Digits & deviations of dynamical systems

Imbierski, J.F.

Digits & Deviations of Dynamical Systems

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op vrijdag 4 juli 2025
klokke 11:30 uur

door

Jonny Imbierski

geboren te Reading, Engeland
in 1998

Promotor:

Prof. Dr. W.T.F. den Hollander

Co-promotores:

Dr. C.C.C.J. Kalle

Dr. D. Terhesiu

Promotiecommissie:

Prof. Dr. Ir. G.L.A. Derks

Prof. Dr. F.M. Spieksma

Dr. K. Dajani (Utrecht Universiteit)

Prof. Dr. L. Liao (University of Wuhan)

Prof. Dr. M. Todd (University of St Andrews)

The front cover artwork was designed by Ryan Imbierski.

Contents

Chapter 1

Setting the Scene

1.1 Motivation

A *discrete-time dynamical system* is an abstract space along with a collection of rules that describes how the space can evolve from one discrete time-step to the next. The abstract space is called the *state space* of the system as it gives the different states that the system may be in, and the evolution of the state space with respect to the collection of rules is referred to as the *dynamics* of the system. In this thesis, we will quantify the unlikeliness of probabilistically unlikely events defined via certain families of discrete-time dynamical systems. There are several ways in which the unlikeliness of a dynamically defined event may be quantified. The most natural of course is to use a probability measure to assign a probability that the event occurs. We will also see that other notions may be used to deduce more information in the situation when the probability is zero. In this case, the events that we consider are examples of *fractals*, which are sets often characterised by having a geometrically complicated boundary and repeating structure. We will use the *Hausdorff dimension* – a function which quantifies the geometric complexity of a fractal boundary – to distinguish them.

This thesis is split into two parts each consisting of two chapters.

Part I is devoted to studying 'Birkhoff average level sets' for different dynamical systems. A *Birkhoff average* is the average over time of some fixed potential along the trajectory of a dynamical system and is one of the most critical quantities appearing in *ergodic theory*: the study of statistical properties of dynamical systems expressed through the behaviour of such averages. The Birkhoff average *level sets* are the elements of a partition of a dynamical system's state space into subsets determined by the different values taken by Birkhoff averages evaluated at each state. In other words, there is a one-to-one correspondence between each level set and each distinct Birkhoff average. Given the importance of Birkhoff averages to ergodic theory, it is a relevant question to ask how large the size of the level sets of Birkhoff averages can be. For many dynamical systems, including *ergodic* dynamical systems (see §1.3.4) and the systems that we will consider throughout Part I, one level set has full probability and the rest therefore have zero probability so the size of Birkhoff average level sets may be further quantified using the Hausdorff dimension. The Hausdorff dimension of sets of Birkhoff averages have been studied in various contexts; see e.g. [PW01; BSS02a; JS07; BM08; FLMW10; KR14; IJ15; BJKR21; Jur21; JR21; Rus23].

In Chapter 2, the Birkhoff averages will be given with respect to a type of dynamical system called an *iterated function system* where the dynamics are given by a collection of self-maps on the same state space that do not increase distances between states. The particular iterated function systems that we consider will be finite collections of *affine* maps of the unit square, meaning that each map may be written as a linear transformation composed with a translation. In [BJKR21], the authors consider iterated function systems of affine maps under certain 'irreducibility conditions' on the linear parts and obtain a formula for the Hausdorff dimension of the level sets of Birkhoff averages corresponding to these systems for any continuous potential. In the first half of Chapter 2, we will obtain the same formula in this setting but with the irreducibility conditions replaced with one of two alternate conditions that each concern how rarely iterates of the iterated function system overlap. This setting is a subcase of the main result in [Ree11]; however, given the extra conditions on our system, we will be able to obtain more information than is given there.

For the rest of Chapter 2 and then in Chapter 3, we consider dynamical systems called *number systems* that are used to generate and assign number expansions, i.e. symbolic representations, to each of the numbers in a state space. Using these number systems, we may define a family of events by considering the different frequencies of occurrence of symbols, called *digits*, that can appear in the number expansions of the numbers in the state space. These events are types of sets called *Besicovitch-Eggleston* sets due to the pioneering works of Besicovitch and Eggleston; in [Bes35], Besicovitch studied such events defined using binary (base-2) expansions, and Eggleston continued this study in [Egg49] to events defined via any other integer base expansion such as the decimal (base-10) expansion. We will see in §1.3.4 that the frequency of occurrence of a digit in a number expansion may be written as a Birkhoff average with respect to a certain potential so Besicovitch-Eggleston sets are particular level sets of Birkhoff averages; hence, we will use the Hausdorff dimension to study them further. There is a plethora of literature concerning the Hausdorff dimension of sets determined by digit properties of number expansions; see e.g. [Bes35; Egg49; PW99; BSS02a; BI09; FLMW10; Mun11; FJLR15; NT24]. Most relevant to this thesis is [FLMW10, Theorem 1.2], where the authors obtain the Hausdorff dimension of Besicovitch-Eggleston sets defined for dynamical systems generating symbolic representations that draw from a countably infinite pool of symbols. As an application, they obtain [FLMW10, Theorem 1.1] by transferring the former result on symbolic dynamical systems to number systems given by maps on the unit interval that are piecewise linear and piecewise bijective – the so-called *Generalised Lüroth Series* (GLS) maps introduced in [BBDK96]. GLS maps will be particularly relevant throughout Part I.

An important feature of the number systems considered in Part I is that they are 'non-autonomous'. A dynamical system is called *autonomous* if its dynamics are given in a time-independent manner, meaning that the rules used to evolve the system are the same at any point in time, and is called *non-autonomous* otherwise. There are two types of autonomous dynamical systems: deterministic and random. A dynamical system is *deterministic* if every state of the system has exactly one possible evolution at each time-step so that there is no uncertainty as to how the system will evolve over time. In contrast, a *random* dynamical system is an autonomous but non-deterministic dynamical system that consists of two parts: an autonomous random process and a non-autonomous dynamical system driven by the random process. The dynamical systems that we will consider in the latter half of Chapter 2 and in Chapter 3 will be non-autonomous so the dynamics will be time-dependent, meaning that the rules used to evolve the system may change from one time-step to the next. Non-deterministic dynamical systems are important to consider because they better capture the behaviour of real-world phenomena, which are rarely governed by time-independent processes. There are many results on number expansions generated by non-deterministic dynamical systems, e.g. [DK03; DV05; DK07; Kem14; KKV17; DO18; DK20; KM22b; NT24; GKKS25; BDKK+24]. The latter half of Chapter 2 is devoted to proving a result similar to [FLMW10, Theorem 1.1] (where the authors treat autonomous and deterministic number systems driven by GLS maps) but for non-autonomous number systems constructed from GLS maps whose associated number expansions draw from a finite pool of digits. We generalise this in Chapter 3 by allowing the pool of digits to be countably infinite under an additional minor assumption on the system. The result obtained in Chapter 3 contains the autonomous deterministic case considered in [FLMW10, Theorem 1.1] as a subcase.

The results in Part I will require a blend of techniques from measure theory, ergodic theory, fractal geometry and symbolic dynamics, the foundations of which will all be introduced in §1.3.

In Part II, we consider a broad class of abstract deterministic dynamical systems that includes dynamical systems of great interest such as *Gibbs-Markov maps* and *AFU maps*; see §1.2. In particular, this includes some of the GLS number systems used in Part I to construct the non-autonomous systems considered there. We will use this broad class of dynamical systems to define random variables and then analyse the probability that the value of the (suitably centred and scaled) partial sums of these random variables falls within an unlikely range. The study of such behaviour for partial sums of random variables is known as *large deviations theory* and in a classical setting defines an unlikely range to be that which is outside of what is dictated by the classical *Central Limit Theorem*, i.e. the statement that the distributions of these partial sums converge to a normal distribution under conditions on the mean and variance of the random variables. We will consider *stable large deviations*, which defines an unlikely range to be that which is outside of what is dictated by a *stable limit theorem* – a generalisation of the classical central limit theorem which states that the distributions of partial sums converge under certain conditions to an *α-stable distribution*, where $\alpha \in (0, 2]$ is called the *stability parameter* of the stable distribution. We give further details concerning stable distributions and stable large deviations in §1.4. The complete statement for stable large deviations of independent and identically distributed (i.i.d.) random variables is a culmination of the results in [Hey67; Nag69; Nag79; Roz89; Ber19], which we will recall in Theorem 1.8. In the dependent stable setting, there are only a few large deviations results; see e.g. [Gan96; MW13; Tak19] and the brief discussion of these in §4.2. For references concerning classical (exponential) large deviations and different applications, see [Hol00] and the references therein.

Throughout Part II, we will be interested in large deviations in the setup of α-stable distributions when the classical CLT does not hold. In particular, we will present a new large deviations result in the context of α-stable dependent random variables with $\alpha \in (0, 1) \cup (1, 2]$ (with infinite variance) in Chapter 4 by first reproving the i.i.d. case in a way that better caters for dynamically defined random variables and then using that and the assumptions on the class of dynamical systems to deduce the dependent case. This result holds for the majority of the range of large deviations given in Theorem 1.8 and for all stability parameters $\alpha \in (0, 1) \cup (1, 2]$, avoiding the significantly difficult case of $\alpha = 1$ corresponding to the distributions of the scaled and centred partial sums converging to a Cauchy distribution. This is seemingly the first stable large deviations result for dynamical systems with natural conditions that is close to the generality of Theorem 1.8. The method in Chapter 4 suggests a way to obtain error rates for general stable large deviations for the first time. In Chapter 5, we use the method from Chapter 4 in the i.i.d. setting to give error rates for large deviations in the particular setup of Cauchy distributed i.i.d. random variables (i.e. when $\alpha = 1$). The result that we obtain holds for the majority of the range of large deviations given in Theorem 1.8. We also obtain error rates over the full range by slightly modifying the type of deviations considered by introducing a large upper bound to the partial sums; we refer to these deviations as *quasi-large deviations*. As an application of the results obtained in Chapter 4 and Chapter 5, we prove stable large and quasi-large deviations statements and give error rates on the different means of digits appearing in some of the deterministic number expansions seen in Part I.

The necessary background and preliminaries for Part II will be provided in §1.4.

1.2 Dynamical Systems

We shall explain more precisely what defines a dynamical system and then give some of the basic concepts in dynamical systems theory that will be used throughout this thesis.

A *dynamical system* consists of an abstract space X representing the various states that the system can be in, referred to as the *state space*, along with a way to describe the evolution of X over time. In this thesis, we will only consider *discrete-time dynamical systems* so one way to define the dynamics of the system is by specifying a set $\mathcal{T} = \{T_s : X \to X\}_{s\in\mathcal{S}}$ of transformations indexed by a subset $\mathcal{S} \subset \mathbb{N}_0 := \{0, 1, 2, \dots\}$ of non-negative integers in such a way that T_{ω_n} describes the possible evolutions of the state space from the $(n-1)^{\text{th}}$ time-step to the n^{th} time-step for each $n \in \mathbb{N}$ and each choice of *fibre* $\omega = (\omega_n)_{n\in\mathbb{N}} \in \Omega := \mathcal{S}^{\mathbb{N}}$. We will adopt the measure-theoretic framework for our dynamical systems: that X, along with a sigma-algebra $\mathcal{A}$ and a measure μ, is a measure space and $\mathcal{T}$ is a collection of measurable non-singular transformations with respect to $(X, \mathcal{A}, \mu)$, i.e. each $T \in \mathcal{T}$ satisfies $T^{-1}(A) \in \mathcal{A}$ and $\mu(A) = 0 \iff \mu(T^{-1}(A)) = 0$ for all $A \in \mathcal{A}$. From now on, all dynamical systems mentioned in this thesis will be assumed to be discrete-time dynamical systems, and all transformations associated with dynamical systems will be assumed to be measurable and non-singular with respect to the underlying measure space.

In the above notation, if $\mathcal{T}$ contains only a single transformation $T : X \to X$, then we will denote the dynamical system by $(X, \mathcal{A}, \mu, T)$. In this case, the dynamics are autonomous and deterministic; the transformation T dictates the evolution of the state space X over time – a state $x \in X$ will be in state $T(x)$ at the next time-step, state $T(T(x))$ at the time-step after that, and so on. Writing

$$T^n := \overbrace{T \circ \cdots \circ T}^{n \text{ times}}, \qquad n \in \mathbb{N},$$

for the n^{th} iterate of T, we define the *orbit of x under T* to be the set $\{T^n(x) : n \in \mathbb{N}\}$ of states that x evolves into over time. As an example, consider the deterministic dynamical system $([0,1], \mathcal{B}([0,1]), \lambda, T)$, where $\mathcal{B}([0,1])$ denotes the Borel sigma-algebra on $[0,1]$, λ denotes the Lebesgue measure restricted to $[0,1]$ and $T : [0,1] \to [0,1]$ is the doubling map defined by $T(x) = 2x \bmod 1$. Figure 1.1(a) shows a graph of the doubling map; notice that the unit interval is partitioned into two subintervals $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1]$, and that T is linear and bijective on each of these subintervals. In general, if a map $[0,1] \to [0,1]$ is injective and continuous on the interior of each element of a countable partition of $[0,1]$, then each injective continuous part is called a *branch* and each partition element is called a *branch support* or just a support when the context is clear. Furthermore, if a branch is additionally surjective, then the branch is called *full*. So, the doubling map has two full linear branches with supports $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1]$. Figure 1.1(b) illustrates how to generate the orbit of a state x with respect to the doubling map.

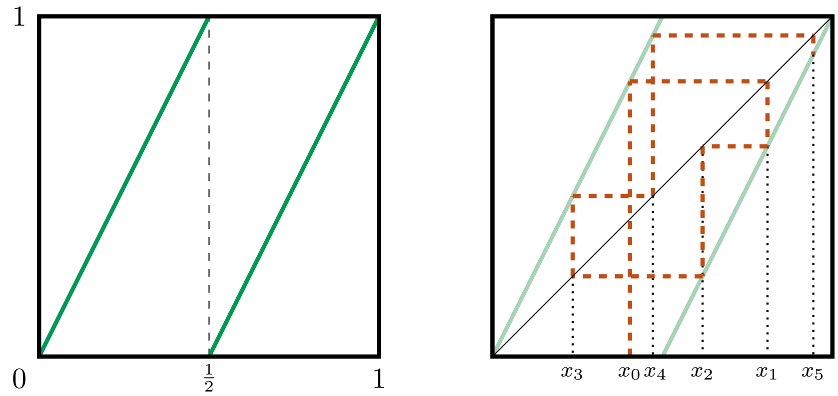
A dynamical system on a measure space $(X, \mathcal{A}, \mu)$ with dynamics dictated by a collection of transformations $\mathcal{T} = \{T : X \to X\}_{s\in\mathcal{S}}$ is *non-autonomous* since each state $x \in X$ may be mapped to different states at each time-step by the maps of $\mathcal{T}$ so the evolution over time of each x is ambiguous; see Figure 1.2(b). We will denote a non-autonomous dynamical system by $(X, \mathcal{A}, \mu, \mathcal{T})$. It is possible to obtain from $(X, \mathcal{A}, \mu, \mathcal{T})$ a deterministic system by considering the trajectories along a fixed fibre $\omega \in \Omega$. These trajectories are defined with respect to ω via the operator

$$T_\omega^n := T_{\omega_n} \circ \cdots \circ T_{\omega_2} \circ T_{\omega_1}, \qquad n \in \mathbb{N},$$

allowing us to define the *orbit of a state $x \in X$ along the fibre ω* by $\{T_\omega^n(x) : n \in \mathbb{N}\}$.

