetilde{I}_{s^2}(x)| + O(1/s).$$

In other words, Proposition 3.2.1 says that the issue of studying the asymptotic of $\widetilde{I}_{s^2}$ is reduced to the study of the behavior of the function L on the family of affine lattices $\Lambda_{s^2}(t)$.

At this point, it remains "only" to complete the details of the items 9, 10, 11 of the program. This will be performed in the next two subsections. Morally, these items essentially say that the study of L on the affine lattices $\Lambda_{s^2}(t)$ can be done through the Ergodic Theory of random affine lattices (in particular, Ratner's theorems will be useful during this task).

3.3 Study of L via Ratner theorems

In the sequel, we will use Ratner's theorems (about the Ergodic Theory of homogeneous flows) in order to understand the values of L along the family of (affine) lattices $\Lambda_{s^2}(t)$. Namely, we will invoke the following result saying that the family $\{\Lambda_{s^2}(t) : t \in [0,1]\}$ of "circles" of affine lattices become equidistributed on the space of affine lattices E when $s \to \infty$:

Theorem 3.3.1. *For any $f \in C_0(E)$ we have*

$$\int_0^1 f(\Lambda_{s^2}(t))dt \to \int_E f d\mu_E \quad \textit{when} \quad s \to \infty.$$

For the discussion of this subsection, we will assume this equidistribution theorem and we will see how it helps in the determination of the asymptotic gap distribution F of $\{\sqrt{n}\}$.

3.3.1 Computation of F assuming Theorem 3.3.1

We recall that Proposition 3.2.1 proved in the previous section says that the size of the subset of $t \in [0,1]$ such that $L(\Lambda_{s^2}(t)) \leq x$ is $|\widetilde{I}_{s^2}(x)| + O(1/s)$. Putting this information together with the theorem 3.3.1 above, we have the following consequence:

Proposition 3.3.1. *For $x \in [0, \infty)$ we have*

$$|\widetilde{I}_{s^2}(x)| \to \mu_E(\{\Lambda \in E : L(\Lambda) \leq x\} \quad \textit{when} \quad s \to \infty.$$

Proof. Consider $E_x := \{\Lambda \in E : L(\Lambda) \leq x\}$. Using this notation, the fact that the size of the subset of $t \in [0,1]$ with $L(\Lambda_{s^2}(t)) \leq x$ is $|\widetilde{I}_{s^2}(x)| + O(1/s)$ can be rewritten as:

$$\int_0^1 \chi_{E_x}(\Lambda_{s^2}(t))dt = |\widetilde{I}_{s^2}(x)| + O(1/s)$$

This reduces our task to show that $\int_0^1 \chi_{E_x}(\Lambda_{s^2}(t))dt$ converges to $\mu_E(E_x)$. To do so, the basic idea is to use the theorem 3.3.1. However, it is not possible to make a direct application of this result because the characteristic function χ_{E_x} is not continuous. A simple remedy to this difficulty is the classical L^1-approximation argument of χ_{E_x} and $1 - \chi_{E_x}$ by some continuous function in the space $C_0(E)$ combined with the theorem 3.3.1. At this stage, it remains only to know whether such compactly supported continuous L^1-approximations exist. From the basic result of Real Analysis, we know that the functions χ_{E_x} and χ_{E-E_x} can be approximated by some functions in $C_0(E)$ whenever $\mu_E(\partial E_x) = 0$.

In resume, the proof of this proposition is complete once we can show that $\mu_E(\partial E_x) = 0$. In this direction, we invite the reader to prove that $L : E \to [0,\infty]$ is a *submersion* for almost all points of E: more precisely, L isn't a submersion only at the affine lattices Λ containing the origin $(0,0)$ or some point of the horizontal side $w_2 = 1$ of its maximal triangle Δ_{c_-,c_+}. In particular, for each x, the corresponding points E_x where L isn't a submersion is a closed subset of μ_E-measure zero. Thus, by the inverse theorem, we see that the level sets of L have zero μ_E-measure and, *a fortiori*, it follows that $\mu_E(\partial E_x) = 0$. This ends the proof. □

A direct corollary of this proposition (and the results of the previous section) is:

Proposition 3.3.2. *Suppose that the asymptotic gap distribution $F(\xi)$ of $\{\sqrt{n}\}$ is* continuous. *Then,*

$$\lim_{N\to\infty} |I_N(x)| = \lim_{N\to\infty} |\widetilde{I}_N(x)| = \mu_E(\{\Lambda \in E : L(\Lambda) \leq x\})$$

for $x \in [0,\infty)$. Moreover, this convergence is uniform on compact sets (with respect to x).

Proof. Assuming that F is continuous, we can combine Lemma 3.2.1, Theorem 3.2.3 and Proposition 3.3.1 to get the desired result. □

Although the statement of Proposition 3.3.2 seems quite promising (since we wrote I_N asymptotically in terms of the measure μ_E of the subset $L^{-1}([0,x])$), we are still not ready to compute the gap distribution F: in fact, the discussion of the previous section tells us only that $I_N(x) \to$

$\int_0^x \xi F(\xi)d\xi$, so that in order to get a concrete formula for F in terms of $\mu_E(L^{-1}([0,x]))$ we should extract the two derivatives (with respect to the x variable) of this function. However, at the present moment, it is not clear even that $\mu_E(L^{-1}([0,x]))$ is differentiable!

Hence, we should analyze more carefully the subsets $L^{-1}([0,x])$. Keeping this goal in mind, we introduce the subset S_{c_-,c_+} of E formed by the affine lattices Λ with some point inside the triangle Δ_{c_-,c_+}, where $c_- < 0 < c_+$. Observe that $\mu_E(S_{c_-,c_+})$ depends only on the area $c_+ - c_-$ of the triangle Δ_{c_-,c_+} because any two triangles with the same area are equivalent under some element of $ASL_2(\mathbb{R})$ and the measure μ_E is $ASL_2(\mathbb{R})$-invariant. In particular, we can define a function $p : [0,\infty] \to [0,1]$ by

$$p(c_+ - c_-) := \mu_E(S_{c_-,c_+})$$

with the conventions that $p(0) = 0$ and $p(\infty) = \infty$.