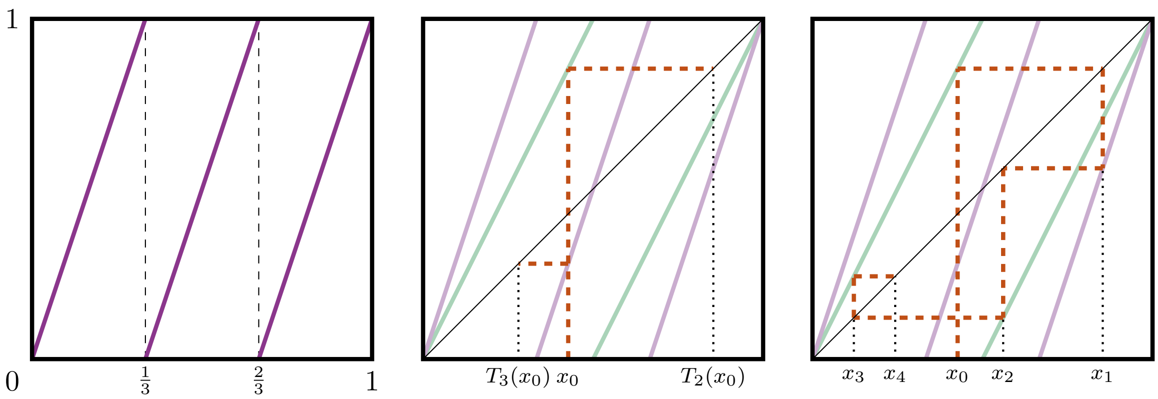
As an example, consider the following non-autonomous dynamical system $([0,1], \mathcal{B}([0,1]), \lambda, \mathcal{T})$ along a fibre ω, where $([0,1], \mathcal{B}([0,1]), \lambda)$ is the Lebesgue measure space on the unit interval as above and $\mathcal{T} = \{T_s\}_{s\in\{2,3\}}$ consists of the maps $T_s : [0,1] \to [0,1]$, $x \mapsto sx \bmod 1$, $s \in \{2,3\}$; see Figure 1.1(a) and Figure 1.2(a) for graphs of T_2 and T_3 respectively. Figure 1.2(c) shows the orbit of a number $x_0 \in [0,1]$ along a fibre ω of the form $\omega = (2,3,2,2,\dots) \in \{2,3\}^{\mathbb{N}}$.

Given a non-autonomous dynamical system $(X, \mathcal{A}, \mu, \mathcal{T})$, one can obtain an associated random dynamical system by introducing a map $\sigma : \Omega \to \Omega$ defined via some random process that drives the non-autonomous dynamics. The dynamics of a random dynamical system are then determined autonomously by the *skew-product* $R : \Omega \times X \to \Omega \times X$, $(\omega, x) \mapsto (\sigma(\omega), T_{\omega_1}(x))$. The first coordinate keeps track of the driving system and the second coordinate gives the dynamics of the non-autonomous system according to the driving system.



(a) Doubling map (b) Iteration

Figure 1.1: Iteration of the doubling map $T : [0,1] \to [0,1]$, $x \mapsto 2x \bmod 1$. (a) shows the graph of T. The green lines are the branches of T. (b) shows the iteration of T to obtain the orbit of a state x_0 up to the 5^{th} iterate. The solid, black line has equation $y = x$. The red, dashed lines show the iteration process. For $n = 1, \dots, 5$, $x_n := T^n(x_0)$.



(a) Trebling map (b) Non-deterministic iteration (c) Deterministic iteration

Figure 1.2: Iteration of the non-autonomous dynamical system $([0,1], \mathcal{B}([0,1]), \lambda, \mathcal{T})$, where $\mathcal{T} = \{T_s\}_{s\in\{2,3\}}$ consists of the maps $T_s(x) = sx \bmod 1$, $s \in \{2,3\}$. (a) shows the graph of the trebling map T_3. (b) shows the different images of a number $x_0 \in [0,1]$ under the maps of $\mathcal{T}$. (c) shows the beginning of an orbit of a number x_0 along a fibre ω of the form $\omega = (2,3,2,2,\dots) \in \{2,3\}^{\mathbb{N}}$, where $x_j := T^j(x_0)$ for $j = 1,2,3,4$.

For the majority of Part I, we will be interested in the fibres of non-autonomous dynamical systems constructed from families of piecewise linear, full-branched unit interval maps such as the doubling map and trebling map above as well as maps that have a countable infinity of branches.

In Part II, we will treat a broad class of abstract deterministic dynamical systems including Gibbs-Markov maps. Let $(X, \mathcal{A}, \mu)$ be a probability space with a countable measurable partition $\{I_j\}_{j\in\mathbb{N}}$, and let $T : X \to X$ be a *measure-preserving* transformation, meaning that $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{A}$. We further assume that μ is 'ergodic' with respect to T; see Definition 1.6 in §1.3.4. Define $s(y, y')$ to be the least integer $n \geq 0$ such that $T^n(y)$ and $T^n(y')$ lie in distinct partition elements. Assuming that $s(y, y') = +\infty$ if and only if $y = y'$, it follows that $d_\theta(y, y') = \theta^{s(y,y')}$ defines a metric for each $\theta \in (0, 1)$. Put $g := \frac{\mathrm{d}\mu}{\mathrm{d}\mu\circ T} : X \to \mathbb{R}$.

Definition 1.1 (Gibbs-Markov Map)**.** *Let $(X, \mathcal{A}, \mu, T)$ be a deterministic dynamical system with μ ergodic with respect to the measure-preserving transformation T. We say that T is a Gibbs-Markov map and that $(X, \mathcal{A}, \mu, T)$ is a Gibbs-Markov dynamical system if T has the following three properties.*

- (Markov Map) *$T(I_j)$ is a union of partition elements and $T|_{I_j} : I_j \to T(I_j)$ is a measurable bijection for each $j \in \mathbb{N}$*
- (Big Images) $\inf\{\mu(T(I_j)) : j \in \mathbb{N}\} > 0$
- (Bounded Distortion) *There are constants $C > 0$ and $\theta \in (0, 1)$ such that, for each $j \in \mathbb{N}$ and $y, y' \in I_j$, $|\log g(y) - \log g(y')| \leq C d_\theta(y, y')$*

The maps with a countable infinity of branches that we will see throughout the thesis, such as certain GLS maps and the continued fraction map defined in §1.3.1, are examples of Gibbs-Markov maps. For a more detailed background on Gibbs-Markov maps, see for instance [Aar97, Chapter 4] and [AD01].

The broad class of dynamical systems considered in Part II also includes AFU maps.

Definition 1.2 (AFU Map)**.** *Let $\{I_j\}$ be a measurable partition of $[0, 1]$ into open intervals. We call a map $T : [0, 1] \to [0, 1]$ an AFU map if $T|_{I_j}$ is $\mathcal{C}^2$ and strictly monotone for each I_j, and*

(A) (Adler's Condition) *$T''/(T')^2$ is bounded on $\bigcup_j I_j$, where T' and T'' denote the first and second derivatives of T respectively.*

(F) (Finite Images) *The set of images $\{T(I_j)\}$ is finite.*

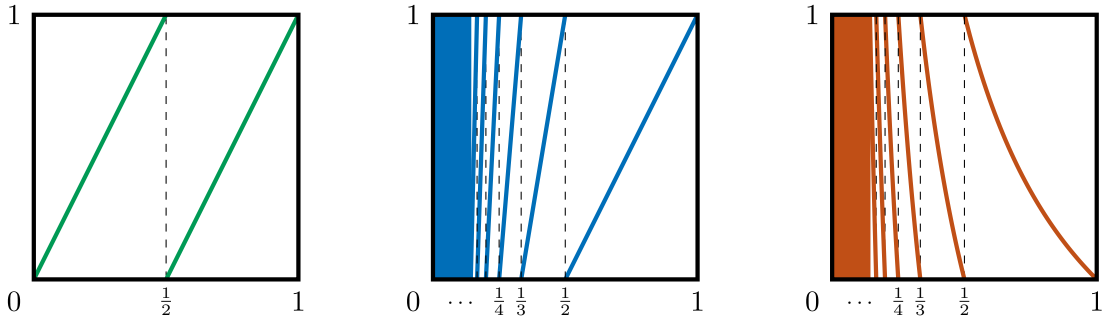
(U) (Uniform Expansion) *There exists $\rho > 1$ such that $|T'| \geq \rho$ on $\bigcup_j I_j$.*

Such maps were introduced in [Zwe98] (see [ADSZ04] also).

1.3 Expansions, Digit Frequencies and Dimensions

1.3.1 Number Systems

A number expansion is a representation of a real number using a finite or infinite string of *digits* from a given set of symbols called a *digit set*. Common examples include the decimal expansion, which uses the digit set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, and the binary expansion, which uses the digit set $\{0, 1\}$. We can generate expansions of numbers by using a type of dynamical system known as a *number system*. This is a dynamical system on a set of numbers that assigns each number



(a) Doubling map (b) Lüroth map (c) Continued fraction map

Figure 1.3: Examples of dynamical systems that generate number expansions

a symbolic representation. The symbols used in these representations determine the associated digit set. For example, the number system $([0,1], \mathcal{B}([0,1]), \lambda, T)$ with λ the Lebesgue measure and T the doubling map as defined in §1.2 (see Figure 1.3(a)) generates binary expansions for the numbers in $[0,1]$. By assigning the digit 0 to the branch of T with support $[0, \frac{1}{2})$, the digit 1 to the branch of T with support $[\frac{1}{2}, 1]$ and then iterating the system, we follow the orbit of each number $x \in [0,1]$ and track the digits met along the way. In this way, we obtain a sequence of digits $(b_j(x))_{j\in\mathbb{N}}$ given by $b_j(x) = 0$ whenever $T^{j-1}(x) \in [0, \frac{1}{2})$ and $b_j(x) = 1$ whenever $T^{j-1}(x) \in [\frac{1}{2}, 1]$. From Figure 1.1(b), one can see that the number $x_0 \in [0,1]$ has $b_1(x_0) = 0, b_2(x_0) = 1, b_3(x_0) = 1$, and so on. For each x, the sequence $(b_j(x))_{j\in\mathbb{N}}$ can be put into the *binary series expansion* to recover x:

$$x = \sum_{j\in\mathbb{N}} \frac{b_j(x)}{2^j} = \frac{b_1(x)}{2^1} + \frac{b_2(x)}{2^2} + \frac{b_3(x)}{2^3} + \cdots .$$

For each number x, the number of representations of x can range from none at all to uncountably many depending on the number system. For example, the number $\pi - 3 \in [0,1]$ has a unique (infinite) decimal expansion:

$$\pi - 3 = \frac{1}{10^1} + \frac{4}{10^2} + \frac{1}{10^3} + \frac{5}{10^4} + \frac{9}{10^5} + \frac{2}{10^6} + \frac{6}{10^7} + \frac{5}{10^8} + \frac{3}{10^9} + \cdots = 0.141592653\cdots$$

whilst $\frac{1}{2}$ has two decimal expansions – one with an infinite tail of 0s (so the expansion is finite) and one with an infinite tail of 1s:

$$\frac{1}{2} = \frac{1}{2^1} + \frac{0}{2^2} + \frac{0}{2^3} + \frac{0}{2^4} + \cdots = 0.5, \qquad \frac{1}{2} = \frac{0}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots = 0.499999\cdots .$$

The binary and decimal number systems belong to the family of *integer-base number systems*. The branches of a map giving one such number system are all full and linear with the same integer-valued slope, called the *base*. One can also consider number systems where the supports of the branches do not equally partition the unit interval; such partitions may admit a countable infinity of branches. For example, the Lüroth series expansions, which were first studied by J. Lüroth in [Lür83] and have the form

$$x = \frac{1}{b_1(x)} + \frac{1}{b_1(x)(b_1(x)+1)b_2(x)} + \cdots + \frac{1}{b_n(x)\prod_{1\le k<n} b_k(x)(b_k(x)+1)} + \cdots \tag{1.1}$$

for sequences $(b_n(x))_{n\in\mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$, can be generated by the number system $([0,1], \mathcal{B}([0,1]), \lambda, L)$, where $L : [0,1] \to [0,1]$ is the Lüroth map, which was first introduced in [JV69] and is defined by $L(0) = 0$ and piecewise on the intervals $\left(\frac{1}{k+1}, \frac{1}{k}\right]$, $k \in \mathbb{N}$, by

$$L(x) = k(k+1)x - k \quad \text{whenever} \quad x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]. \tag{1.2}$$

Figure 1.3(b) shows a graph of the Lüroth map. A sequence $(A_k)_{k\in\mathbb{N}}$ of subsets of $[0,1]$ has an *accumulation point* at c if for any $\varepsilon > 0$, there exists a $k_0 \in \mathbb{N}$ so that $A_k \subset (c-\varepsilon, c+\varepsilon)$ for all $k \geq k_0$. Notice then that the partition $\{(\frac{1}{k+1}, \frac{1}{k}] : k \in \mathbb{N}\}$ of supports of branches of the Lüroth map is countably infinite with an accumulation point at $c = 0$. In [BBDK96], the authors generalise Lüroth series expansions by introducing Generalised Lüroth Series (GLS) expansions, which are generated by a family of dynamical systems of the form $([0,1], \mathcal{B}([0,1]), \lambda, T)$ with $T : [0,1] \to [0,1]$ any piecewise linear full-branched interval map. Such maps are called *GLS maps* and have countably many branches, each of positive or negative slope, that are supported on some partition of the unit interval. If we index the branches with a set $\mathcal{B} \subset \mathbb{N}$ and write $(-1)^{\varepsilon_b} N_b$, $b \in \mathcal{B}$, for the signed slope of the b^{th} branch, where $\varepsilon_b \in \{0,1\}$ and $N_b > 1$, then the associated number expansions generated by GLS number systems have the form

$$x = \frac{t_{b_1}}{N_{b_1}} + (-1)^{\varepsilon_{b_1}} \frac{t_{b_2}}{N_{b_1} N_{b_2}} + \cdots + (-1)^{\varepsilon_{b_1} + \cdots + \varepsilon_{b_{n-1}}} \frac{t_{b_n}}{\prod_{1 \leq \ell \leq n} N_{b_\ell}} + \cdots, \tag{1.3}$$

where t_b is a non-negative quantity depending on the position of the b^{th} branch on the unit interval. As the name suggests, we may obtain Lüroth series expansions as a type of GLS expansion by taking the partition $\{(\frac{1}{k+1}, \frac{1}{k}] : k \in \mathbb{N}\} \cup \{0\}$ of the unit interval and taking all branches to have positive slopes. We will define GLS maps explicitly in §2.1.

Even more distinct from the previous examples of number expansions are regular continued fraction expansions. Instead of representing each number with a series expansion, a number is instead represented with a nest of finitely or infinitely continuing fractions. For example, the number $\pi - 3 \in [0,1]$ has a unique (infinite) continued fraction expansion

$$\pi - 3 = \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \cdots}}}}.$$

The dynamics of continued fraction expansions are given on the same partition as the Lüroth map by the continued fraction map $G : [0,1] \to [0,1]$ defined by $G(0) = 0$ and piecewise on the intervals $(\frac{1}{k+1}, \frac{1}{k}]$, $k \in \mathbb{N}$, by

$$G(x) = \frac{1}{x} - k \quad \text{whenever} \quad x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]. \tag{1.4}$$

Figure 1.3(c) shows a graph of the continued fraction map.