As we already mentioned, we will eventually take two derivatives of p in order to find an explicit formula for F:

Lemma 3.3.1. *Suppose that $p \in C^2$ (i.e., p is twice differentiable and p'' is continuous). Then,*

$$F(x) = -p''(x).$$

Proof. We write $\mu_E(\{\Lambda \in E : L(\Lambda) < x\})$ as a telescopic sum:

$$\mu_E(S_{0,x}) - \lim_{M\to\infty} \sum_{j=0}^{M-1} [\mu_E(S_{\frac{(j+1)x}{M}-x,\frac{jx}{M}}) - \mu_E(S_{\frac{jx}{M}-x,\frac{jx}{M}})].$$

Putting this equation in terms of p, we get

$$\mu_E(\{\Lambda \in E : L(\Lambda) < x\}) = p(x) - \lim_{M\to\infty} M\left(p(x) - p(x - \frac{x}{M})\right).$$

Since p is differentiable, it follows that

$$\mu_E(\{\Lambda \in E : L(\Lambda) < x\}) = p(x) - xp'(x).$$

On the other hand, assuming that p is twice differentiable, we know that $\frac{d}{dx}(p(x) - xp'(x)) = -xp''(x)$. Moreover, since $p(0) = 0$, we see that $p(x) - xp'(x) = 0$ on $x = 0$. Combining these two facts, we get $p(x) - xp'(x) = \int_0^x -\xi p''(\xi)d\xi$.

Using these two identities, we obtain

$$\mu_E(\{\Lambda \in E : L(\Lambda) < x\}) = \int_0^x -\xi p''(\xi)d\xi.$$

This completes the proof of the lemma in view of Proposition 3.3.2 and the fact that $|I_N(x)| \to \int_0^x \xi F(\xi)d\xi$. □

Remark 3.3.1. *Still assuming that $p \in C^2$, from the definition of p and Lemma 3.3.1, we see that*

$$F(x) = -p''(x) = -\frac{\partial^2}{\partial c_- \partial c_+}\mu_E(S_{c_-,c_+})$$

for any $c_- < 0 < c_+$ with $c_+ - c_- = x$. This gives the following geometrical interpretation of $F(x)$ in terms of μ_E: the value $F(c_+ - c_-)dc_-dc_+$ is the measure of the subset of affine lattices $\Lambda \in E$ intersecting Δ_{c_-,c_+} into exactly two points - one of them with coordinates (w_1, w_2) verifying $w_1/2w_2 \in (c_-, c_- + dc_-)$ and the other one with coordinates (w_1, w_2) verifying $w_1/2w_2 \in (c_+ - dc_+, c_+)$.

From Lemma 3.3.1, the calculation of the gap distribution of F of $\{\sqrt{n}\}$ is reduced to the explicit computation of the function p and the verification of the fact $p \in C^2$. In this direction, we will recall some known facts about the theory of unimodular lattices.

We denote by B the space of unimodular lattices of $\mathbb{R}^2$ (i.e., discrete subgroups Λ^0 isomorphic to $\mathbb{Z}^2$ with covolume 1) and μ_B the Haar probability measure of B. A vector $w \in \Lambda^0$ of a lattice $\Lambda^0 \in B$ is called *primitive* whenever one can find $w' \in \Lambda^0$ such that $\{w, w'\}$ is a $\mathbb{Z}^2$-basis of Λ^0. Equivalently, $w \in \Lambda^0$ is primitive when $w/k \notin \Lambda^0$ for any $k > 1$. In the sequel, we are going to use the following facts:

- the subset $Z_w \subset B$ of lattices with w as a primitive vector is a circle (in fact it is a closed horocycle);
- given $K \subset \mathbb{R}^2$ a convex compact subset, the area of K is $\zeta(2) \times \int_B f_K(\Lambda^0)d\mu_B$ where $f_K(\Lambda^0)$ is the quantity of primitive vectors of Λ^0 inside K;
- in particular, taking K sufficiently small so that $f_K(\Lambda^0) \leq 1$ for all $\Lambda^0 \in B$, we see that the subset of lattices with a primitive vector inside K has μ_B-measure equal to $1/\zeta(2)$ times the area of K;
- moreover, we can desintegrate the μ_B-measure of a measurable subset $\widetilde{B} \subset B$ as follows: $\mu_B(\widetilde{B}) = \frac{1}{\zeta(2)}\int_{w \in K}\mu_w(\widetilde{B} \cap Z_w)$, where μ_w is the (normalized) Lebesgue measure on the circle Z_w.

At this point, our objective is to combine the remark 3.3.1 with the desintegration of μ_B in order to express F as a double integral. In this direction, in view of the geometrical interpretation of F (see remark 3.3.1), we look at the lattices $\Lambda \in E$ intersecting the triangle Δ_{c_-,c_+} into two points whose coordinates $(w_1^{(i)}, w_2^{(i)})$ $(i = 1, 2)$ verify $w_1^{(1)}/2w_2^{(1)} \in (c_-, c_- + dc_-)$ and $w_1^{(2)}/2w_2^{(2)} \in (c_+ - dc_+, c_+)$. Note that the difference between these two points is a primitive vector of Λ: otherwise, Λ would contain a third point inside the line segment determined by these two points; since Δ_{c_-,c_+} is convex (since it is a triangle) it would follow that Λ intersects Δ_{c_-,c_+} into three points, a contradiction with our hypothesis. Using this primitive vector, we apply the desintegration of μ_B to write F as an integral on the w_2-coordinates v_-, v_+ of the vectors of Λ at the boundary of Δ_{c_-,c_+}: for $v_-, v_+ \in (0,1)$, we write $w = (2c_+v_+, v_+) - (2c_-v_-, v_-)$ and we recall that Z_w parametrize the subset of lattices containing w; next, we denote by $q_x(v_-, v_+) \in [0,1]$ the (μ_w)-measure of the subset of Z_w formed by the lattices do not touching the interior of Δ_{c_-,c_+}. Observe that we write q_x instead of q_{c_-,c_+} because this quantity depends only on $x = c_+ - c_-$. In this notation, we can express F as a double integral:

Proposition 3.3.3. *The function $(x, v_-, v_+) \mapsto q_x(v_-, v_+)$ is continuous except at a certain subset of $\{v_- = v_+\}$. Moreover, for $x \in [0,\infty)$, we have*

$$-p''(x) = F(x) = \frac{1}{\zeta(2)} \int_0^1 \int_0^1 4v_-v_+q_x(v_-, v_+)dv_-dv_+.$$

In particular, it follows that F is continuous.