So far, the number systems encountered have all been examples of dynamical systems given by a single map so only that one map and its associated digit set are used to generate the different number expansions. We can also consider non-autonomous number systems that encode numbers in a time-dependent manner by using more than one map – each with its own digit set. As a simple example, one could consider the non-autonomous number system

$([0,1], \mathcal{B}([0,1]), \lambda, \mathcal{T})$ along the fibre ω from Figure 1.2. Number expansions in this system have the form

$$x = \frac{b_{\omega_1}}{N_{\omega_1}} + \frac{b_{\omega_2}}{N_{\omega_1} N_{\omega_2}} + \cdots = \sum_{n\in\mathbb{N}} \frac{b_{\omega_n}}{\prod_{1\le k\le n} N_{\omega_k}},$$

for sequences $(b_n)_{n\in\mathbb{N}}$ such that $b_n \in \{0,1\}$ for all $n \in \mathbb{N}$ with $\omega_n = 2$ and $b_n \in \{0,1,2\}$ otherwise, where N_s, $s \in \{2,3\}$, denotes the slope of the branches of the s^{th} map, i.e. $N_2 \coloneqq 2$ and $N_3 \coloneqq 3$. So, the denominators appearing in these expansions are products of 2s and 3s. For a more general example, consider the collection $\mathcal{T} = \{T_1, T_2\}$ of any two GLS maps. For each $s \in \{1,2\}$, index the branches of each map T_s with elements from an indexing set $\mathcal{B}_s \subset \mathbb{N}$, and, for each $b \in \mathcal{B}_s$, denote by $(-1)^{\varepsilon_{s,b}} N_{s,b}$ the signed slope of the b^{th} branch of the map T_s, where $\varepsilon_{s,b} \in \{0,1\}$ and $N_{s,b} > 1$. Then any number expansion of $x \in [0,1]$ generated by the number system $([0,1], \mathcal{B}([0,1]), \lambda, \mathcal{T})$ along a fibre $\omega \in \{1,2\}^{\mathbb{N}}$ has the form

$$x = \frac{t_{\omega_1,b_1}}{N_{\omega_1,b_1}} + (-1)^{\varepsilon_{\omega_1,b_1}} \frac{t_{\omega_2,b_2}}{N_{\omega_1,b_1} N_{\omega_2,b_2}} + \cdots + (-1)^{\varepsilon_{\omega_1,b_1}+\cdots+\varepsilon_{\omega_{n-1},b_{n-1}}} \frac{t_{\omega_n,b_n}}{\prod_{1\le \ell\le n} N_{\omega_\ell,b_\ell}} + \cdots \qquad (1.5)$$

with $t_{s,b}$ a non-negative quantity depending on the position of the b^{th} branch of T_s on the unit interval, where $(b_n)_{n\in\mathbb{N}}$ is an *ω-compatible* sequence, i.e. $b_n \in \mathcal{B}_{\omega_n}$ for all $n \in \mathbb{N}$. Depending on the context, we will refer to expansions of the form in (1.5) as *GLS IFS expansions* or as *Non-autonomous GLS* (NGLS) *expansions*. We write $(b_n(\omega,x))_{n\in\mathbb{N}}$ for the particular ω-compatible sequence giving the expansion of x in (1.5) and refer to it as the *digit sequence of x along ω* with respect to the number system $([0,1], \mathcal{B}([0,1]), \lambda, \mathcal{T})$. In other words, we have for each $n \in \mathbb{N}$ that $b_n(\omega,x) \in \mathcal{B}_{\omega_n}$ is such that $b_n(\omega,x) = b$ if and only if $T_\omega^{n-1}(x)$ belongs to the support of the b^{th} branch of T_{ω_n}. Observe that expansions of the form (1.5) are determined completely by pairs of indices (s,b) with $b \in \mathcal{B}_s$ and $s \in \mathcal{S}$. Consequently, we define

$$\mathcal{D} \coloneqq \{(s,b) : b \in \mathcal{B}_s,\ s \in \mathcal{S}\}$$

to be the *digit set* of a non-autonomous number system and we call its elements *digits*.

In Part I, we will be interested in the frequency of occurrence of the digits of $\mathcal{D}$ appearing in expansions of the form (1.5) generated by non-autonomous number systems constructed from finite families of GLS maps.

1.3.2 Symbolic Dynamics

Let $(X, \mathcal{A}, \mu, \mathcal{T})$ be a non-autonomous number system, and let $\mathcal{D}$ be the associated digit set. So, $\mathcal{T}$ is a collection of maps indexed by a set $\mathcal{S}$ and $\mathcal{D} = \{(s,b) : b \in \mathcal{B}_s, s \in \mathcal{S}\}$, where $\mathcal{B}_s$ is the index set of the branches of $T_s \in \mathcal{T}$. Fix a fibre $\omega \in \Omega \coloneqq \mathcal{S}^{\mathbb{N}}$. As explained in §1.3.1, the numbers $x \in X$ can naturally be associated with sequences $(b_n(\omega,x))_{n\in\mathbb{N}}$ of digits in $\mathcal{D}$. In the examples of non-autonomous number systems that we will consider later in the thesis, this gives a conjugacy between the dynamics of the system and the symbolic dynamics of the left-shift map $\sigma_{\mathcal{D}}$ defined on the space of infinite sequences $(d_n)_{n\in\mathbb{N}}$ of digits in $\mathcal{D}$ by $(d_n)_{n\in\mathbb{N}} \mapsto (d_{n+1})_{n\in\mathbb{N}}$. Consequently, much of the study of these number systems may be reduced to studying properties of the symbolic dynamics of the left-shift map $\sigma_{\mathcal{D}}$.

We give some general definitions and notation for symbolic dynamical systems. Let Υ be a countable set of symbols, which we call an *alphabet*, and denote by $\Upsilon^{\mathbb{N}}$ the set of one-sided infinite sequences $(\upsilon_1, \upsilon_2, \ldots)$ of elements in Υ. For each $n \in \mathbb{N}_0$, the set Υ^n is the set of *words* of length n, where we let $\Upsilon^0 \coloneqq \{\varnothing\}$ be the set containing only the *empty word* denoted by $\varnothing$. Let $\Upsilon^* \coloneqq \bigcup_{n\in\mathbb{N}_0} \Upsilon^n$ be the set of all words. For a word $\boldsymbol{\upsilon} \in \Upsilon^*$, we use the notation $|\boldsymbol{\upsilon}|$ for its length so that $|\boldsymbol{v}| = n$ if $\boldsymbol{v} \in \Upsilon^n$. If $\boldsymbol{v} = v_1 \cdots v_n \in \Upsilon^*$, then for each $k \leq n$, we use the notation $\boldsymbol{v}|_k := v_1 \cdots v_k$. Similarly, for a sequence $\xi = (\xi_n)_{n\in\mathbb{N}} \in \Upsilon^{\mathbb{N}}$ and any $n \in \mathbb{N}$, we set $\xi|_n := \xi_1 \cdots \xi_n$.

We may specify a topology on $\Upsilon^{\mathbb{N}}$ by considering the compact metric space $(\Upsilon^{\mathbb{N}}, \rho)$, where $\rho : \Upsilon^{\mathbb{N}} \times \Upsilon^{\mathbb{N}} \to [0,1]$ is the metric given by

$$\rho(\xi, \zeta) = \begin{cases} 2^{-\min\{n\in\mathbb{N}:\, \xi_n \neq \zeta_n\}}, & \xi \neq \zeta, \\ 0, & \xi = \zeta. \end{cases}$$

The *level-n cylinder set* corresponding to a word $\boldsymbol{v} \in \Upsilon^n$, $n \in \mathbb{N}_0$, is denoted by

$$[\boldsymbol{v}] = \{\xi \in \Upsilon^{\mathbb{N}} : \xi|_n = \boldsymbol{v}\}.$$

and is both open and closed with respect to the topology on $\Upsilon^{\mathbb{N}}$ specified by ρ. The collection of all cylinder sets generates the Borel sigma-algebra on $\Upsilon^{\mathbb{N}}$.

The *left-shift map* σ_Υ *on the alphabet* Υ is the map $\sigma_\Upsilon : \Upsilon^{\mathbb{N}} \to \Upsilon^{\mathbb{N}}$ given by $(\xi_n)_{n\in\mathbb{N}} \mapsto (\xi_{n+1})_{n\in\mathbb{N}}$. With a slight abuse of notation, we will use σ to denote the left-shift map on any sequence space without specifying the alphabet as a subscript whenever no confusion can arise. Let $\mathcal{M}(\Upsilon^{\mathbb{N}}, \sigma)$ denote the set of all shift-invariant Borel probability measures on $\Upsilon^{\mathbb{N}}$. For $\mu \in \mathcal{M}(\Upsilon^{\mathbb{N}}, \sigma)$, we use $h_\mu(\sigma)$ to denote the measure-theoretic entropy of μ with respect to σ. As a consequence of the Kolmogorov-Sinai Theorem (see e.g. [DK21]),

$$h_\mu(\sigma) = -\lim_{n\to+\infty} \frac{1}{n} \sum_{\boldsymbol{v}\in\Upsilon_n} \mu([\boldsymbol{v}]) \log \mu([\boldsymbol{v}]),$$

where $0 \log 0 := 0$. Given a probability vector $\mathbf{p} = (p_v)_{v\in\Upsilon}$, the $\mathbf{p}$-*Bernoulli measure* $\mu_{\mathbf{p}}$ is the probability measure on $(\Upsilon^{\mathbb{N}}, \sigma)$ defined on each cylinder $[\boldsymbol{v}] = [v_1 \cdots v_n]$ by $\mu_{\mathbf{p}}([\boldsymbol{v}]) = p_{v_1} \cdots p_{v_n}$. It is well-known (see e.g. [DK21, p158]) that the measure-theoretic entropy of $\mu_{\mathbf{p}}$ with respect to σ is given by

$$h_{\mu_{\mathbf{p}}}(\sigma) = -\sum_{v\in\Upsilon} p_v \log p_u. \tag{1.6}$$

For each $\xi \in \Upsilon^{\mathbb{N}}$, $v \in \Upsilon$ and $n \in \mathbb{N}$, we denote by $\tau_v(\xi, n) := \#\{1 \leq \ell \leq n : \xi_\ell = v\}$ the number of times that the symbol v occurs in the first n elements of the sequence ξ, and we write

$$\tau_v(\xi) := \lim_{n\to+\infty} \frac{\tau_v(\xi, n)}{n}$$

for the frequency of the digit v in ξ when it exists.

1.3.3 Fractal Geometry

A *fractal* is a mathematical object that typically displays detailed geometric structure at arbitrarily small scales and is usually characterised by having a rough boundary and a repeating structure. A fractal can often be defined as the *attractor* of an *iterated function system* (IFS), which we define as follows.

Definition 1.3. *Fix* $m \in \mathbb{N}$.

(i) (Contraction) *A map* $g : \mathbb{R}^m \to \mathbb{R}^m$ *is a contraction if* $\big|g(\mathbf{x}) - g(\mathbf{y})\big| \leq r|\mathbf{x} - \mathbf{y}|$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ *and some* $r \in (0, +\infty)$. *We call* r *the contraction ratio of* g.

(ii) (Iterated Function System) *An IFS is a countable collection $\mathcal{I} = \{g_u : \mathbb{R}^m \to \mathbb{R}^m\}_{u\in\mathcal{U}}$ of contractions.*

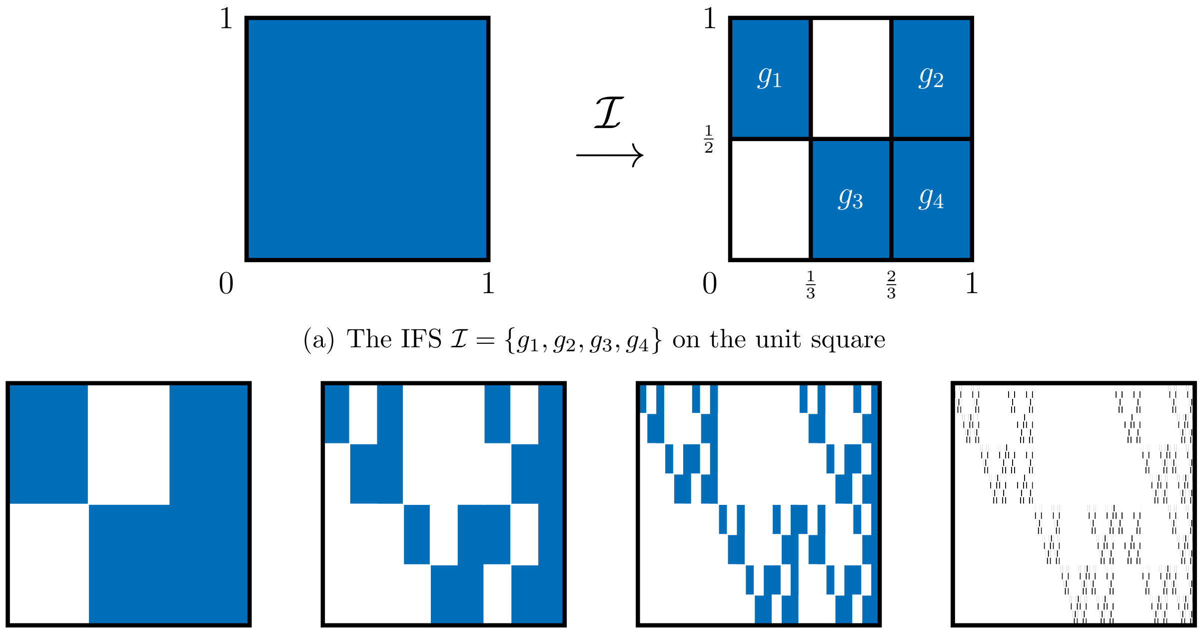
(iii) (Attractor) *An attractor of an IFS $\mathcal{I} = \{g_u : \mathbb{R}^m \to \mathbb{R}^m\}_{u\in\mathcal{U}}$ is a set $\Lambda \subset \mathbb{R}^m$ with the property that $g_u(\Lambda) = \Lambda$ for all $u \in \mathcal{U}$.*

It was shown in [Hut81] that an IFS $\mathcal{I}$ of contractions on $\mathbb{R}^m$ always has a unique non-empty attractor. If the contractions are all *affine*, i.e. if each $g \in \mathcal{I}$ may be written in the form $g(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{v}$ with $\mathbf{A} \in \mathrm{GL}_m(\mathbb{R})$ and $\mathbf{v} \in \mathbb{R}^m$, then the IFS is called affine and the resulting attractor is called a *self-affine set*. In addition, if the matrices $\mathbf{A} \in \mathrm{GL}_m(\mathbb{R})$ are all diagonal, i.e. if the non-diagonal entries of each matrix are all zero, then the IFS is called *diagonally affine*. In the specific case when $m = 2$, we call a self-affine set a *carpet*. A simple family of carpets are the *Bedford-McMullen carpets* first studied independently in [Bed84; McM84]. These sets are defined to be the attractors of affine IFSs $\mathcal{I} = \{g_u : [0,1]^2 \to [0,1]^2\}_{u\in\mathcal{U}}$ with each map g_u, $u \in \mathcal{U}$, of the form

$$g_u(\mathbf{x}) = \begin{bmatrix} 1/r & 0 \\ 0 & 1/s \end{bmatrix}\mathbf{x} + \begin{bmatrix} (i_u - 1)/r \\ (j_u - 1)/s \end{bmatrix},$$

where $(i_u, j_u) \in \{(i_1, j_1), \ldots, (i_k, j_k)\} \subset \{1, \ldots, r\} \times \{1, \ldots, s\}$ for integers $r \geq s \in \mathbb{N}$ and $k \leq rs$. In other words, the unit square is first divided into rs rectangles each with the same height and the same width and then the k rectangles with indices in the above set are selected and the others discarded. This process is repeated within each of the k selected rectangles by subdividing each of them into rs rectangles and choosing in each the k rectangles with the correct indices. After infinitely many iterations, the Bedford-McMullen carpet is the collection of points that have not been discarded; see Figure 1.4.

In Chapter 2, we consider a construction similar to that of the Bedford-McMullen carpets but we instead partition the unit square into horizontal strips and then *independently* subdivide each



(a) The IFS $\mathcal{I} = \{g_1, g_2, g_3, g_4\}$ on the unit square

(b) 1st Iteration (c) 2nd Iteration (d) 3rd Iteration (e) Attractor

Figure 1.4: The generation of a Bedford-McMullen carpet as the attractor of an IFS $\mathcal{I}$. The blue regions indicate the interior and the white regions indicate void. (a) shows the action of each g_u, $u = 1, 2, 3, 4$, on $[0,1]^2$. (b)-(d) show the first three iterations of $\mathcal{I}$ on $[0,1]^2$. (e) shows an approximation of the attractor of $\mathcal{I}$, a Bedford-McMullen carpet, on $[0,1]^2$.

of the strips into rectangles whose widths are all strictly less than their heights (or vice versa). If we were to continue as with the construction of Bedford-McMullen carpets by selecting a collection of rectangles and repeating the same process to these rectangles infinitely many times, then we would obtain an attractor known as a *Lalley-Gatzouras carpet*; see [LG92]. We study level sets on Lalley-Gatzouras carpets under particular conditions concerning the overlaps of functions in the IFS. Without discarding any of the rectangles, the resulting attractor is just the unit square. Such IFSs that generate the unit square, which we will later refer to as *GLS IFSs*, may be associated to a skew-product map defined on the unit square by a finite collection of GLS maps with finitely many branches; we study the level sets defined by fibres of these systems as well.

The usual notions of size, e.g. length, area, volume, are often unable to capture the geometric complexity of fractals. Instead, a more complex measure of size is needed, which we give in the next definition.

Definition 1.4. *Fix $m \in \mathbb{N}$ and $A \subset \mathbb{R}^m$. For each $B \subset \mathbb{R}^m$, denote by $|B|$ the Euclidean diameter of B.*

(i) (δ-cover) *For any $\delta > 0$, a cover $\mathcal{U} = \{U_j\}_{j \in \mathcal{J}}$ of A is called a δ-cover of A if $|U_j| < \delta$ for all $j \in \mathcal{J}$.*

(ii) (Hausdorff Outer Measure) *For any $t \geq 0$, the t-Hausdorff outer-measure of A is given by*

$$\mathcal{H}^t(A) := \lim_{\delta \to 0} \inf \left\{ \sum_{U \in \mathcal{U}} |U|^t : \mathcal{U} \text{ is a } \delta\text{-cover of } A \right\}.$$

(iii) (Hausdorff Dimension) *The Hausdorff dimension $\dim_H A$ of A is given by*

$$\dim_H A := \inf\{t \geq 0 : \mathcal{H}^t(A) = 0\}.$$

Equivalently (see e.g. [Fal14]), $\dim_H A$ can be defined as $\dim_H A := \sup\{t \geq 0 : \mathcal{H}^t(A) = +\infty\}$ or as the number $t_0 \geq 0$ with the property that

$$\mathcal{H}^t(A) = \begin{cases} +\infty, & t < t_0, \\ 0, & t > t_0. \end{cases}$$

In other words, $\dim_H A$ is the value of $t \geq 0$ where $\mathcal{H}^t(A)$ transitions from $+\infty$ to 0; see Figure 1.5. Note that $\mathcal{H}^{t_0}(A)$ can be any value in $[0, +\infty]$.

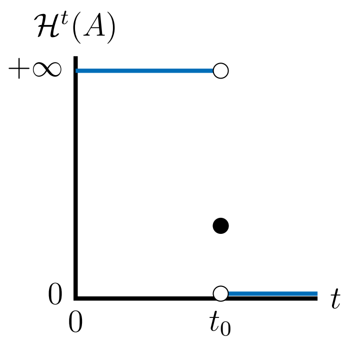


Figure 1.5: A plot of $\mathcal{H}^t(A)$ over $t \geq 0$ for a set A with $\dim_H A = t_0$.

As an example, the Hausdorff dimension of a general Bedford-McMullen carpet is given by the formula (see [Bed84; McM84])

$$\frac{\log \sum_{1\leq j\leq s} \#\{1 \leq \ell \leq k : j_\ell = j\}^{\frac{\log s}{\log r}}}{\log s}.$$

In particular, the Hausdorff dimension of the Bedford-McMullen carpet Λ in Figure 1.4 is

$$\dim_H \Lambda = \frac{\log(2 \cdot 2^{\frac{\log 2}{\log 3}})}{\log 2} = 1 + \frac{\log 2}{\log 3} \approx 1.6309.$$

The Hausdorff dimension of a set is often very difficult to calculate directly due to the need to find an optimal cover within the definition of the Hausdorff outer-measure; therefore, much of the focus of fractal geometry is to find different techniques that bypass this complication, allowing the Hausdorff dimensions of fractals to be calculated explicitly even when the optimal cover is unknown. One such strategy is to relate the Hausdorff dimension of a set with certain dimension quantities for suitably chosen measures – we introduce these dimension quantities now. For a Borel probability measure μ supported on a subset $X \subset \mathbb{R}^m$, $m \in \mathbb{N}$, we define the *Hausdorff dimension of* μ to be

$$\dim_H \mu = \inf\{\dim_H A : A \subset X \text{ a Borel set, } \mu(A) = 1\}.$$

A fact that we will use several times is the following: for any Borel set $A \subset X$ with $\mu(A) = 1$, it follows that $\dim_H A \geq \dim_H \mu$. We will also make use of the following equivalent definition of $\dim_H \mu$ (see e.g. [Pes97, p41]):

$$\dim_H \mu = \lim_{\varepsilon\to 0} \inf\{\dim_H A : A \subset X \text{ a Borel set, } \mu(A) > 1 - \varepsilon\}. \tag{1.7}$$

In addition, we define the *lower pointwise dimension* of μ at a point $\mathbf{x} \in A$ by

$$\underline{d}_\mu(\mathbf{x}) := \liminf_{r\to 0} \frac{\log \mu(B(\mathbf{x}, r))}{\log r},$$

where $B(\mathbf{x}, r)$ denotes the open Euclidean ball in $\mathbb{R}^m$ with radius r centred at $\mathbf{x}$.

We will use the next result to relate the above concepts of Hausdorff dimension and pointwise dimension (see e.g. [You84], [Pes97, Theorems 7.1 and 7.2]).

Lemma 1.5. *Fix $m \in \mathbb{N}$. Let μ be a finite Borel measure supported on $X \subset \mathbb{R}^m$, and take a Borel subset $A \subseteq X$.*

(i) *If $\underline{d}_\mu(\mathbf{x}) \leq c$ for some $c > 0$ and every $\mathbf{x} \in A$, then $\dim_H A \leq c$.*

(ii) *If $\underline{d}_\mu(\mathbf{x}) \geq c$ for some $c > 0$ and μ-a.e. $\mathbf{x} \in X$, then $\dim_H \mu \geq c$.*

1.3.4 Ergodic Theory

Let $(X, \mathcal{A}, \mu, T)$ be a dynamical system driven by a single transformation $T : X \to X$, and recall that T is therefore assumed to be $(X, \mathcal{A})$-measurable. We will assume throughout this section that μ is a probability measure. Recall that $(X, \mathcal{A}, \mu, T)$ is measure-preserving if we have $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{A}$. In this case, we call T a *μ-preserving transformation* and μ a *T-invariant measure*.

1

Ergodic theory is the study of the average behaviour of measure-preserving dynamical systems over time. Central to ergodic theory are ergodic transformations: the measure-preserving transformations $T : X \to X$ that, for any $A \in \mathcal{A}$ with $T^{-1}(A) = A$, can only be decomposed into measurable transformations $T_A : A \to A$ and $T_{X \setminus A} : X \setminus A \to X \setminus A$ trivially, i.e. such that T is the same as one of T_A or $T_{X \setminus A}$ up to μ-null sets. This means that, when studying the average behaviour of measure-preserving dynamical systems, it is only necessary to consider the ergodic transformations. To be more precise, we give the following definition of ergodicity.

Definition 1.6 (Ergodicity)**.** *Let $(X, \mathcal{A}, \mu, T)$ be a measure-preserving deterministic dynamical system. Then T is ergodic if, for every $A \in \mathcal{A}$ with $T^{-1}(A) = A$, either $\mu(A) = 0$ or $\mu(A) = 1$.*

When T is ergodic, then we say that μ is ergodic with respect to T or that μ is T-ergodic, and we call $(X, \mathcal{A}, \mu, T)$ an *ergodic dynamical system.* A fundamental result in ergodic theory for ergodic dynamical systems is the Birkhoff Ergodic Theorem, which states that the time average of an integrable potential (function) converges to the space average.

Theorem 1.7 (Birkhoff Ergodic Theorem)**.** *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic dynamical system. Then for any potential $\Phi : X \to \mathbb{R}^m$, $m \in \mathbb{N}$, that belongs to the set $\mathcal{L}^1(X, \mathcal{B}, \mu)$ of μ-integrable functions,*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{1 \leq \ell \leq n} \Phi \circ T^{\ell - 1}(x) = \int_X \Phi \, \mathrm{d}\mu, \qquad \mu\text{-a.e. } x \in X. \tag{1.8}$$

The left-hand side of (1.8) is called the *Birkhoff average of* Φ *at* x. Given a dynamical system $(X, \mathcal{A}, \mu, T)$ and a potential $\Phi : X \to \mathbb{R}^m$, one may consider the $\boldsymbol{\alpha}$-level sets

$$F_\Phi(\boldsymbol{\alpha}) := \left\{ x \in X : \lim_{n \to +\infty} \frac{1}{n} \sum_{1 \leq \ell \leq n} \Phi \circ T^{\ell - 1}(x) = \boldsymbol{\alpha} \right\}, \qquad \boldsymbol{\alpha} \in \mathbb{R}^m,$$

of points whose Birkhoff average of Φ is given by the vector $\boldsymbol{\alpha}$. If μ is T-ergodic, then the Birkhoff Ergodic Theorem implies that $\mu(F_\Phi(\overline{\boldsymbol{\alpha}})) = 1$ for $\overline{\boldsymbol{\alpha}} := \int_X \Phi \, \mathrm{d}\mu$ and $\mu(F_\Phi(\boldsymbol{\alpha})) = 0$ for all other $\boldsymbol{\alpha} \in \mathbb{R}^m$. It is therefore logical to quantify the size of Birkhoff average $\boldsymbol{\alpha}$-level sets with a function that distinguishes between zero measure sets such as the Hausdorff dimension from Definition 1.4.

The main object of study in Part I is the *Birkhoff dimension spectrum of a potential* Φ, which is the function $\boldsymbol{\alpha} \mapsto \dim_H F_\Phi(\boldsymbol{\alpha})$. We will be interested in finding precise formulae for Birkhoff dimension spectra with respect to the two different types of dynamical systems mentioned in §1.3.3, namely the Lalley-Gatzouras IFSs under certain overlap conditions and the fibres of a skew-product associated with a GLS IFS. For the former system, we study the sets $F_\Phi(\boldsymbol{\alpha})$ for any continuous potential Φ and appropriate vector $\boldsymbol{\alpha}$. For the latter system, we consider the specific choice of potential Φ that allows us to write the digit frequencies of digits in $\mathcal{D}$ appearing in the random GLS expansions as Birkhoff averages. In this setting, the Birkhoff average level sets are *Besicovitch-Eggleston* sets as mentioned in §1.1. Along the fibres $y \in [0, 1]$ of GLS IFSs (which we map from Ω to $[0, 1]$ for convenience in this setting), these sets are

$$F_y(\boldsymbol{\alpha}) := \pi\big(\{\xi = (\omega_n, b_n)_{n \in \mathbb{N}} \in \mathcal{D}^{\mathbb{N}} : \tau_d(\xi) = \alpha_d \, \forall d \in \mathcal{D}\}\big),$$

where $\boldsymbol{\alpha} = (\alpha_d)_{d \in \mathcal{D}}$ is a fixed probability vector, $\omega = (\omega_n)_{n \in \mathbb{N}} \mapsto y$ is the fixed fibre and π is the map $\mathcal{D}^{\mathbb{N}} \to [0, 1]^2$ that takes digit sequences to their corresponding points via the series expansions in (1.5). We call the probability vectors $\boldsymbol{\alpha}$ that define the different Besicovitch-Eggleston sets *frequency vectors.* In Chapter 3, where we study the same fibres of skew-products defined by collections of GLS maps but for GLS maps with countably many branches, it is no longer useful to consider the GLS IFS representation of the system so there is no benefit from mapping the fibre to $[0, 1]$. Consequently, the relevant Besicovitch-Eggleston sets, which are defined in the same way, are denoted by $F_\omega(\boldsymbol{\alpha})$ instead for each fixed fibre $\omega \in \Omega$.

1.4 Stable Large Deviations

Classical large deviations theory is an area of probability theory devoted to analysing the probabilistic behaviour of sums of events that deviate by an amount more than is prescribed by the classical central limit theorem (CLT). The CLT may be stated as follows. Let $(X_j)_{j\in\mathbb{N}}$ be a sequence of independent identically distributed (i.i.d.) random variables on a probability space $\big(\mathbb{R},\mathcal{B}(\mathbb{R}),\mathbb{P}\big)$, and let $\mathbb{E}X_1$ (resp. $\operatorname{Var}X_1$) denote the *expectation* (resp. *variance*) of X_1 with respect to the measure $\mathbb{P}$. Suppose that $\mathbb{E}X_1$ is finite and that $\operatorname{Var}X_1$ is finite and non-zero. Then, the partial sums $S_n \coloneqq X_1 + \cdots + X_n$ $(n\in\mathbb{N})$ satisfy

$$\frac{S_n - n\mathbb{E}X_1}{\sqrt{n\operatorname{Var}X_1}} \to Z$$

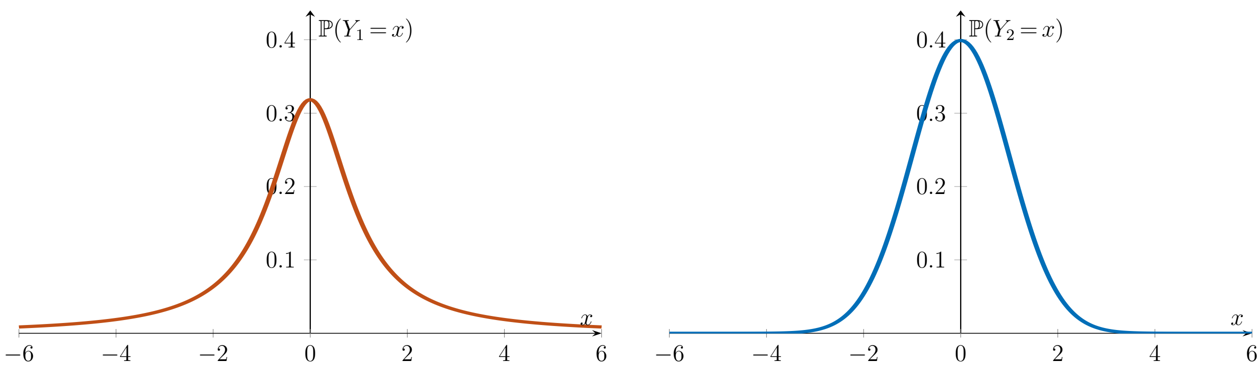
as $n\to+\infty$ in distribution with respect to $\mathbb{P}$, where $Z\sim\mathcal{N}(0,1)$ is a standard normal random variable. The CLT can be generalised further allowing for the same conclusion to be drawn even for sequences of dependent random variables such as those given by a mixing random process. In any case, the CLT determines the probability that the partial sums S_n of a sequence $(X_j)_{j\in\mathbb{N}}$ of random variables deviates from $n\mathbb{E}X_1$ by a quantity of order $\sqrt{n}$ as $n\to+\infty$. Such deviations are called normal deviations, and, unsurprisingly, a deviation of S_n from $n\mathbb{E}X_1$ by a quantity of order strictly greater than $\sqrt{n}$ is therefore called a *large deviation* (LD).

If the condition that the variance of X_1 be finite is dropped in the CLT, then the attractor random variable of the (suitably rescaled and renormalised) partial sums is a (non-normal) *stable distribution*. We say that a random variable Y has a stable distribution if any linear combination of independent Y-distributed random variables also has the same distribution as Y (up to scaling and centring). Associated with each stable distribution is a *stability parameter* $\alpha\in(0,2]$. Note that $\alpha=1$ corresponds to Y having a Cauchy distribution and $\alpha=2$ corresponds to Y having a normal distribution; see Figure 1.6.

We set some notation. Define the scaling sequence $(a_n)_{n\in\mathbb{N}}$ and centring sequence $(b_n)_{n\in\mathbb{N}}$ by

$$a_n^\alpha\big(1+o(1)\big) = \begin{cases} n\ell(a_n), & \alpha\in(0,2),\\ n\widehat{\ell}(a_n), & \alpha=2,\end{cases} \qquad b_n \coloneqq \begin{cases} 0, & \alpha\in(0,1),\\ n\mathbb{E}(X_1\mathbb{1}_{\{|X_1|\le a_n\}}), & \alpha=1,\\ n\mathbb{E}X_1, & \alpha\in(1,2],\end{cases} \tag{1.9}$$

where $\widehat{\ell}(x) \coloneqq 1+\int_1^{1+x}\ell(u)/u\,\mathrm{d}u$ with ℓ a *slowly varying function*, i.e. a non-negative $\mathbb{P}$-measurable function $\ell:(0,+\infty)\to\mathbb{R}$ with the property that $\ell(cx)/\ell(x)\to 1$ as $x\to+\infty$



(a) Standard Cauchy Distribution

(b) Standard Normal Distribution

Figure 1.6: Probability density functions of the symmetric, centred, non-skewed α-stable distribution Y_α for $\alpha=1$ and $\alpha=2$.

for any $c > 0$ (see [BGT87, §1.2.1]). For a partial sum S_n of a sequence $(X_j)_{j\in\mathbb{N}}$ of i.i.d. random variables, it is known that $(S_n - b_n)/a_n$ is in the domain of attraction of an α-stable distribution, i.e. that

$$\frac{S_n - b_n}{a_n} \to Y_\alpha \text{ in distribution as } n \to +\infty \tag{1.10}$$

for some α-stable random variable Y_α, if and only if the random variable X_1 satisfies the following condition on its tails:

$$\mathbb{P}(X_1 > x) = px^{-\alpha}\ell(x)\big(1 + o(1)\big) \quad \text{and} \quad \mathbb{P}(X_1 < -x) = qx^{-\alpha}\ell(x)\big(1 + o(1)\big), \tag{1.11}$$

where $p, q \geq 0$ with $p + q = 1$. If $p = 0$ in (1.11), then $\mathbb{P}(X_j > x) = px^{-\alpha}\ell(x)\big(1 + o(1)\big)$ is to be interpreted as $\mathbb{P}(X_j > x) = o\big(x^{-\alpha}\ell(x)\big)$, and the case of $q = 0$ in (1.11) has an analogous interpretation. Equation (1.10) can be seen as a CLT with non-standard normalisation; thus, the usual starting point for the study of large deviations in the α-stable i.i.d. setting is to assume (1.11).