Proof. The fact that $q_x(v_-, v_+)$ is continuous is immediate except when the vector w is horizontal (in particular, it is parallel to the third side of Δ_{c_-,c_+}). This shows the first part of this proposition since w is horizontal implies $v_- = v_+$. Next, since $0 \leq q_x(v_-, v_+) \leq 1$, the double integral above exists and it varies continuously on the x-variable. Finally, in order to see that this integral coincides with $-p''(x)$ and $F(x)$, we use the geometrical interpretation of F (see the paragraph before the statement of this proposition) combined with the fact that $4v_-v_+$ is the product of the lengths of the line segments

$$\{(w_1, v_-) : 2c_-v_- < w_1 < 2(c_- + dc_-)v_-\}$$

and

$$\{(w_1, v_+) : 2(c_+ - dc_+)v_+ < w_1 < 2c_+\,v_+\}$$

where the vectors of the lattices reside, and the desintegration formula of μ_B. □

In order to make Proposition 3.3.3 more useful, we need to compute $q_x(v_-, v_+)$. The idea is to do a series of geometrical considerations after a affine change of variables from (w_1, w_2) to (z, z') sending the triangle Δ_{c_-,c_+} on the isosceles triangle

$$\Delta_0 := \{(z, z') \in \mathbb{R}^2 : z, z' > 0, z + z' < 1\}$$

of area 1/2. Since the triangle Δ_{c_-,c_+} has area $c_+ - c_-$, this affine transformation multiplies the area by the factor

$$r := 1/2x.$$

Although the argument is not complicated, we will refer the curious reader to Lemma 3.12 of Elkies and McMullen for a detailed proof of the following fact:

Lemma 3.3.2. *For any $0 < v, v' \leq 1$ and $x > 0$ it holds $q_x(v, v') = q_x(v', v)$. Moreover, for $v \geq v'$, we have*

$$q_x(v,v') = \max\left\{0, \min\left(1, \frac{r}{vv'}\right) - \max\left(0, \frac{v(1-v')-r}{v(v-v')}\right)\right\}$$

with $r = 1/2x$. Here we are using the following convention

$$\max\left(0, \frac{v(1-v')-r}{v(v-v')}\right) = \begin{cases} \infty & \text{if } v = v' \text{ and } r < v(1-v') \\ 0 & \text{if } v = v' \text{ and } r \geq v(1-v') \end{cases}$$

Once this fact is available, the task of finding an explicit formula for F (the asymptotic gap distribution of $\{\sqrt{n}\}$) becomes a Calculus I exercise. Indeed, combining Proposition 3.3.3 with Lemma 3.3.2 and computing some integrals (as in the proof of Theorem 3.14 of Elkies and McMullen paper), the reader will eventually prove the following result:

Theorem 3.3.2. *It holds*

$$F(t) = \begin{cases} 6/\pi^2, & t \in [0, 1/2], \\ F_2(t), & t \in [1/2, 2], \\ F_3(t), & t \in [2, \infty), \end{cases}$$

where $F_2(t)$ and $F_3(t)$ are

$$F_2(x) = \frac{6}{\pi^2}(\frac{2}{3}(4r-1)^{\frac{3}{2}}\psi(r) + (1-6r)\log r + 2r - 1)$$

and

$$F_3(x) = \frac{6}{\pi^2}(f(\alpha) - g(\alpha) - h(\alpha)).$$

Here $r := 1/2x$ and $\psi(r) = \tan^{-1}[(2r-1)/\sqrt{4r-1}] - \tan^{-1}[1/\sqrt{4r-1}]$, $\alpha = (1-\sqrt{1-4r})/2$, $f(\alpha) = 4(1-4\alpha)(1-\alpha)^2\log(1-\alpha)$, $g(\alpha) = 2(1-2\alpha)^3\log(1-2\alpha)$ and $h(\alpha) = 2\alpha^2$.

In other words, we completed the proof of Elkies and McMullen Theorem 3.1.1 *modulo* Theorem 3.3.1 (which we were assuming during the entire subsection)!

Now, we end this subsection and we pass to the study of the relation between Theorem 3.3.1 with the Ergodic Theory of homogenous flows.

3.3.2 Theorem 3.3.1 and homogenous flows

We recall that Theorem 3.3.1 concerns the equidistribution of the family of circles of affine lattices $\{\Lambda_{s^2}(t) : t \in [0,1]\}$ when $s \to \infty$. In order to reformulate this theorem into a more appropriate language, we observe that the entire action takes places at the special affine group $ASL_2(\mathbb{R})$ which we will denote by

$$G(\mathbb{R}) := \left\{ \begin{pmatrix} a & b & x \\ c & d & y \\ 0 & 0 & 1 \end{pmatrix} : ad - bc = 1 \right\} \subset SL_3(\mathbb{R}).$$

Note that this group acts on $\mathbb{R}^2$ via the conservative affine transformations

$$\begin{pmatrix} X \\ Y \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix}.$$

Denote by $G(\mathbb{Z}) \subset G(\mathbb{R})$ the subgroup of matrices with integer entries and observe that the space of unimodular affine lattices E is naturally identified with $G(\mathbb{R})/G(\mathbb{Z})$: we take the integral lattice $\mathbb{Z}^2$ as a base point and for each $g \in G(\mathbb{R})$ we associate the affine lattice

$$\Lambda(g) := \left\{ (w_1, w_2) \in \mathbb{R}^2 : \begin{pmatrix} w_1 \\ w_2 \\ 1 \end{pmatrix} \in g \begin{pmatrix} \mathbb{Z} \\ \mathbb{Z} \\ 1 \end{pmatrix} \right\}.$$

This map is surjective and $\Lambda(g) = \Lambda(h)$ if and only if $h \in g \cdot G(\mathbb{Z})$ (as the reader can easily check), so that this map is an isomorphism between E and $G(\mathbb{R})/G(\mathbb{Z})$.

In the particular case of the affine lattices $\Lambda_{s^2}(t)$, the corresponding elements of $G(\mathbb{R})/G(\mathbb{Z})$ under this isomorphism can be explicitly calculated as follows: recall that

$$\Lambda_{s^2}(t) := \{(s(b - 2ta - t^2), (a+t)/s)\},$$

so that the points $(w_1, w_2) \in \Lambda_{s^2}(t)$ in matricial notation are:

$$\begin{pmatrix} w_1 \\ w_2 \\ 1 \end{pmatrix} = \begin{pmatrix} s & -2st & -st^2 \\ 0 & 1/s & t/s \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b \\ a \\ 1 \end{pmatrix} = A_s U(t) \begin{pmatrix} b \\ a \\ 1 \end{pmatrix},$$

where $A_s = diag(s, 1/s, 1)$ is the diagonal matrix

$$A_s := \begin{pmatrix} s & 0 & 0 \\ 0 & 1/s & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$U(t) := \begin{pmatrix} 1 & -2t & -t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, $\Lambda_{s^2}(t)$ is the affine lattice

$$\Lambda_{s^2}(t) = \left\{ (w_1, w_2) \in \mathbb{R}^2 : \begin{pmatrix} w_1 \\ w_2 \\ 1 \end{pmatrix} \in A_s U(t) \begin{pmatrix} \mathbb{Z} \\ \mathbb{Z} \\ 1 \end{pmatrix} \right\}.$$