Precise large deviations in the α-stable setting have been known since [Hey67; Nag69; Nag79] when $\alpha = (0, 1) \cup (1, 2)$ and only recently since [Ber19] when $\alpha = 1$. The case of $\alpha = 2$ is contained in [Roz89, Theorem 6]. We collect these results in the following theorem. Let

$$\Phi(x) := \frac{1 + o(1)}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

be the distribution function of a standard normal random variable.

Theorem 1.8 ([Hey67; Nag69; Nag79; Roz89; Ber19])**.** *Assume* (1.11), *and let* $(a_n)_{n\in\mathbb{N}}$ *and* $(b_n)_{n\in\mathbb{N}}$ *be the sequences defined in* (1.9). *Let* $N : \mathbb{N} \to (0, +\infty)$ *be such that* $N(n)/a_n \to +\infty$ *as* $n \to +\infty$. *Then the following statements hold as* $n \to +\infty$*:*

(i) *If* $\alpha \in (0, 2)$, *then*

$$\mathbb{P}(S_n - b_n > N) = n\mathbb{P}(X_1 > N)\big(1 + o(1)\big). \tag{1.12}$$

(ii) *If* $\alpha = 2$, *then*

$$\mathbb{P}(S_n - b_n > N) = \left(n\mathbb{P}(X_1 > N) + \Phi\left(\frac{N}{a_n}\right)\right)\big(1 + o(1)\big). \tag{1.13}$$

Similar statements hold for $\mathbb{P}(S_n - b_n < -N)$.

Remark 1.9. To illustrate the role of Φ, we can take for example $\ell(x) \equiv 1$ so that $\widehat{\ell}(x) = \log x$ and $N = a_n \log N\big(1 + o(1)\big)$. Then the term with $\Phi(N/a_n)$ dominates the term $n\mathbb{P}(X_1 > N)$ by approximately a factor of $\log N$.

Throughout Part II, we will be interested in large deviations in the setup of α-stable distributions when the classical CLT does not hold. In particular, we will present in Chapter 4 a new large deviations result similar to Theorem 1.8 but in the context of α-stable dependent random variables with $\alpha \in (0, 1) \cup (1, 2]$ (with infinite variance). In Chapter 5, we focus on the much harder case of $\alpha = 1$ and give error rates for large deviations in the setup of i.i.d. Cauchy distributions. We also introduce the notion of quasi-large deviations, where the range of unlikeliness considered for the partial sums S_n has a large upper bound. Central to the method of proof of these results will be the Fourier inversion formula, used to rephrase the problems in terms of characteristic functions; see e.g. [Bil12, Theorem 26.2].

Theorem 1.10 (Fourier Inversion Formula)**.** *For a random variable* X *on the probability space* $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$, *let* $\Psi_X(t) := \int_{\mathbb{R}} e^{itx}\, \mathrm{d}\mathbb{P}(X \leq x)$, $t \in \mathbb{R}$, *denote the characteristic function of* X. *If* $\mathbb{P}(\{a\}) = \mathbb{P}(\{b\}) = 0$, *then*

$$\mathbb{P}\big(X \in [a, b]\big) = \frac{1}{2\pi} \lim_{s\to+\infty} \int_{-s}^{+s} \frac{e^{-ita} - e^{-itb}}{it} \Psi_X(t)\, \mathrm{d}t.$$

1.5 Outline of Thesis

We conclude this chapter by summarising the content to follow.

Part I: Birkhoff Spectra for Autonomous and Non-autonomous Number Systems

In Chapter 2, we study Birkhoff dimension spectra for two different types of dynamical systems and different families of Birkhoff potentials. The first type of system is the family of IFSs known as Lalley-Gatzouras carpets, which we study for all continuous Birkhoff potentials. We determine the set of parameters for which the Birkhoff level sets are non-empty and, for such parameters, we compute the Birkhoff dimension spectrum under one of two additional assumptions on the way in which the functions of the Lalley-Gatzouras IFS overlap. This result gives the same formula as in [BJKR21] but under different conditions. This result is also a subcase of [Ree11], but we are able to obtain more information given our additional assumptions. The main difficulty in this (two-dimensional) setting is that the systems we treat are *non-conformal*, i.e. the contraction ratios in the relevant self-affine IFSs differ in the two stable directions, which means that many of the usual methods used in fractal geometry do not apply.

The second type of dynamical system that we consider in Chapter 2 is the non-autonomous system given by a family of GLS IFSs used to generate GLS expansions with redundancy. Specifically, we consider these systems along fixed fibres. For the choice of potential, we take the one that captures the digit frequencies of digits appearing in the GLS IFS expansions. The resulting Birkhoff average level sets are then a family of Besicovitch-Eggleston sets i.e. sets of points whose digit frequencies are given by some specified frequency vector. As in the setting above, we determine the set of parameters for which the Birkhoff level sets are non-empty and, for such parameters, compute the Birkhoff dimension spectrum. As a corollary, we may use the Birkhoff dimension spectrum along the fibres to deduce a lower bound for the Hausdorff dimension of the Besicovitch-Eggleston sets for the full system. In contrast to the first setting, this non-autonomous setting is conformal so the standard methods of fractal geometry may be used; however, non-autonomy means that fundamental results such as the Birkhoff Ergodic Theorem cannot be used.

In Chapter 3, we extend the results in Chapter 2 on the non-autonomous number systems by dropping the condition that the interval maps giving each number system have only finitely many branches. As before, we determine the set of parameters for which the Besicovitch-Eggleston sets of these more general systems are non-empty and, for such parameters, compute the Birkhoff dimension spectrum. To do this, we make a minor additional assumption on the system that is satisfied for a wide variety of systems. The result that we obtain contains the deterministic setting [FLMW10, Theorem 1.1] as a subcase. Since the underlying symbolic dynamics of the system now operates on an infinite alphabet, many of the arguments used in the finite setting of Chapter 2 fail to carry over. Consequently, we do not rely on the methods in Chapter 2 but instead adapt some of the techniques used in the deterministic setting [FLMW10], which also operates on an infinite alphabet.

Part II: Stable Large Deviations for Deterministic Dynamical Systems

In Chapter 4, we will be interested in obtaining a result similar in spirit to Theorem 1.8 for dependent sequences of dynamically defined random variables. Current results in the literature that attempt this either fail to capture the full range of large deviations as in Theorem 1.8 or have conditions that are extremely difficult to verify for dynamical systems; see §4.2 for a brief discussion of some of these results. In contrast, the generalisation of Theorem 1.8 for dependent sequences that we obtain in Chapter 4, namely Theorem 4.11, has natural conditions known to be satisfied for a broad class of (autonomous) dynamical systems and treats the full range of large deviations, giving the optimal behaviour in the majority of this range; Theorem 4.11 holds for all stability parameters $\alpha \in (0,1) \cup (1,2]$. We obtain these results by using the Fourier inversion formula to rephrase the problem in terms of characteristic functions and then by applying Nagaev's spectral method (see e.g. [Gou10b]), allowing us to write the characteristic function in terms of a related characteristic function for an i.i.d. sequence of random variables plus some well-behaved quantities that we may obtain good estimates for. The results may be applied to the class of Gibbs-Markov dynamical systems, including the continued fraction map, allowing us to obtain a statement on the large deviations of the Hölder mean of continued fraction digits, as well as other dynamical systems such as AFU maps; see §4.4.

In Chapter 5, we investigate the specific case of $\alpha = 1$ in the above setting, i.e. when the partial sums are in the domain of attraction of a Cauchy distribution. This case is much harder due to the inherent pathological behaviour of the Cauchy distribution and means that results of the generality of those in Chapter 4 are very difficult to obtain in this setting. Instead, we obtain error rates for the large deviations in Theorem 1.8 of Cauchy i.i.d. random variables over the majority of the range; see Theorem 5.1. We also obtain error rates for quasi-large deviations of Cauchy i.i.d. random variables over the full range of large deviations seen in Theorem 1.8; see Theorem 5.2. To do this, we use the same method as in Chapter 4 (when treating i.i.d. sequences), but we adapt the techniques to work in the $\alpha = 1$ setting and so as to give error rates. We then apply the above results to find the error rates for the large and quasi-large deviations of the arithmetic mean of digits coming from Lüroth series expansions and other GLS expansions.

Part I

Birkhoff Spectra for Autonomous and Non-autonomous Number Systems

Chapter 2

Birkhoff Spectra for Certain Diagonally Self-affine Sets

The results in this chapter are based on the article [IKM24]:

Birkhoff Spectrum for Diagonally Self-affine Sets and Besicovitch-Eggleston Sets for GLS Systems with Redundancy

Abstract

We calculate the Birkhoff spectrum in terms of the Hausdorff dimension of level sets for Birkhoff averages of continuous potentials for a certain family of diagonally affine IFSs. Also, we study Besicovitch-Eggleston sets for finite GLS number systems with redundancy. The redundancy refers to the fact that each number $x \in [0,1]$ has uncountably many expansions in the system. We determine the Hausdorff dimension of digit frequency sets for such expansions along fibres.

2.1 Introduction

We consider affine IFSs $\{\mathbf{A}_u + \mathbf{v}_u\}_{u\in\mathcal{U}}$ indexed by a finite set $\mathcal{U}$ that are given by a collection $\{\mathbf{A}_u : u \in \mathcal{U}\} \in \mathrm{GL}_2(\mathbb{R})^{\#\mathcal{U}}$ of matrices of the form

$$\mathbf{A}_u = \begin{bmatrix} b_u & 0 \\ 0 & c_u \end{bmatrix}, \quad \text{with } 0 < |b_u|, |c_u| < 1,\ u \in \mathcal{U}, \tag{2.1}$$

and a collection $(\mathbf{v}_u)_{u\in\mathcal{U}}$ of vectors of the form

$$\mathbf{v}_u = \begin{bmatrix} \beta_u \\ \gamma_u \end{bmatrix} \in \mathbb{R}^2, \quad u \in \mathcal{U}.$$

For a fixed IFS as above, we will be interested in the multifractal properties of its self-affine attractor Λ. Let $\sigma : \mathcal{U}^{\mathbb{N}} \to \mathcal{U}^{\mathbb{N}}$ denote the left shift. For any continuous potential $\Phi : \mathcal{U}^{\mathbb{N}} \to \mathbb{R}^d$, $d \in \mathbb{N}$, and for a given vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, consider the symbolic α-level sets

$$E_\Phi(\boldsymbol{\alpha}) := \left\{ \xi \in \mathcal{U}^{\mathbb{N}} : \lim_{n\to+\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Phi \circ \sigma^i(\xi) = \boldsymbol{\alpha} \right\}. \tag{2.2}$$

For $\xi \in \mathcal{U}^{\mathbb{N}}$ and $n \in \mathbb{N}$, put $\mathbf{A}_{\xi|_{n-1}} = \mathbf{A}_{\xi_1} \cdots \mathbf{A}_{\xi_{n-1}}$ and define $\pi : \mathcal{U}^{\mathbb{N}} \to \Lambda$ by

$$\pi(\xi) = \sum_{n\in\mathbb{N}} \mathbf{A}_{\xi|_{n-1}} \mathbf{v}_{\xi_n}. \tag{2.3}$$

Then the $\boldsymbol{\alpha}$-*level set* for Φ on Λ is the set $F_\Phi(\boldsymbol{\alpha}) := \pi(E_\Phi(\boldsymbol{\alpha}))$. There are various known results on the dimension spectra of $F_\Phi(\boldsymbol{\alpha})$, i.e. the size of level sets with respect to dimension quantities such as the Hausdorff dimension or topological entropy (in the sense of [Bow73]); see for instance [BS01; FFW01; BSS02a; BSS02b; BM08; FH10; Ree11; KR14; Moh22; Moh24]. Multifractal results for self-affine sets in terms of Lyapunov dimensions were obtained in [BJKR21] for collections of matrices $(\mathbf{A}_u)_{u\in\mathcal{U}}$ under certain strong-irreducibility and proximality conditions. One can find more information about the multifractal formalism in [BG11; Cli14; Moh23]. In this chapter, we will only treat the *Hausdorff dimension spectra* $\boldsymbol{\alpha} \mapsto \dim_H F_\Phi(\boldsymbol{\alpha})$, where the size of our level sets is quantified by the Hausdorff dimension $\dim_H$.

Let $\mathcal{I} = \{f_u : \mathbb{R} \to \mathbb{R}\}_{u\in\mathcal{U}}$ be an IFS of real-valued affine maps of the form $f_u(y) = d_u y + \delta_u$, so $|d_u| < 1$ for each $u \in \mathcal{U}$. For a sequence $\mathbf{u} = u_1 \cdots u_n \in \mathcal{U}^n$, $n \in \mathbb{N}$, we write

$$f_{\mathbf{u}}(y) := f_{u_1} \circ \cdots \circ f_{u_n}(y) = d_{\mathbf{u}} y + \delta_{\mathbf{u}}.$$

We say that $\mathcal{I}$ has *exact overlaps* if there are $\mathbf{u}, \mathbf{u}' \in \mathcal{U}^n$ for some $n \in \mathbb{N}$ such that $f_{\mathbf{u}} = f_{\mathbf{u}'}$. For $\mathbf{u}, \mathbf{u}' \in \mathcal{U}^n$, we define the distance

$$\operatorname{dist}(f_{\mathbf{u}}, f_{\mathbf{u}'}) := \begin{cases} |\delta_{\mathbf{u}} - \delta_{\mathbf{u}'}|, & \text{if } d_{\mathbf{u}} = d_{\mathbf{u}'}; \\ +\infty, & \text{otherwise.} \end{cases}$$

In his breakthrough result, Hochman [Hoc14] introduced the *Exponential Separation Condition* (ESC) to calculate the dimension of self-similar measures, i.e. measures associated to the IFS $\mathcal{I}$ that are of the form

$$\mu(A) = \sum_{i\in\Lambda} p_i \mu \circ f_i^{-1}(A), \qquad A \in \mathcal{B}(\mathbb{R}^d),$$

for some fixed probability vector $(p_i)_{i\in\Lambda}$, where Λ is the attractor of $\mathcal{I}$. We say that $\mathcal{I}$ satisfies the Exponential Separation Condition if there exists a constant $c > 0$ and infinitely many integers $n \in \mathbb{N}$ such that $\operatorname{dist}(f_{\mathbf{u}}, f_{\mathbf{u}'}) \geq c^n$ for all $\mathbf{u}, \mathbf{u}' \in \mathcal{U}^n$.

We say that a diagonally affine IFS $\{\mathbf{A}_u + \mathbf{v}_u\}_{u\in\mathcal{U}}$ satisfies the *Strong Open Set Condition* (SOSC) if there is an open set $V \subseteq \mathbb{R}^2$ such that all the sets $(\mathbf{A}_u + \mathbf{v}_u)(V)$ are pairwise disjoint, $\bigcup_{u\in\mathcal{U}}(\mathbf{A}_u + \mathbf{v}_u)(V) \subseteq V$ and $\Lambda \cap V \neq \emptyset$, where Λ is the attractor of the IFS.

Our first result is for the following class of diagonally affine IFSs. Let $\mathcal{G}$ be the collection of all IFSs $\{\mathbf{A}_u + \mathbf{v}_u\}_{u\in\mathcal{U}}$ with matrices as in (2.1) that satisfy the SOSC together with either (D) or (D'):

(D) $|b_u| > |c_u|$ for all $u \in \mathcal{U}$ and

- (a) the IFS obtained from projecting to the first coordinate $\mathcal{G}_1 := \{g_{1,u}(y) = b_u y + \beta_u\}_{u\in\mathcal{U}}$ satisfies the ESC, or,
- (b) b_u is algebraic for all $u \in \mathcal{U}$ and the IFS obtained by projecting to the first coordinate $\mathcal{G}_1 := \{g_{1,u}(y) = b_u y + \beta_u\}_{u\in\mathcal{U}}$ has no exact overlaps;

(D') $|b_u| < |c_u|$ for all $u \in \mathcal{U}$ and

- (a) the IFS obtained by projecting to the second coordinate $\mathcal{G}_2 := \{g_{2,u}(y) = c_u y + \gamma_u\}_{u\in\mathcal{U}}$ satisfies the ESC, or,
- (b) c_u is algebraic for all $u \in \mathcal{U}$ and the IFS obtained from projecting to the second coordinate $\mathcal{G}_2 := \{g_{2,u}(y) = c_u y + \gamma_u\}_{u\in\mathcal{U}}$ has no exact overlaps.