In other words, $\Lambda_{s^2}(t)$ is identified with $A_s U(t)$ via the isomorphism Λ. In resume, we see that Theorem 3.3.1 is equivalent to:

Theorem 3.3.3. *The circles $\{A_s U(t) : t \in [0,1]\}$ become equidistributed in $G(\mathbb{R})/G(\mathbb{Z})$ when $s \to \infty$, i.e., for all $f \in C_0(E)$, it holds*

$$\lim_{s\to\infty} \int_0^1 f(A_s U(t))dt = \int_E f d\mu_E.$$

Once we translated Theorem 3.3.1 to Theorem 3.3.3, our plan is to use the Ergodic Theory of the homogenous flow A_s on the space $E = G(\mathbb{R})/G(\mathbb{Z})$: more precisely, we will exploit the fact that the circle $\{U(t) : t \in [0,1]\}$ is a *non-linear horocycle* (a concept to be introduced later) to derive Theorem 3.3.3 from a more general result about the equidistribution of the non-linear horocycles of the geodesic flow A_s.

A detailed explanation of this plan is the content of the next section.

3.4 Equidistribution of non-linear horocycles

During this section, we will discuss the Ergodic Theory of the homogenous flow A_s in the space of affine lattices $G(\mathbb{R})/G(\mathbb{Z})$. To do so, we recall some definitions of the last subsection. As we saw in the last section, the special affine group $ASL_2(\mathbb{R})$ is naturally identified with the following subgroup of $SL_3(\mathbb{R})$:

$$G(\mathbb{R}) := \left\{ \begin{pmatrix} a & b & x \\ c & d & y \\ 0 & 0 & 1 \end{pmatrix} : ad - bc = 1 \right\}$$

which is the semi-direct product $G(\mathbb{R}) = SL_2(\mathbb{R}) \ltimes V_2(\mathbb{R})$ where

$$SL_2(\mathbb{R}) \simeq \left\{ \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\} \text{ and } V_2(\mathbb{R}) = \left\{ \begin{pmatrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \right\} \simeq \mathbb{R}^2.$$

Moreover, we identify the space of affine lattices E with $G(\mathbb{R})/G(\mathbb{Z})$ and we define

$$A_s := \begin{pmatrix} s & 0 & 0 \\ 0 & 1/s & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } U(t) := \begin{pmatrix} 1 & -2t & -t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.1}$$

Finally, we concluded that these identifications reduce our task to the proof of the following result:

Theorem 3.4.1. *For all $f \in C_0(E)$ it holds*

$$\int_0^1 f(A_s \cdot U(t))dt \to \int_E f d\mu_E.$$

As we already told, this theorem will be a consequence of a more general result about the equidistribution of non-linear horocycles. In order to state precisely this general result, let us introduce the following definition:

Definition 3.4.1. *A horocyclic section (or horocycle) is a map $\sigma : \mathbb{R} \to G(\mathbb{R})$ of the form*

$$\sigma(t) = \begin{pmatrix} 1 & t & x(t) \\ 0 & 1 & y(t) \\ 0 & 0 & 1 \end{pmatrix} \tag{3.2}$$

such that

$$\sigma(t + p_0) = \sigma(t)\gamma_0$$

for some integer $p_0 \geq 1$ and some element $\gamma_0 \in G(\mathbb{Z})$.

Remark 3.4.1. *Given a horocycle σ, there exists a* minimal *integer $p \geq 1$ such that $\sigma(t+p) = \sigma(t)\gamma$ for some $\gamma \in G(\mathbb{Z})$. This integer p is the* period *of σ in $E = G(\mathbb{R})/G(\mathbb{Z})$.*

Remark 3.4.2. *The name* horocycle *is motivated by the fact that the natural projection of the space of affine lattices E to the space of lattices B sends a horocyclic section of E to a (usual) horocycle around a cusp of B.*

Definition 3.4.2. *A horocycle σ is called* linear *(over the rationals $\mathbb{Q}$) whenever for all $\alpha, \beta \in \mathbb{Q}$ it holds*

$$m\left(\{t \in [0,p] : x(t) = \alpha t + \beta\}\right) > 0,$$

where m stands for the Lebesgue measure. Otherwise, the horocycle σ is called non-linear*.*

Remark 3.4.3. *The behavior of $y(t)$ doesn't have any influence in the definition of linear horocycles.*

Remark 3.4.4. *A real-analytic horocycle σ is linear if and only if $x(t) \equiv \alpha t + \beta$ for some $\alpha, \beta \in \mathbb{Q}$ (since any non-constant real-analytic function has a discrete set of zeroes).*

Comparing the equation (3.1) and (3.2) and using the remark 3.4.4, we see that

$$\sigma(t) := U(-t/2) := \begin{pmatrix} 1 & t & -t^2/4 \\ 0 & 1 & -t/2 \\ 0 & 0 & 1 \end{pmatrix}$$

is a non-linear horocycle with period $p = 2$ and $x(t) = -t^2/4$. Therefore, the theorem 3.4.1 is an immediate consequence of the following more general fact:

Theorem 3.4.2 (Equidistribution of non-linear horocycles)**.** *Let $\sigma : \mathbb{R} \to G(\mathbb{R})$ be a non-linear horocycle of period p. Then, the circle $A_s \cdot \sigma$ become equidistributed in E, i.e.,*

$$\lim_{s\to\infty} \frac{1}{p} \int_0^p f(A_s \cdot \sigma(t))dt = \int_E f(x) d\mu_E(x).$$

Remark 3.4.5. *The main ingredients in this result are: the "linear part" of the horocycle is an unipotent matrix and the horocycle is non-linear. Indeed, during the proof of Theorem 3.4.2, we will use the fact that the linear part of the horocycle is unipotent to apply Ratner's theorem in order to reduce the list of candidates to the distribution law μ of the horocycle to a countable quantity of possibilities (among them μ_E). Then, we will use the non-linearity to exclude all exotic possibilities.*

Remark 3.4.6. *The assumption of non-linearity of the horocycle is essential: when it is linear, the conclusion of Theorem 3.4.2 is simply false! We will come back to this point after the proof of this theorem.*

After this considerations, we will dedicate the rest of this last section of the last chapter of this book to the proof of Theorem 3.4.2. To do so, we will use the following scheme:

- during the next subsection, we will revise some basic facts about invariant measure and we will see some properties of the probability measure μ associated to the distribution law of $A_s \cdot \sigma(t)$;

- in the sequel, we will use Ratner theorem to show that there are only a countable quantity of possibilities to the distribution law μ;

- finally, in the last subsection, we will use the non-linearity of the horocycle σ to show that $\mu = \mu_E$.