An IFS satisfying condition (D')(a) is shown in Figure 2.1(a).

Let $\mathcal{M}(\mathcal{U}^{\mathbb{N}},\sigma)$ denote the set of all σ-invariant Borel probability measures on $\mathcal{U}^{\mathbb{N}}$, and let

$$\begin{aligned} L_\Phi &:= \left\{ \boldsymbol{\alpha} \in \mathbb{R}^d : \exists\, \xi \in \mathcal{U}^{\mathbb{N}} \text{ with } \lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \Phi\circ\sigma^i(\xi) = \boldsymbol{\alpha} \right\} \\ &= \left\{ \boldsymbol{\alpha} \in \mathbb{R}^d : \exists\, \mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}},\sigma) \text{ with } \int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\mu = \boldsymbol{\alpha} \right\} \end{aligned} \tag{2.4}$$

be the collection of all vectors $\boldsymbol{\alpha}$ for which the corresponding level set is non-empty, known as the *spectrum of* Φ. We use $\operatorname{int} L_\Phi$ to denote the interior of L_Φ. Let P denote the topological pressure, and let $\dim_L \mu$ denote the Lyapunov dimension of $\mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}},\sigma)$; see §2.2 for definitions. Our first result is as follows.

2

Theorem 2.1. *Let* $\{\mathbf{A}_u + \mathbf{v}_u\}_{u\in\mathcal{U}} \in \mathcal{G}$*, and let* $\Phi : \mathcal{U}^{\mathbb{N}} \to \mathbb{R}^d$*,* $d \in \mathbb{N}$*, be a continuous potential. Then for each* $\boldsymbol{\alpha} \in \operatorname{int} L_\Phi$*,*

$$\begin{aligned} \dim_H F_\Phi(\boldsymbol{\alpha}) &= \sup\left\{ \dim_L \mu : \mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}},\sigma),\ \int_{\mathcal{U}^{\mathbb{N}}} \Phi\,\mathrm{d}\mu = \boldsymbol{\alpha} \right\} \\ &= \sup\left\{ s \geq 0 : \inf_{\mathbf{q}\in\mathbb{R}^d} P(\log\phi^s + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle) \geq 0 \right\}. \end{aligned}$$

The second family of IFSs that we consider in this chapter is motivated by the representations of real numbers known as *Generalised Lüroth Series* (GLS) *expansions* introduced in [BBDK96]. Recall from §1.3.1 that, whilst Lüroth series expansions take their digits from the infinite digit set $\mathbb{N}$ and all terms in the expansion are positive, a GLS number system can have either finite or infinite digit sets and the corresponding GLS expansions can have both positive and negative terms. Given a finite or countably infinite digit set $\mathcal{B}$, a partition $\{[\ell_b, r_b]\}_{b\in\mathcal{B}}$ of $[0,1]$ into closed intervals and a vector $(\varepsilon_b)_{b\in\mathcal{B}} \in \{0,1\}^{\#\mathcal{B}}$, one can consider the IFS

$$\left\{ g_b : [0,1] \to [0,1] \right\}_{b\in\mathcal{B}}, \tag{2.5}$$

where g_b maps the interval $[0,1]$ affinely onto $[\ell_b, r_b]$ in an orientation-preserving manner if $\varepsilon_b = 0$ and in an orientation-reversing manner if $\varepsilon_b = 1$. In other words, if we write $N_b := (r_b - \ell_b)^{-1}$, then $g_b(x) = \ell_b + \frac{\varepsilon_b + (-1)^{\varepsilon_b} x}{N_b}$. Since $g_b([0,1]) = [\ell_b, r_b]$ for each $b \in \mathcal{B}$, it follows for each $x \in [0,1]$ that there is a sequence $(b_n)_{n\in\mathbb{N}}$ such that

$$x = \lim_{n\to+\infty} g_{b_1} \circ \cdots \circ g_{b_n}(0).$$

In this notation, x can thus be expressed as

$$x = \sum_{n\in\mathbb{N}} (-1)^{\sum_{i=1}^{n-1} \varepsilon_{b_i}} \frac{\ell_{b_n} N_{b_n} + \varepsilon_{b_n}}{\prod_{i=1}^{n} N_{b_i}}, \tag{2.6}$$

which is the *GLS expansion* of x with digit set $\mathcal{B}$ from (1.3); here, we have set $\sum_{i=1}^{0} \varepsilon_{b_i} = 0$. One recovers the Lüroth series expansions by taking $\mathcal{B} = \mathbb{N}$, $[\ell_b, r_b] = [\frac{1}{b+1}, \frac{1}{b}]$ and $\varepsilon_b = 0$ for each $b \in \mathbb{N}$, and one obtains integer base N-expansions, $N \geq 2$, by setting $\mathcal{B} = \{0, 1, \ldots, N-1\}$ and taking $[\ell_b, r_b] = [\frac{b}{N}, \frac{b+1}{N}]$ and $\varepsilon_b = 0$. The expansions from (2.6) can also be seen as signed versions of Cantor base expansions, as introduced by Cantor in [Can69]. GLS expansions have been considered previously in [Mun11; Arr15; KKSS15; JMS18; HK24; BK24] and recently also in relation to neural networks [BKSN19; RHN23]. Level sets for Lüroth series expansions

(a) An IFS satisfying (D')(a)

(b) Example of a GLS IFS

Figure 2.1: Two examples of IFSs. The coloured rectangles indicate the images of the unit square under the maps in the IFS.

and more generally GLS expansions have been considered in particular with respect to digit frequencies; see [BI09; FLMW10]. These are the *Besicovitch-Eggleston sets* for such systems.

In the above setting, for any given GLS number system, all but countably many numbers in $[0, 1]$ have a unique GLS expansion in that system and the numbers that do not have a unique expansion have exactly two expansions. In this chapter, we consider IFSs that correspond to GLS number systems with redundancy, that is, in which all numbers have uncountably many different representations in the system. Number systems with redundancy have proven interesting in several settings, including signed binary expansions where they are used to find so-called minimal weight expansions, i.e. expansions that maximise the number of digits 0, see e.g. [MO90; KT93; DK20], and in non-integer base expansions in relation to applications in analogue-to-digital converters and random number generation, see e.g. [DDGV06; JM16]. Number systems with redundancy have also been considered in [KKV17; KMTV22; KM22b] for continued fraction expansions and Lüroth series expansions.

To obtain a GLS number system with redundancy, we take a finite index set $\mathcal{S} = \{1, \ldots, S\} \subset \mathbb{N}$ and start with S IFSs that correspond to S different GLS number systems with finite digit sets. We combine these into one diagonally affine IFS on $\mathbb{R}^2$, which we call a *GLS IFS*, by using a positive probability vector $(p_s)_{s\in\mathcal{S}}$ – a vector with $p_s > 0$ for all $s \in \mathcal{S}$ and $\sum_{s\in\mathcal{S}} p_s = 1$. This vector $(p_s)_{s\in\mathcal{S}}$ can be thought of as the probabilities with which the s^{th} GLS number system is chosen to generate the n^{th} digit in the expansions for each $n \in \mathbb{N}$. Therefore, a GLS IFS is given by the following data:

- an index set $\mathcal{S} \subset \mathbb{N}$ and a positive probability vector $(p_s)_{s\in\mathcal{S}}$
- for each $s \in \mathcal{S}$, a branch-indexing set $\mathcal{B}_s := \{1, \ldots, B_s\} \subset \mathbb{N}$ with $B_s \geq 2$, a partition $0 = r_{(s,1)} < r_{(s,2)} < \cdots < r_{(s,B_s)} = 1$ and a vector $(\varepsilon_{(s,b)})_{b\in\mathcal{B}_s} \in \{0, 1\}^{B_s}$

If we set $\mathcal{D} := \{(s, b) : b \in \mathcal{B}_s,\ s \in \mathcal{S}\}$ and, for each $(s, b) \in \mathcal{D}$, define $q_{(s,b)} := r_{(s,b+1)} - r_{(s,b)}$ and

$$\mathbf{A}_{(s,b)} := \begin{bmatrix} (-1)^{\varepsilon_{(s,b)}} q_{(s,b)} & 0 \\ 0 & p_s \end{bmatrix}, \quad \mathbf{v}_{(s,b)} := \begin{bmatrix} r_{(s,b)} + \varepsilon_{(s,b)} q_{(s,b)} \\ \sum_{1\leq \ell < s} p_\ell \end{bmatrix},$$

then we call the IFS $\{\mathbf{A}_d + \mathbf{v}_d\}_{d\in\mathcal{D}}$ a GLS IFS; see Figure 2.1(b) for an example.

We mention a few particular properties of GLS IFSs. Each GLS IFS satisfies the strong open set condition and has $\Lambda = [0, 1]^2$. For the projection onto the first coordinate, we use the collection of maps $\mathcal{G}_1 = \{h_d : [0, 1] \to [0, 1]\}_{d\in\mathcal{D}}$, where, for each $d \in \mathcal{D}$,

$$h_d(x) := r_d + q_d(\varepsilon_d + (-1)^{\varepsilon_d} x).$$

Without additional assumptions, the GLS IFS need not fall into one of the categories (D)(a) or (D)(b). The projection on the second coordinate $\mathcal{G}_2 = \{g_{2,(s,b)}(y) = p_s y + \sum_{1\leq \ell < s} p_\ell\}_{(s,b)\in\mathcal{D}}$ of a GLS IFS contains several duplicates of each map. Therefore, GLS IFSs do not fall into the class of diagonally affine IFSs that satisfy (D'), but, by removing these duplicates, they can potentially contain a sub-collection of contractions that satisfies (D') as shown in Figure 2.1.

We can obtain number expansions from a GLS IFS in the following way. For each $x \in [0,1]$, there are sequences $(d_m)_{m\in\mathbb{N}} \in \mathcal{D}^{\mathbb{N}}$ such that x can be written as

$$x = \lim_{m\to+\infty} h_{d_1} \circ \cdots \circ h_{d_m}(0). \tag{2.7}$$

If we write for each $n \in \mathbb{N}$

$$\epsilon_n = \varepsilon_{d_n}, \quad N_n = q_{d_n}^{-1}, \quad t_n = r_{d_n} + \varepsilon_{d_n} q_{d_n}, \tag{2.8}$$

then it follows from (2.7) that

$$x = \sum_{n\in\mathbb{N}} (-1)^{\sum_{1\leq \ell<n} \epsilon_\ell} \frac{t_n}{\prod_{1\leq \ell\leq n} N_\ell} \tag{2.9}$$

and the resemblance with Lüroth series expansions becomes clear. It is shown in Proposition 2.7 below that, under the additional assumption on the GLS IFS that $h_d \neq h_{d'}$ whenever $d \neq d'$, indeed all numbers $x \in [0,1]$ have uncountably many different representations of the form (2.7). We give several examples of GLS IFSs and the associated number expansions at the end of the chapter.

For GLS IFSs, we consider the potential that captures digit frequencies. Let $\mathbb{1}_{[d]} : \mathcal{D}^{\mathbb{N}} \to \{0,1\}$ denote the indicator function on $[d]$. Define the continuous potential $\mathbb{1} : \mathcal{D}^{\mathbb{N}} \to \{0,1\}^{\#\mathcal{D}}$ by $\mathbb{1}(\xi) = (\mathbb{1}_{[d]}(\xi))_{d\in\mathcal{D}}$. Fix a frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}} \in [0,1]^{\#\mathcal{D}}$, and set $F(\boldsymbol{\alpha}) := \pi(E_{\mathbb{1}}(\boldsymbol{\alpha}))$. Then

$$F(\boldsymbol{\alpha}) = \{(x,y) \in [0,1]^2 : \exists\, \xi \in \pi^{-1}\{(x,y)\} \text{ s.t. } \tau_d(\xi) = \alpha_d \text{ for all } d \in \mathcal{D}\} \tag{2.10}$$

is the *GLS digit frequency level set* or *Besicovitch-Eggleston set* for $\boldsymbol{\alpha}$. Results on the Hausdorff dimension spectrum $\boldsymbol{\alpha} \mapsto \dim_H F(\boldsymbol{\alpha})$ have been obtained in [Nie99, Theorem 1] in the specific case of Bedford-McMullen carpets, i.e. with $p_s = S^{-1}$ for each $s \in \mathcal{S}$ and $(-1)^{\varepsilon_d} = 1$ and $q_d = N^{-1}$ for each $d \in \mathcal{D}$, where $N \geq 2$ is some fixed integer. This result was extended in [Ree11, Corollary 1] to Lalley-Gatzouras carpets, which are similar to our setting but have the additional requirements that $\varepsilon_{(s,b)} = 0$ and $q_{(s,b)} \leq p_s$ for all $(s,b) \in \mathcal{D}$. In the current setting, a lower bound for $\dim_H F(\boldsymbol{\alpha})$ in terms of the Ledrappier-Young formula for the Bernoulli measure $\mu_{\boldsymbol{\alpha}}$ can be deduced from [BK17, Theorem 2.3 and Corollary 2.8] in the case that the two Lyapunov exponents of the Bernoulli measure of the system differ and the frequency vector $\boldsymbol{\alpha}$ is strictly positive. In the case that the two Lyapunov exponents of the Bernoulli measure $\mu_{\boldsymbol{\alpha}}$ are equal, one may apply [FH09] to obtain a similar lower bound in terms of the Ledrappier-Young formula for $\mu_{\boldsymbol{\alpha}}$.

Here, we will instead focus for fixed $y \in [0,1]$ on the *fibre Besicovitch-Eggleston sets*

$$F_y(\boldsymbol{\alpha}) := \left\{x \in [0,1] : \exists\, \xi \in \pi^{-1}\{(x,y)\} \text{ s.t. } \tau_d(\xi) = \alpha_d \text{ for all } d \in \mathcal{D}\right\},$$

recalling from §1.3.2 that $\tau_d(\xi)$ denotes the frequency of occurrence of the digit d in the sequence ξ. We only consider frequency vectors $\boldsymbol{\alpha}$ with $\alpha_s := \sum_{b\in\mathcal{B}_s} \alpha_{(s,b)} > 0$ for all $s \in \mathcal{S}$ (otherwise we could just as well have considered a smaller GLS IFS). Set

$$Y(\boldsymbol{\alpha}) := \{y \in [0,1] : F_y(\boldsymbol{\alpha}) \neq \emptyset\}. \tag{2.11}$$

Let $\mu_{\boldsymbol{\alpha}}$ be the $\boldsymbol{\alpha}$-Bernoulli measure on $\mathcal{D}^{\mathbb{N}}$. For each $s \in \mathcal{S}$, let $f_s : [0,1] \to [0,1]$ be the map given by

$$f_s(y) = p_s y + \sum_{1 \le \ell < s} p_\ell \tag{2.12}$$

and define the map $\pi_2 : \mathcal{D}^{\mathbb{N}} \to [0,1]$ by

$$((\omega_n, b_n))_{n\in\mathbb{N}} \mapsto \lim_{n\to+\infty} f_{\omega_1} \circ \cdots \circ f_{\omega_n}(0).$$

Set $\nu_{\boldsymbol{\alpha}} := \mu_{\boldsymbol{\alpha}} \circ \pi_2^{-1}$. As we will see later, $\nu_{\boldsymbol{\alpha}}(Y(\boldsymbol{\alpha})) = 1$. We have the following results on the Hausdorff dimension of the fibre Besicovitch-Eggleston sets.

2

Theorem 2.2. *Let $\{\mathbf{A}_d + \mathbf{v}_d\}_{d\in\mathcal{D}}$ be a GLS IFS and $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}} \in [0,1]^{\#\mathcal{D}}$ a frequency vector. Then*

$$\dim_H F_y(\boldsymbol{\alpha}) \le \frac{\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d - \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{\sum_{d\in\mathcal{D}} \alpha_d \log q_d}$$

for all $y \in Y(\boldsymbol{\alpha})$. In addition,

$$\dim_H F_y(\boldsymbol{\alpha}) \ge \frac{\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d - \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{\sum_{d\in\mathcal{D}} \alpha_d \log q_d}$$

for $\nu_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$.