Now we start to formalize this program.

3.4.1 The distribution law of a loop

Given a loop $\sigma : \mathbb{R}/p\mathbb{Z} \to E$, we denote by $m(\sigma)$ the natural probability measure supported on the image of σ:

$$\int_E f dm(\sigma) := \frac{1}{p}\int_0^p f(\sigma(t))dt$$

for any $f \in C_0(E)$.

Furthermore, given $\sigma : \mathbb{R}/p\mathbb{Z} \to E$ a non-linear horocycle of period p, we denote by $\sigma_s := A_s \cdot \sigma$, so that Theorem 3.4.2 is equivalent to:

Theorem 3.4.3 (Equidistribution of non-linear horocycles, 2nd version)**.** *For any non-linear horocycle σ it holds*

$$m(\sigma_s) = (A_s)_* m(\sigma) \to \mu_E$$

when $s \to \infty$.

Here the convergence means weak-* convergence. By the Banach-Alaoglu theorem, we know that $m(\sigma_s)$ possesses a subsequence converging to a measure μ. In particular, our task consists into showing that $\mu = \mu_E$ is the unique possible limit of all convergent subsequences.

In this direction, consider the "derivative" map D from the space of affine lattices E to the space of lattices B assigning to each element $g \in E$ its linear part $D(g) \in B$, i.e.,

$$D\begin{pmatrix} a & b & x \\ c & d & y \\ 0 & 0 & 1 \end{pmatrix} := \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Observe that, *a priori*, the projection of the Haar probability measure μ_E of E by D isn't necessarily equal to the Haar probability of μ_B of B. Thus, as a preliminary work in the direction of the proof of Theorem 3.4.3, let us verify that the projection of μ by D is correct:

Proposition 3.4.1. *We have $D_*\mu = \mu_B$.*

Proof. The image H of $D \circ \sigma$ is a horocycle (in the usual sense) of the space B. On the other hand, D sends the orbits of the "Teichmuller geodesic flow" A_s of E to the geodesics of B and D sends the measure $m(\sigma)$ to the Haar measure μ_H of H. Finally, a simple argument shows that the geodesic flow of B pushes H far from the cusps of B so that H becomes equidistributed (for further details see Theorem 2.4 of Elkies and McMullen paper). Putting these facts together, it follows that

$$D_*\mu = \lim(A_s)_*\mu_H = \mu_B.$$

This concludes the proof. □

Remark 3.4.7. *A direct consequence of Proposition 3.4.1 is the fact that μ is a* probability *measure on E, i.e., $\mu(E) = 1$. In particular, the mass of the probabilities $m(\sigma_s)$ is* conserved *at the limit. This is a non-trivial remark because E is* non-compact *(so that the mass could escape to infinity* a priori*)!*

As we are going to see later, in order to fit the context of Ratner theorem, we need to know that μ is invariant by an unipotent subgroup of $SL_2(\mathbb{R})$. In this direction, we introduce the subgroup

$$N(t) := \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that this unipotent subgroup appears naturally in view of the formula $D \circ \sigma(t) = N(t)$ whenever $\sigma(t)$ is a horocycle. The following preparatory result will put us in the setting of Ratner theorem:

Proposition 3.4.2. *The probability measure μ is $N(\mathbb{R})$-invariant.*

Proof. Fix $\tau \in \mathbb{R}$. Consider $\sigma_s(t) = A_s \cdot \sigma(t)$ and $\eta_s(t) = N_\tau \cdot \sigma_s(t)$ where $\sigma(t)$ is a horocycle. We have that

$$\sigma_s(t) = \begin{pmatrix} s & st & sx(t) \\ 0 & \frac{1}{s} & \frac{y(t)}{s} \\ 0 & 0 & 1 \end{pmatrix}, \eta_s(t) = \begin{pmatrix} s & st + \frac{\tau}{s} & sx(t) + \frac{\tau y(t)}{s} \\ 0 & \frac{1}{s} & \frac{y(t)}{s} \\ 0 & 0 & 1 \end{pmatrix}.$$

In order to compare $\sigma_s(t)$ and $\eta_s(t)$, we perform a change of variables to get the same linear parts. More precisely, we define $u = \tau/s^2$ and we consider

$$\rho_s(t) := \eta_s(t-u) := \begin{pmatrix} s & st & sx(t-u) + s^{-1}\tau y(t-u) \\ 0 & 1/s & y(t-u) \\ 0 & 0 & 1 \end{pmatrix}.$$

Recall that $m(\sigma_s) \to \mu$, so that

$$m(\rho_s) = m(\eta_s) = (N_\tau)_* m(\sigma_s) \to (N_\tau)_* \mu. \tag{3.3}$$

On the other hand, we have that $D \circ \rho_s = D \circ \sigma_s$, so that the distance between ρ_s and σ_s is given by the distance between the vectors obtained from the third column of the matrices:

$$d(\rho_s, \sigma_s) = \left| \begin{pmatrix} sx(t-u) + \tau y(t-u)/s \\ y(t-u)/s \\ 1 \end{pmatrix} - \begin{pmatrix} sx(t) + \tau y(t)/s \\ y(t)/s \\ 1 \end{pmatrix} \right|$$

Next, we use the fact that $x(t)$ is Lipschitz, $y(t)$ is bounded and $u = \tau/s^2$ to get

$$|sx(t) - sx(t-u)| \leq s|x(t) - x(t-u)| \leq O(su) = O(1/s)$$

and

$$|y(t)/s - y(t-u)/s| \leq (|y(t)| + |y(t-u)|)/s = O(1/s).$$

Thus, we see that $d(\rho_s, \sigma_s) \to 0$ when $s \to \infty$. In particular, it follows that $\lim m(\rho_s) = \lim m(\sigma_s) = \mu$. Putting this information together with (3.3), we get

$$(N_\tau)_* \mu = \mu$$

so that the proof is complete. □

Once we know that μ is invariant by the unipotent subgroup $N(\mathbb{R})$, we move to the discussion of Ratner theorem.