Fibrewise results similar in spirit to Theorem 2.2 were obtained in [NT24], where the authors study real numbers with a semi-regular continued fraction expansion that satisfies a certain growth condition on its digits.

This chapter is outlined as follows. In §2.2, we provide the necessary preliminaries. We prove Theorem 2.1 in §2.3. §2.4 is devoted to GLS IFSs. Here, we show given a GLS IFS that has $h_d \neq h_{d'}$ whenever $d \neq d'$ that all $x \in [0,1]$ have uncountably many expansions of the form (2.7). We then continue with some results on the spectra of the Besicovitch-Eggleston sets $F(\boldsymbol{\alpha})$ and on the sets $Y(\boldsymbol{\alpha})$ with respect to the appropriate frequency potentials. These results will be used in the proof of Theorem 2.2. This section also contains the proof of Theorem 2.2. Finally, §2.5 contains some examples.

2.2 Preliminaries

In this section, we introduce some notation and definitions to be used in the rest of the chapter.

Let $(\mathbf{A}_u)_{u\in\mathcal{U}} \in \mathrm{GL}_2(\mathbb{R})^{\#\mathcal{U}}$ be a collection of matrices as in (2.1). Recall that, for a sequence $\xi = (\xi_n)_{n\in\mathbb{N}} \in \mathcal{U}^{\mathbb{N}}$ and $n \in \mathbb{N}$, we set $\mathbf{A}_{\xi|_n} = \mathbf{A}_{\xi_1} \cdots \mathbf{A}_{\xi_n}$. For the entries on the diagonal of $\mathbf{A}_{\xi|_n}$, write $b_{\xi|_n} = b_{\xi_1} \cdots b_{\xi_n}$ and $c_{\xi|_n} = c_{\xi_1} \cdots c_{\xi_n}$. For $\mathbf{u} = u_1 \cdots u_n \in \mathcal{U}^n$, we similarly write $\mathbf{A}_{\mathbf{u}} = \mathbf{A}_{u_1} \cdots \mathbf{A}_{u_n}$ with $b_{\mathbf{u}} = b_{u_1} \cdots b_{u_n}$ and $c_{\mathbf{u}} = c_{u_1} \cdots c_{u_n}$ for the diagonal entries.

Let $\mathbb{P}^1$ be the real projective line, which is the set of all lines through the origin in $\mathbb{R}^2$. We say that a proper subset $\mathcal{C} \subset \mathbb{P}^1$ is a *cone* if it is a closed projective interval and a *multicone* if it is a finite union of cones. The collection $(\mathbf{A}_u)_{u\in\mathcal{U}}$ of diagonal matrices as in (2.1) is called *dominated* if there exists a multicone $\mathcal{C} \subset \mathbb{P}^1$ such that $\bigcup_{u\in\mathcal{U}} \mathbf{A}_u \mathcal{C} \subset \operatorname{int} \mathcal{C}$. It was shown in [BG09, Theorem B] that $(\mathbf{A}_u)_{u\in\mathcal{U}}$ is dominated if and only if there exist constants $C > 0$ and $0 < \tau < 1$ such that

$$\frac{b_{\mathbf{u}} \cdot c_{\mathbf{u}}}{\max\{b_{\mathbf{u}}, c_{\mathbf{u}}\}^2} \le C\tau^n, \qquad \forall \mathbf{u} \in \mathcal{U}^n,\ n \in \mathbb{N}. \tag{2.13}$$

For each diagonal matrix $\mathbf{A}_u$ as in (2.1), the *singular value function* is given by

$$\phi^s(\mathbf{A}_u) := \begin{cases} \max\{b_u, c_u\}^s, & \text{if } 0 \le s \le 1, \\ \max\{b_u, c_u\} \min\{b_u, c_u\}^{s-1}, & \text{if } 1 \le s < 2. \end{cases}$$

The *Lyapunov exponents* of the collection $(\mathbf{A}_u)_{u\in\mathcal{U}}$ with respect to a measure $\mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma)$ are defined as

$$\chi_1(\mu) := -\lim_{n\to+\infty} \frac{1}{n} \int_{\mathcal{U}^{\mathbb{N}}} \log \max\{b_{\xi|_n}, c_{\xi|_n}\} \,\mathrm{d}\mu(\xi),$$
$$\chi_2(\mu) := -\lim_{n\to+\infty} \frac{1}{n} \int_{\mathcal{U}^{\mathbb{N}}} \log \min\{b_{\xi|_n}, c_{\xi|_n}\} \,\mathrm{d}\mu(\xi).$$

The *Lyapunov dimension* of $\mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma)$ is defined to be

$$\dim_L \mu := \min\left\{\frac{h_\mu(\sigma)}{\chi_1(\mu)}, 1 + \frac{h_\mu(\sigma) - \chi_1(\mu)}{\chi_2(\mu)}\right\}.$$

For a continuous potential $\Phi : \mathcal{U}^{\mathbb{N}} \to \mathbb{R}^d$, $d \in \mathbb{N}$, write $S_n\Phi = \sum_{k=0}^{n-1} \Phi \circ \sigma^k$ for its *Birkhoff sum*. The *topological pressure* of Φ and $(\mathbf{A}_u)_{u\in\mathcal{U}}$ is given by

$$P(\log \phi^s + \Phi) = \lim_{n\to+\infty} \frac{1}{n} \log \sum_{\mathbf{u}\in\mathcal{U}^n} \phi^s(\mathbf{A}_{\mathbf{u}}) \sup_{\xi\in[\mathbf{u}]} \exp S_n\Phi(\xi),$$

where the existence of the limit is guaranteed by sub-additivity.

For any subset $U \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, recall that we use the notation $|U| = \sup\{|x - y| : x, y \in U\}$ to denote the diameter of U. In many cases, U will be an interval so $|U|$ will just be its length.

2.3 Dominated, Diagonally Affine IFSs

In this section, we prove Theorem 2.1. Recall the definition of the natural projection $\pi : \mathcal{U}^{\mathbb{N}} \to \Lambda$ from (2.3) and the definitions of the sets $E_\Phi(\boldsymbol{\alpha})$ and L_Φ from (2.2) and (2.4), respectively. Additionally recall that $F_\Phi(\boldsymbol{\alpha}) := \pi(E_\Phi(\boldsymbol{\alpha}))$. We have the following upper bound for the Hausdorff dimension of $F_\Phi(\boldsymbol{\alpha})$. In the proof, we make use of [BJKR21, Proposition 3.2], which holds for general affine IFSs (including the diagonally affine case) on $\mathbb{R}^2$.

Lemma 2.3. *Let $\{\mathbf{A}_u + \mathbf{v}_u\}_{u\in\mathcal{U}}$ be a diagonally affine IFS on $\mathbb{R}^2$ such that the collection $(\mathbf{A}_u)_{u\in\mathcal{U}}$ is dominated. Let $\Phi : \mathcal{U}^{\mathbb{N}} \to \mathbb{R}^d$, $d \in \mathbb{N}$, be a continuous potential. Then for each $\boldsymbol{\alpha} \in \operatorname{int} L_\Phi$,*

$$\dim_H F_\Phi(\boldsymbol{\alpha}) \le \sup\left\{ \dim_L \mu : \mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma), \int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\mu = \boldsymbol{\alpha} \right\}$$
$$= \sup\left\{ s \ge 0 : \inf_{\mathbf{q}\in\mathbb{R}^d} P\big(\log \phi^s + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle\big) \ge 0 \right\}.$$

Proof. It follows directly from [BJKR21, Lemma 3.1 and Proposition 3.2] for any diagonally affine IFS $\{\mathbf{A}_u + \mathbf{v}_u\}_{u\in\mathcal{U}}$ on $\mathbb{R}^2$, any continuous potential $\Phi : \mathcal{U}^{\mathbb{N}} \to \mathbb{R}^d$, $d \in \mathbb{N}$, and any $\boldsymbol{\alpha} \in \operatorname{int} L_\Phi$ that

$$\dim_H F_\Phi(\boldsymbol{\alpha}) \le \sup\left\{ s \ge 0 : \inf_{\mathbf{q}\in\mathbb{R}^d} P\big(\log \phi^s + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle\big) \ge 0 \right\}. \tag{2.14}$$

2

A measure $\nu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma^n)$ is called an *n-step Bernoulli measure* if it is a Bernoulli measure on $(\mathcal{U}^{\mathbb{N}}, \sigma^n)$. For n-step Bernoulli measures $\nu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma^n)$, write

$$\tilde{\nu} = \frac{1}{n}\sum_{k=0}^{n-1} \nu \circ \sigma^{-k}. \tag{2.15}$$

Then $\tilde{\nu} \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma)$ and $\tilde{\nu}$ is ergodic. Since $(\mathbf{A}_u)_{u\in\mathcal{U}}$ is dominated, it follows by [BJKR21, Proposition 4.3] and (2.14) that for any $\boldsymbol{\alpha} \in \operatorname{int} L_\Phi$,

$$\begin{aligned}
&\sup\Big\{ s \geq 0 : \inf_{\mathbf{q}\in\mathbb{R}^d} P\big(\log\phi^s + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle\big) \geq 0 \Big\} \\
&\quad \leq \sup\left\{ \dim_L \tilde{\nu} : \nu \text{ fully supported } n\text{-step Bernoulli and } \int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\tilde{\nu} = \boldsymbol{\alpha} \right\} \\
&\quad \leq \sup\Big\{ \dim_L \mu : \mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma) \text{ and } \int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\mu = \boldsymbol{\alpha} \Big\}.
\end{aligned}$$

On the other hand, if $\mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma)$ such that $\int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\mu = \boldsymbol{\alpha}$, then for any $0 \leq t < \dim_L \mu$, it holds by the sub-additive variational principle (see [CFH08]) that, for all $\mathbf{q} \in \mathbb{R}^d$,

$$P\big(\log\phi^t + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle\big) \geq h_\mu(\sigma) + \lim_{n\to+\infty} \frac{1}{n}\int_{\mathcal{U}^{\mathbb{N}}} \log\phi^t\big(\mathbf{A}_{\xi|_n}\big)\,\mathrm{d}\mu(\xi) > 0.$$

Hence,

$$t \leq \sup\Big\{ s \geq 0 : \inf_{\mathbf{q}\in\mathbb{R}^d} P\big(\log\phi^s + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle\big) \geq 0 \Big\}$$

and thus

$$\sup\Big\{ \dim_L \mu : \mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma), \int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\mu = \boldsymbol{\alpha} \Big\} \leq \sup\Big\{ s \geq 0 : \inf_{\mathbf{q}\in\mathbb{R}^d} P\big(\log\phi^s + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle\big) \geq 0 \Big\}$$

holds, giving the result. □

Remark 2.4. Note that the proof of Lemma 2.3 in fact shows that

$$\begin{aligned}
&\sup\Big\{ s \geq 0 : \inf_{\mathbf{q}\in\mathbb{R}^d} P(\log\phi^s + \langle \mathbf{q}, \Phi - \boldsymbol{\alpha}\rangle) \geq 0 \Big\} \\
&\quad = \sup\left\{ \dim_L \tilde{\nu} : \nu \text{ fully supported } n\text{-step Bernoulli and } \int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\tilde{\nu} = \boldsymbol{\alpha} \right\} \\
&\quad = \sup\Big\{ \dim_L \mu : \mu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma), \int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\mu = \boldsymbol{\alpha} \Big\}.
\end{aligned}$$

Under the additional conditions mentioned in the statement of Theorem 2.1, we can prove that this upper bound in fact equals the Hausdorff dimension of the level set. Note that it would be possible to combine Hochman [Hoc14] and Jordan and Simon [JS07] to obtain a similar result for almost all vectors $\mathbf{v}_u$, but our Theorem 2.1 is proved for all vectors $\mathbf{v}_u$. [BJKR21] proved a similar result for affine IFSs satisfying the SOSC under the assumption that the set of matrices $(\mathbf{A}_u)_{u\in\mathcal{U}}$ is strongly irreducible such that the generated subgroup of the normalised matrices is not relatively compact; Theorem 2.1 is inspired by their result.

Proof of Theorem 2.1. Let $\{\mathbf{A}_u + \mathbf{v}_u\}_{u\in\mathcal{U}} \in \mathcal{G}$. For each $u \in \mathcal{U}$, it holds that

$$\frac{b_u \cdot c_u}{\max\{b_u, c_u\}^2} = \frac{\min\{b_u, c_u\}}{\max\{b_u, c_u\}} < 1,$$

since either $|b_u| > |c_u|$ for all $u \in \mathcal{U}$ or $|b_u| < |c_u|$ for all $u \in \mathcal{U}$. Take

$$\tau := \max_{u\in\mathcal{U}} \left\{ \frac{\min\{b_u, c_u\}}{\max\{b_u, c_u\}} \right\}.$$

Then $\tau < 1$, and, since each $\mathbf{A}_u$ is a diagonal matrix, we get (2.13) with $C = 1$. Hence, $\{\mathbf{A}_u\}_{u\in\mathcal{U}}$ is dominated and the desired upper bound for the Hausdorff dimension of $F_\Phi(\boldsymbol{\alpha})$ is therefore given by Lemma 2.3.

For the lower bound, suppose that $\nu \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma^n)$ is a fully supported n-step Bernoulli measure with $\int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\tilde{\nu} = \boldsymbol{\alpha}$ with $\tilde{\nu}$ as defined in (2.15). The existence of the measure ν is guaranteed by [BJKR21, Proposition 4.3]. Then $\tilde{\nu} \in \mathcal{M}(\mathcal{U}^{\mathbb{N}}, \sigma)$ and is ergodic. Therefore, from $\int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\tilde{\nu} = \boldsymbol{\alpha}$, we get that

$$\tilde{\nu}\left(\left\{\xi \in \mathcal{U}^{\mathbb{N}} : \lim_{n\to+\infty} \frac{1}{n} S_n \Phi(\xi) = \boldsymbol{\alpha}\right\}\right) = 1. \tag{2.16}$$

Let $\hat{\nu} = \tilde{\nu} \circ \pi^{-1}$. Assume that $|b_u| > |c_u|$ for all $u \in \mathcal{U}$ so that we are in the situation of condition (D) (the proof for the case (D') goes similarly). Then the strong stable direction of the collection $(\mathbf{A}_u)_{u\in\mathcal{U}}$ is equal to the subspace parallel to the y-axis (see [BG09]). Let $P_x\hat{\nu}$ be the measure on $[0, 1]$ given by the canonical projection onto the x-coordinate of $\hat{\nu}$. Since the matrices $\mathbf{A}_u$ are diagonal, $P_x\hat{\nu}$ is a self-similar measure for the IFS $\mathcal{G}_1$, i.e. there is a probability vector $\hat{\boldsymbol{p}} = (\hat{p}_u)_{u\in\mathcal{U}}$ such that

$$P_x\hat{\nu}(B) = \sum_{u\in\mathcal{U}} \hat{p}_u P_x\hat{\nu}(g_{1,u}^{-1}(B))$$

for each Borel set $B \subseteq [0, 1]$. Then condition (D)(a) together with [Hoc14, Theorem 1.1] or condition (D)(b) together with [Rap22, Theorem 1.2] yields

$$\dim_H P_x\hat{\nu} = \min\left\{1, \frac{h_{\tilde{\nu}}(\sigma)}{\chi_1(\tilde{\nu})}\right\}. \tag{2.17}$$

It then follows from [BK17, Corollaries 2.7 and 2.8] and (2.17) that $\dim_H \hat{\nu} = \dim_L \tilde{\nu}$. This and (2.16) yield $\dim_H F_\Phi(\boldsymbol{\alpha}) \geq \dim_L \tilde{\nu}$. Since this holds for arbitrary fully supported n-step Bernoulli measures ν with $\int_{\mathcal{U}^{\mathbb{N}}} \Phi \,\mathrm{d}\tilde{\nu} = \boldsymbol{\alpha}$, the result follows from Remark 2.4. □

Remark 2.5. We make a small remark on the conditions (D) and (D'). It was shown by Hochman in [Hoc14, proof of Theorem 1.5] that an IFS satisfies the ESC if it does not have exact overlaps and all parameters $b_u, c_u, \beta_u, \gamma_u$ are algebraic numbers over $\mathbb{Q}$. In [Che21; Bak21; BK21] it was shown that there exist iterated function systems that do not contain exact overlaps while there are cylinders which are super-exponentially close at all small scales, i.e. the ESC does not hold. What is needed in the proof of Theorem 2.1 is (2.17), which is also guaranteed by [Rap22] under the assumption of having algebraic b_u, c_u and no exact overlaps.