3.4.2 Ratner theorem and classification of μ

Ratner theorem can be stated as follows:

Theorem 3.4.4. *Let Γ be a discrete subgroup of a connected Lie group G and N be an unipotent subgroup. Let ν be a N-invariant ergodic probability on G/Γ and denote by J the biggest subgroup of G leaving ν invariant. Then, there exists $x \in G/\Gamma$ such that $\nu(J \cdot x) = 1$. Furthermore, ν is the Haar measure of $J \cdot x$ and its support is $J \cdot x$ (so that $J \cdot x$ is closed in G/Γ).*

The relevance of Ratner's theorem in the context of Elkies and McMullen is evident: since μ is invariant by the unipotent subgroup N, we can *classify* μ by listing all closed subgroups of E.

Obviously Ratner's theorem has a beautiful history including several applications to several areas of Mathematics. In particular, it is nearly impossible to describe its importance in a brief discussion, so that we recommend the interested reader the nice exposition of D. Morris [11] (and also Terence Tao's blog "What's new" http://terrytao.wordpress.com/ for a series of blog posts about Ratner's theorems).

In any case, we are going to use Ratner theorem as follows. Denoting by F a fiber of $E \to B$, we observe that F is a complex torus $\mathbb{C}/\Lambda$. For each integer $n \geq 1$ we define $F[n] = \left(\frac{1}{n} \cdot \Lambda\right)/\Lambda \subset F$ the torsion points of order n with respect to the group structure of F and we denote by $E[n]$ the subbundle of E whose fibers are $F[n]$.

Definition 3.4.3. $\bigcup E[n]$ *is the subset of torsion points of E.*

Next we introduce $H(\mathbb{R}) \subset G$ the subgroup of horizontal translations, i.e., translations by vectors of the form $(x, 0) \in \mathbb{R}^2$ and $H(r, \varepsilon) \subset G$ the subset of translations by vectors (x, y) of the form $|x| < r$ and $|y| < \varepsilon$.

The goal of this subsection is the application of Ratner theorem to show the following classification result:

Theorem 3.4.5 (Classification of μ)**.** *We have that $\mu = \mu_E$ or $\mu(H(\mathbb{R}) \cdot E[n]) > 0$ for some $n \geq 1$.*

Unfortunately, this result *isn't* an immediate consequence of Ratner theorem because we don't know that μ is ergodic. However, this is not a big deal since we can use the ergodic decomposition theorem to write μ as a ("unique") convex combination of $N(\mathbb{R})$-invariant ergodic measures:

$$\mu = \int \nu dP(\nu).$$

Remark 3.4.8. *Usually the ergodic decomposition theorem concerns compact spaces. In the specific case of E (a non-compact space), we apply the ergodic decomposition theorem to the one-point compactification of E and we restrict it to E.*

Now, for each $N(\mathbb{R})$-invariant ergodic probability ν on E, we define

$$J(\nu) := \{g \in G(\mathbb{R}) : g_* \nu = \nu\},$$

that is, $J(\nu)$ is the biggest subgroup of $G(\mathbb{R})$ leaving ν invariant. Observe that $J(\nu)$ is closed and $N(\mathbb{R}) \subset J(\nu)$.

Proposition 3.4.3. *For almost every ν of the ergodic decomposition of μ, we have*

$$D_* \nu = \mu_B \quad \text{and} \quad D(J(\nu)) = SL_2(\mathbb{R}).$$

Proof. From Proposition 3.4.1, we know that $\mu_B = D_* \mu = \int D_* \nu dP(\nu)$. Since the action of $N(\mathbb{R})$ on (B, μ_B) is ergodic (because it is the action of the horocyclic flow on B), it follows that $D_* \nu = \mu_B$ for almost every ν.

On the other hand, by Ratner theorem, we know that ν is supported on a closed orbit $J(\nu) \cdot x \subset E$. Thus,

$$D(J(\nu)) \cdot D(x) = D(J(\nu) \cdot x) = D(\text{supp}(\nu)) = \text{supp}(D_* \nu).$$

Since $D_* \nu = \mu_B$, we obtain

$$D(J(\nu)) \cdot D(x) = \text{supp}(\mu_B) = B = SL_2(\mathbb{R})/SL_2(\mathbb{Z}).$$

Therefore, $D(J(\nu)) = SL_2(\mathbb{R})$. This completes the proof. □

Now we recall the following proposition about $SL_2(\mathbb{R})$-actions:

Proposition 3.4.4. *Every affine action of $SL_2(\mathbb{R})$ on $\mathbb{R}^k$ has fixed points.*

Proof. By Weyl's unitary trick, this action can be extended to a $SL_2(\mathbb{C})$-action on $\mathbb{C}^k$. On the other hand, a fixed point $p \in \mathbb{C}^k$ of the compact subgroup $SU_2(\mathbb{C})$ can be easily constructed (e.g., by averaging). Since $\mathbb{C} \cdot su_2(\mathbb{C}) = sl_2(\mathbb{C})$, the point p is also fixed by the $SL_2(\mathbb{C})$ and, *a fortiori*, by the $SL_2(\mathbb{R})$. Hence, the real part of p is a fixed point of $SL_2(\mathbb{R})$ on $\mathbb{R}^k$. □

Proposition 3.4.5. *If $H \subset G(\mathbb{R})$ is a subgroup such that $D(H) = SL_2(\mathbb{R})$, then $H = G(\mathbb{R})$ or H is conjugated to $SL_2(\mathbb{R})$.*

Proof. Since $D(H) = SL_2(\mathbb{R})$, the kernel K of the derivative map $D : H \to SL_2(\mathbb{R})$ is a $SL_2(\mathbb{R})$-invariant subgroup of $V_2(\mathbb{R}) \simeq \mathbb{R}^2$, so that one of the following two possibilities occurs:

- $K = V_2(\mathbb{R})$: in this case, $H = G(\mathbb{R})$;
- $K = \{e\}$: in this case, we have an affine action $D^{-1} : SL_2(\mathbb{R}) \to H \subset G(\mathbb{R}) = ASL_2(\mathbb{R})$ of $SL_2(\mathbb{R})$ on $\mathbb{R}^2$ which has fixed points by Proposition 3.4.4; up to conjugation with an appropriate element of $V_2(\mathbb{R})$, we can assume that this fixed point is the origin and $H = SL_2(\mathbb{R})$.

This ends the proof. □

Corollary 3.4.1. *$J(\nu) = G(\mathbb{R})$ or $J(\nu) = g \cdot SL_2(\mathbb{R}) \cdot g^{-1}$ for some horizontal translation $g \in H(\mathbb{R})$.*

Proof. Since ν is $N(\mathbb{R})$-invariant, we know that $N(\mathbb{R}) \subset J(\nu)$. Moreover, by Proposition 3.4.3, we have that $D(J(\nu)) = SL_2(\mathbb{R})$. Hence, using Proposition 3.4.5, it follows that $J(\nu) = G(\mathbb{R})$ or $J(\nu) = g \cdot SL_2(\mathbb{R}) \cdot g^{-1}$. This concludes the argument. □

Proposition 3.4.6. *$\nu = \mu_E$ or $supp(\nu) \subset g \cdot E[n]$ for some integer $n \geq 1$ and $g \in H(\mathbb{R})$.*

Proof. From the previous corollary, we have that $J(\nu) = G(\mathbb{R})$ or $g \cdot SL_2(\mathbb{R}) \cdot g^{-1}$. In the first case, we see that $\nu = \mu_E$ by $J(\nu)$-invariance of ν. In the second case, $g^{-1}\text{supp}(\nu) = SL_2(\mathbb{R}) \cdot x$ is a closed $SL_2(\mathbb{R})$-orbit in E. Since such orbits are always contained in $E[n]$ for some $n \geq 1$, this ends the proof. □

At this stage, we can conclude this subsection with the proof of Theorem 3.4.5:

Proof. We write the ergodic decomposition of μ as $\mu = \int \nu dP(\nu)$. By the proposition 3.4.6, almost every ergodic component ν of μ satisfy $\nu = \mu_E$ or $\mathrm{supp}(\nu) \subset H(\mathbb{R}) \cdot E[n]$ for some n. Hence, we can write μ as:

$$\mu = a_0 \mu_E + \sum_{n=1}^{\infty} a_n \mu_n,$$

where $\sum_{n=0}^{\infty} a_n = 1$ and $\mathrm{supp}(\mu_n) \subset H(\mathbb{R}) \cdot E[n]$. In particular, if $\mu \neq \mu_E$ then $a_n \neq 0$ for some $n \geq 1$, so that $\mu(H(\mathbb{R}) \cdot E[n]) > 0$. This completes the proof of the theorem. □

In view of the classification of μ provided by Theorem 3.4.5, we see that Theorem 3.4.3 (about the equidistribution of non-linear horocycles) follows once we show that μ doesn't see the torsion points of E. This is the topic of the next subsection.

3.4.3 Non-linearity and torsion points

The main theorem of this subsection is

Theorem 3.4.6. *Given σ a non-linear horocycle and μ an accumulation point of the sequence of probability measures $m(A_s \cdot \sigma)$ (when $s \to \infty$), we have*

$$\mu(H(\mathbb{R}) \cdot E[n]) = 0$$

for every $n \geq 1$.

Proof. Given $\varepsilon > 0$ and $r > 0$, define

$$U = H(r, \varepsilon) \cdot E[n]$$

and

$$T_s = \{t \in [0, p] : \sigma_s(t) \in U\}.$$

We claim that

$$\limsup_{s \to \infty} m(T_s) = O(\varepsilon), \tag{3.4}$$

where m is the Lebesgue measure. In order to compute $m(T_s)$, it is convenient to pass to the universal cover $G = G(\mathbb{R})$ of $E = G/G(\mathbb{Z})$. We start with the observation that $E[n]$ is covered by the $SL_2(\mathbb{R})$-orbit of $G[n] = \bigcup G[n]^{i,j}$ where

$$G[n]^{i,j} = \left\{ \begin{pmatrix} a & b & \frac{i}{n}a + \frac{j}{n}b \\ c & d & \frac{i}{n}c + \frac{j}{n}d \\ 0 & 0 & 1 \end{pmatrix} : ad - bc = 1 \right\}.$$

In particular, the points of $G[n]$ belonging to the same fiber of $\sigma_s(t)$ are

$$\rho_s^{i,j}(t) = \left(\begin{array}{ccc} s & st & \frac{i}{n}s + \frac{j}{n}st \\ 0 & s^{-1} & \frac{j}{n}s^{-1} \\ 0 & 0 & 1 \end{array} \right).$$

Taking the Euclidean metric on the third column of the matrices above, we see that $T_s = \bigcup T_s^{i,j}$ where

$$T_s^{i,j} = \left\{ t : \left(\begin{array}{c} sx(t) \\ s^{-1}y(t) \end{array} \right) - \left(\begin{array}{c} \frac{i}{n}s + \frac{j}{n}st \\ s^{-1}\frac{j}{n} \end{array} \right) \in H(r, \varepsilon) \right\}.$$

In particular, $T_s^{i,j} \subset X_s^{i,j} \cap Y_s^{i,j}$ where

$$X_s^{i,j} = \{t : |x(t) - \frac{i}{n} - \frac{j}{n}t| < r/s\}$$

and

$$Y_s^{i,j} = \{t : |y(t) - \frac{j}{n}| < \varepsilon s\}.$$

At this point, we will use the non-linearity of σ to get that the subset of t with $x(t) = \frac{i}{n} + \frac{j}{n}t$ has zero Lebesgue measure, so that, for each i, j fixed, we have

$$\lim_{s \to \infty} m(X_s^{i,j}) = 0 \tag{3.5}$$

On the other hand, we can use that $x(t)$ is Lipschitz to estimate $m(X_s^{i,j})$ when j is large: more precisely, whenever $|j| > M := 2n \sup\limits_{0 \leq t \leq p} |x'(t)|$, the subset $X_s^{i,j}$ is the pre-image of an interval of size $1/s$ by a map whose derivative has order j/n. Hence,

$$m(X_s^{i,j}) = O(1/s|j|) \quad \text{for all} \quad |j| > M. \tag{3.6}$$

Moreover, we note that

$$Y_s^{i,j} = \emptyset \quad \text{when} \quad |j| \geq J_s := n(s\varepsilon + \sup_{0 \leq t \leq p} |y(t)|). \tag{3.7}$$

and

$$X_s^{i,j} = \emptyset \quad \text{when} \quad |i| \geq I_s(j) := n(\frac{r}{s} + |\frac{j}{n}| + \sup_{0 \leq t \leq p} |x(t)|). \tag{3.8}$$

Finally, we observe that

$$J_s = O(s\varepsilon) \quad \text{and} \quad I_s(j) = O(|j| + 1) \quad \text{for} \quad s \quad \text{large}. \tag{3.9}$$

Keeping these facts in mind, we can estimate $m(T_s)$: by (3.7) and (3.8) it follows that

$$m(T_s) \leq \sum_{|j| < J_s} \sum_{|i| < I_s(j)} m(X_s^{i,j}) \tag{3.10}$$

Now we split the right-hand sum into two parts:

$$\sum_{|j|<J_s}\sum_{|i|<I_s(j)} m(X_s^{i,j}) = \sum_{M<|j|<J_s}\sum_{|i|<I_s(j)} m(X_s^{i,j}) + \sum_{|j|\leq M}\sum_{|i|<I_s(j)} m(X_s^{i,j})$$

Next, we note that the equation (3.6) implies that the first part of this sum is $O(|J_s|\varepsilon/s) = O(\varepsilon^2)$ (because (3.9) says that $|I_s| = O(|j|+1)$ and $|J_s| = O(s\varepsilon)$) and the second part of this sum occurs over a finite set of indices i, j so that (3.5) says that it tends to zero (when s increases). Thus, putting these two estimates together with (3.10), we see that, for large s, it holds

$$m(T_s) = O(\varepsilon),$$

so that the desired estimate (3.4) follows.

Finally, we recall that $m_s(U) = m(T_s)/p$, so that the estimate (3.4) implies $\mu(H(r,\varepsilon)\cdot E[n]) = O(\varepsilon)$ for all $r, \varepsilon > 0$. Making $\varepsilon \to 0$ first and $r \to \infty$ after, it follows that $\mu(H(\mathbb{R})\cdot E[n]) = 0$, so that the desired theorem is proved. □

Once we proved Theorem 3.4.6, it is an easy task to conclude the proof of Theorem 3.4.3 (or equivalently, Theorem 3.4.2). In fact, this is the content of the next (short) final subsection below.

3.4.4 End of the proof of Theorem 3.4.3

Given σ a non-linear horocycle, consider any accumulation point μ of $m(A_s\cdot\sigma)$ when $s\to\infty$. By Theorem 3.4.6, μ gives zero mass to the horizontal translations of the torsion points $\bigcup\limits_{n\geq 1} E[n]$ of E. Hence, the classification theorem (Theorem 3.4.5) implies that $\mu = \mu_E$. In other words, we have that μ_E is the unique accumulation point of the sequence $m(A_s\cdot\sigma)$. This shows that

$$m(A_s\cdot\sigma)\to\mu_E$$

so that the proof of Theorem 3.4.3 is complete.

At this point, our exposition of the proof of Elkies and McMullen is finished! Closing the last section of the final chapter of this book, we make the following remark:

Remark 3.4.9. *The equidistribution theorem (Theorem 3.4.3) is* optimal, *i.e., it* never *holds for* linear *horocycles σ: if $x(t) = \frac{i}{n} + \frac{j}{n}t$ for a positive measure subset of t then $\mu(E[n]) > 0$ so that $m(A_s\cdot\sigma)$ can't converge to μ_E.*

Bibliography

[1] A. Arbieto, C. Matheus and C.G. Moreira, Aspectos Ergódicos da Teoria dos Números. *Pub. Mat. IMPA.*, 26.^{o} *Colóquio Brasileiro de Matemática*, 2007.

[2] V. Brun, La serie $1/5 + 1/7 + 1/11 + 1/13 + 1/17 + 1/19 + 1/29 + 1/31 + 1/41 + 1/43 + 1/59 + 1/61 + ...$, les denominateurs sont nombres premiers jumeaux est convergente ou finie. *Bull. Sci. Math. 43*, 124-128, 1919.

[3] J. Chen, On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica 16*, 157-176, 1973.

[4] M. Elkin, An improved construction of progression-free sets, Preprint 2008, http://arxiv.org/abs/0801.4310.

[5] W. Feller, Introduction to Probability Theory and its applications. Vol. I, 3rd Edition, *Wiley, New York*. 1968.

[6] N. Elkies and C. McMullen, Gaps in $\sqrt{n}$ mod 1 and Ergodic Theory. *Duke Math. Journal 123*, 95-139, 2004.

[7] A. Goldston, J. Pintz and Y. Yildirim, Primes in Tuples I. *Annals of Math. 170*, 819-862, 2009.

[8] A. Goldston, J. Pintz and Y. Yildirim, Primes in Tuples II. Preprint 2007. http://arxiv.org/abs/0710.2728.

[9] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions. *Annals of Math. 167*, 481–547, 2008.

[10] B. Green and J. Wolf, A note on Elkin's improvement of Behrend's construction, Preprint 2008, http://arxiv.org/abs/0810.0732.

[11] D. Morris, Ratner's Theorems on Unipotent Flows. *Chicago Lectures in Mathematics Series*, University of Chicago Press. 2005.

[12] K. Roth, On certain sets of integers. *J. London Math. Soc. 28*, 245-252, 1953.

[13] L. G. Schnirelman, *Uspekhi Math. Nauk 6*, 3-8, 1939.

[14] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression. *Acta Arith. 27*, 299-345, 1975.

[15] T. Tao, Arithmetic Progressions and the Primes (El Escorial Lectures), *Collect. Math.*, 37-88, 2006.

[16] T. Tao, The ergodic and combinatorial approaches to Szemerédi's theorem. *CRM Proc. Lecture Notes 43*, 145-193, 2007.

[17] T. Tao, The Gaussian primes contain arbitrarily shaped constellations. *J. d'Analyse Mathematique 99*, 109-176, 2006.

[18] T. Tao and T. Ziegler, The primes contain arbitrarily long polynomial progressions. *Acta Math. 201*, 213-305, 2008.

[19] J. Van der Corput, Uber Summen von Primzahlen und Primzahlquadraten. *Math. Ann. 116*, 1-50, 1939.

[20] I. Vinogradov, Representation of an odd prime as a sum of three primes. *Comptes Rendus (Doklady) de l'Académie des Sciences de PURSS 15,* 291-294, 1937.

Alexander ARBIETO
UFRJ, Universidade Federal do Rio de Janeiro
Av. Athos da Silveira Ramos, 149, Ilha do Fundão, CEP 68530
Rio de Janeiro, RJ, Brazil
alexande@impa.br
http://www.im.ufrj.br/∼arbieto

Carlos MATHEUS
Collège de France
3, Rue d'Ulm, CEDEX 05
Paris, France
matheus@impa.br
http://www.impa.br/∼cmateus

Carlos Gustavo (Gugu) MOREIRA
IMPA, Instituto de Matemática Pura e Aplicada
Estrada D. Castorina, 110, CEP 22.460-320
Rio de Janeiro, RJ, Brazil
gugu@impa.br
http://www.impa.br/∼gugu