2.4 Digit Frequencies for Finite GLS Expansions

We now move to the second type of IFS we consider. Fix a GLS IFS $\{\mathbf{A}_d + \mathbf{v}_d\}_{d\in\mathcal{D}}$. We start by proving some properties of the expansions from (2.9).

2.4.1 Multiple Representations

First, consider the representations of the points $y \in [0,1]$. Recall the definition of the maps f_s, $s \in \mathcal{S}$, from (2.12). The IFS $\{f_s\}_{s\in\mathcal{S}}$ satisfies the SOSC and has the interval $[0,1]$ as its attractor. Put $\Omega := \mathcal{S}^{\mathbb{N}}$, and let $\pi_{\mathcal{S}} : \Omega \to [0,1]$ be the map given by

$$\pi_{\mathcal{S}}(\omega) = \lim_{n\to+\infty} f_{\omega_1} \circ f_{\omega_2} \circ \cdots \circ f_{\omega_n}(0).$$

One easily sees that to all but countably many $y \in [0,1]$, there corresponds a unique sequence $\omega \in \Omega$ such that $y = \pi_{\mathcal{S}}(\omega)$ and otherwise $\#\pi_{\mathcal{S}}^{-1}\{y\} = 2$ and there is one sequence ending in an infinite string of 0's and one ending in an infinite string of S's. We make the following observation, which we will use later on. Recall the definition of the set $Y(\boldsymbol{\alpha})$ from (2.11).

2

Lemma 2.6. *Let $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}} \in [0,1]^{\#\mathcal{D}}$ be a frequency vector with $\alpha_s > 0$ for each $s \in \mathcal{S}$. Then $\#\pi_{\mathcal{S}}^{-1}\{y\} = 1$ for all $y \in Y(\boldsymbol{\alpha})$.*

Proof. Let $y \in [0,1]$ be such that $\#\pi_{\mathcal{S}}^{-1}\{y\} = 2$, and let $x \in [0,1]$. Recalling that $\mathcal{S} = \{1,\dots,S\}$, it follows that any $\xi = (\omega_n, b_n)_{n\in\mathbb{N}} \in \pi^{-1}\{(x,y)\}$ either has $\omega_n = 1$ for all n large enough or $\omega_n = S$ for all n large enough. In the former case, $\sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi) = 0 \neq \alpha_s$ for all $s \neq 1$. In the latter case, $\sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi) = 0 \neq \alpha_s$ for all $s \neq S$. Hence, $(x,y) \notin F(\boldsymbol{\alpha})$ and thus $Y(\boldsymbol{\alpha}) = \emptyset$. □

For a fixed $y \in [0,1]$, we can consider the expansions one obtains from the GLS IFS for $x \in [0,1]$. We define the *fibre fundamental intervals* corresponding to y by setting for each $n \in \mathbb{N}$ and $b_1,\dots,b_n$ satisfying $b_\ell \in \mathcal{B}_{\omega_\ell}$ for all $1 \le \ell \le n$,

$$\langle b_1 \cdots b_n \rangle_y := h_{(\omega_1,b_1)} \circ \cdots \circ h_{(\omega_n,b_n)}([0,1]), \tag{2.18}$$

where $(\omega_n)_{n\in\mathbb{N}}$ is the lexicographically smallest sequence in $\pi_{\mathcal{S}}^{-1}\{y\}$. For $y \in Y(\boldsymbol{\alpha})$, this means that $\omega = (\omega_n)_{n\in\mathbb{N}} \in \pi_{\mathcal{S}}^{-1}\{y\}$ is the unique sequence that satisfies $\tau_s(\omega) = \alpha_s$ for each $s \in \mathcal{S}$; for $y \in [0,1] \setminus Y(\boldsymbol{\alpha})$, the sequence $(\omega_n)_{n\in\mathbb{N}}$ is the one ending in an infinite string of S's.

If we fix $y \in Y(\boldsymbol{\alpha})$ and take $x \in [0,1]$ such that $\#\pi^{-1}\{(x,y)\} > 1$, then we know by Lemma 2.6 that $\#\pi_{\mathcal{S}}^{-1}\{y\} = 1$, say $y = \pi_{\mathcal{S}}(\omega)$. Consequently, x must have multiple expansions along the fibre y and so must lie on the boundary of a fibre fundamental interval $\langle b_1 \cdots b_n \rangle_y$ for some $b_\ell \in \mathcal{B}_{\omega_\ell}$, $1 \le \ell \le n$, and some $n \in \mathbb{N}$. Since each fibre fundamental interval has two boundary points and there are only countably many fibre fundamental intervals, the set of such points x must be countable.

Now, fix an $x \in [0,1]$. Since the GLS IFS $\{\mathbf{A}_d + \mathbf{v}_d\}_{d\in\mathcal{D}}$ has $[0,1]^2$ as its attractor, to any $y \in [0,1]$, there corresponds a sequence $\xi \in \mathcal{D}^{\mathbb{N}}$ such that $\pi(\xi) = (x,y)$. Therefore, to any sequence $\omega \in \Omega$, there corresponds a sequence $(b_n)_{n\in\mathbb{N}}$ with $b_n \in \mathcal{B}_{\omega_n}$ for each $n \in \mathbb{N}$ such that

$$x = \lim_{n\to+\infty} h_{(\omega_1,b_1)} \circ \cdots \circ h_{(\omega_n,b_n)}(0).$$

We show that if $h_d \neq h_{d'}$ whenever $d \neq d'$, then each of the sequences $\omega \in \Omega$ yields a different GLS expansion for x as in (2.9).

The GLS expansions from (2.9) are given by the triples of digits (ϵ_n, N_n, t_n), $n \in \mathbb{N}$, from (2.8). Therefore, if we set

$$\mathcal{A} = \left\{ \left(\varepsilon_d, q_d^{-1}, r_d + \varepsilon_d q_d\right) : d \in \mathcal{D} \right\},$$

then we can think of $\mathcal{A}$ as the *GLS digit set* corresponding to $\{\mathbf{A}_d + \mathbf{v}_d\}_{d\in\mathcal{D}}$ and we can map sequences $((\omega_n, b_n))_{n\in\mathbb{N}} \in \mathcal{D}^{\mathbb{N}}$ to sequences $(\epsilon_n, N_n, t_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ through the identification given in (2.8). Let $\omega, \omega' \in \Omega$ be two different sequences so that there is an $m \in \mathbb{N}$ such that $\omega_m \neq \omega'_m$. Let $(\omega_n, b_n)_{n\in\mathbb{N}}, (\omega'_n, b'_n)_{n\in\mathbb{N}} \in \mathcal{D}^{\mathbb{N}}$ be two sequences that both project to x in the second coordinate under π. Since $\omega_m \neq \omega'_m$, it holds that $(\omega_m, b_m) \neq (\omega'_m, b'_m)$. If we assume that $h_d \neq h_{d'}$ whenever $d \neq d'$, then it would follow that $\varepsilon_{(\omega_m,b_m)} \neq \varepsilon_{(\omega'_m,b'_m)}$ or $q_{(\omega_m,b_m)} \neq q_{(\omega'_m,b'_m)}$ and thus that the digits from $\mathcal{A}$ corresponding to (ω_m, b_m) and (ω'_m, b'_m) differ. Therefore, we immediately find the following result.

Proposition 2.7. *Let $\{\mathbf{A}_d+\mathbf{v}_d\}_{d\in\mathcal{D}}$ be a GLS IFS with the additional assumption that $h_d \neq h_{d'}$ whenever $d \neq d'$. Then for each $x \in [0,1]$, there are uncountably many different digit sequences $(\epsilon_n, N_n, t_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ with*

$$x = \sum_{n\in\mathbb{N}} (-1)^{\sum_{1\leq \ell<n} \epsilon_\ell} \frac{t_n}{\prod_{1\leq \ell \leq n} N_\ell}.$$

The above also shows that there is a one-to-one correspondence between the sequences in $\mathcal{D}^{\mathbb{N}}$ and in $\mathcal{A}^{\mathbb{N}}$, which justifies considering the elements of $\mathcal{D}$ as digits in GLS expansions.

2

2.4.2 Non-empty Level Sets

In this section, we determine for which frequency vectors $\boldsymbol{\alpha}$ the level set $F(\boldsymbol{\alpha}) := \pi(E_{\mathbb{1}}(\boldsymbol{\alpha}))$ from (2.10) and the set $Y(\boldsymbol{\alpha})$ from (2.11) are non-empty. We first consider the level sets $F(\boldsymbol{\alpha})$.

Proposition 2.8. *The set $F(\boldsymbol{\alpha})$ is non-empty for any frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}} \in [0,1]^{\#\mathcal{D}}$.*

Proof. Fix a frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}} \in [0,1]^{\#\mathcal{D}}$. It is sufficient to construct a sequence $\xi = (\xi_n)_{n\in\mathbb{N}} \in \mathcal{D}^{\mathbb{N}}$ such that $\tau_d(\xi) = \alpha_d$ for each $d \in \mathcal{D}$ since then $\pi(\xi) \in F(\boldsymbol{\alpha})$. Denote by $\lfloor\cdot\rceil$ the nearest-integer function. Order the elements in $\mathcal{D}$ by setting $(s,b) \prec (s',b')$ if either $s < s'$ or if $s = s'$ and $b < b'$. For each $n \in \mathbb{N}$, set

$$E_n = \{d \in \mathcal{D} : \lfloor n\alpha_d \rceil = \lfloor (n-1)\alpha_d \rceil + 1\} = \{e_{n,1} \prec \cdots \prec e_{n,m_n}\},$$

where E_n can be empty and thus $m_n = 0$ for some n. Define $\xi \in \mathcal{D}^{\mathbb{N}}$ by setting for each $n \in \mathbb{N}$ and $1 \leq m \leq m_n$

$$\xi_{m+\sum_{i=1}^{n-1} m_i} = e_{n,m},$$

where we let $\sum_{i=1}^{0} m_i = 0$. Clearly, there are infinitely many n for which $E_n \neq \emptyset$ so ξ is well defined. Now observe that, for each n, the number of terms of ξ we have defined using $\bigcup_{i=1}^{n} E_i$ is

$$\sum_{i=1}^{n} m_i = \sum_{d\in\mathcal{D}} \lfloor n\alpha_d \rceil.$$

Since $\sum_{d\in\mathcal{D}} n\alpha_d = n$ and $\mathfrak{m} := \#\mathcal{D} < +\infty$, we must have

$$n - \mathfrak{m} \leq \sum_{i=1}^{n} m_i \leq n + \mathfrak{m}.$$

Thus for each $d \in \mathcal{D}$,

$$\begin{aligned}\frac{\#\{0 \leq m < n : \xi_m = d\}}{n} &\leq \frac{\#\{0 \leq m < \mathfrak{m} + \sum_{i=1}^{n} m_i : \xi_m = d\}}{n} \\ &\leq \frac{\#\{0 \leq m < \sum_{i=1}^{n} m_i : \xi_m = d\} + \mathfrak{m}}{n} \leq \frac{n\alpha_d + 1 + \mathfrak{m}}{n}.\end{aligned}$$

Similarly, it holds that

$$\frac{\#\{0 \leq m < n : \xi_m = d\}}{n} \geq \frac{n\alpha_d - 1 - \mathfrak{m}}{n}.$$

Taking the limit as $n \to +\infty$ yields $\tau_d(\xi) = \alpha_d$ for all $d \in \mathcal{D}$. □

For a fixed frequency vector $\boldsymbol{\alpha}$, we would like to determine the set $Y(\boldsymbol{\alpha})$ of points $y \in [0,1]$ for which there exists an $x \in [0,1]$ such that the point (x,y) has digit frequencies given by $\boldsymbol{\alpha}$; see (2.11). Recall the definition of the Borel measure $\nu_{\boldsymbol{\alpha}} = \mu_{\boldsymbol{\alpha}} \circ \pi_2^{-1}$ from §2.1. We have the following result.

2

Proposition 2.9. *Let* $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}} \in [0,1]^{\#\mathcal{D}}$. *Then*

$$Y(\boldsymbol{\alpha}) = \left\{ y \in [0,1] : \exists \xi \in \pi_2^{-1}\{y\} \textit{ s.t. } \sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi) = \alpha_s \textit{ for all } s \in \mathcal{S} \right\}.$$

In particular, $\nu_{\boldsymbol{\alpha}}(Y(\boldsymbol{\alpha})) = 1$.

Proof. For the forward direction, set

$$Y = \left\{ y \in [0,1] : \exists \xi \in \pi_2^{-1}\{y\} \text{ s.t. } \sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi) = \alpha_s \text{ for all } s \in \mathcal{S} \right\}$$

and take $y \in Y(\boldsymbol{\alpha})$. This means that $F_y(\boldsymbol{\alpha}) \neq \emptyset$ so there is a $\xi \in \mathcal{D}^{\mathbb{N}}$ such that $\pi(\xi) = (x,y)$ for some $x \in [0,1]$ and $\tau_d(\xi) = \alpha_d$ for all $d \in \mathcal{D}$. Consequently, we find that

$$\sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi) = \sum_{b\in\mathcal{B}_s} \alpha_{(s,b)} = \alpha_s$$

for all $s \in \mathcal{S}$ and so $y \in Y$.

For the reverse direction, take $y \in Y$. Then there is a $\xi' = (\omega_n, b'_n)_{n\in\mathbb{N}} \in \pi_2^{-1}\{y\}$ for which it holds that $\sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi') = \alpha_s$ for all $s \in \mathcal{S}$. For each $s \in \mathcal{S}$ and $n \in \mathbb{N}$, set

$$E_n^{(s)} := \left\{ (s,b) \in \mathcal{D} : \left\lfloor \frac{n\alpha_{(s,b)}}{\alpha_s} \right\rfloor = \left\lfloor \frac{(n-1)\alpha_{(s,b)}}{\alpha_s} \right\rfloor + 1 \right\} = \left\{ e_{n,1}^{(s)} \prec \cdots \prec e_{n,m_n^{(s)}}^{(s)} \right\}$$

and let $\xi^{(s)} \in \mathcal{D}^{\mathbb{N}}$ be the sequence obtained from concatenating all elements from the sets $E_n^{(s)}$ as in Proposition 2.8 so that

$$\xi^{(s)} = e_{1,1}^{(s)} \cdots e_{1,m_1^{(s)}}^{(s)} e_{2,1}^{(s)} \cdots e_{2,m_2^{(s)}}^{(s)} e_{3,1}^{(s)} \cdots.$$

Then as in Proposition 2.8, we obtain for each $b \in \mathcal{B}_s$ that

$$\lim_{n\to+\infty} \frac{\#\{1 \leq \ell \leq n : \xi_\ell^{(s)} = (s,b)\}}{n} = \frac{\alpha_{(s,b)}}{\alpha_s}.$$

We now weave the sequences $\xi^{(s)}$ together to construct a sequence $\xi = (\omega_n, b_n)_{n\in\mathbb{N}} \in \mathcal{D}^{\mathbb{N}}$ that satisfies $\tau_d(\xi) = \alpha_d$ for all $d \in \mathcal{D}$. Then $\pi(\xi) = (x,y)$ for some $x \in F_y(\boldsymbol{\alpha})$, which shows that $F_y(\boldsymbol{\alpha}) \neq \emptyset$. For each $\ell \in \mathbb{N}$, let ξ_ℓ be the $\tau_{\omega_\ell}(\omega,\ell)^{\text{th}}$ element of the sequence $\xi^{(\omega_\ell)}$. So, $\xi_1 = \xi_1^{(\omega_1)} = e_{1,1}^{(\omega_1)}$, ξ_2 either equals $\xi_2^{(\omega_1)}$ if $\omega_1 = \omega_2$ or $\xi_1^{(\omega_2)}$ if $\omega_1 \neq \omega_2$, et cetera. As the sequences in the first coordinates of $\xi = (\omega_\ell, b_\ell)_{\ell\in\mathbb{N}}$ and $\xi' = (\omega_\ell, b'_\ell)_{\ell\in\mathbb{N}}$ coincide, we have for each $s \in \mathcal{S}$ that

$$\sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi) = \sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi') = \alpha_s.$$

Moreover, for each $d = (s, b) \in \mathcal{D}$ and $n \in \mathbb{N}$,

$$\#\{1 \le \ell \le n : \xi_\ell = d\} = \#\{1 \le \ell \le \tau_s(\omega, n) : \xi_\ell^{(s)} = d\}. \tag{2.19}$$

If $\alpha_s > 0$, then $\tau_s(\omega, n) > 0$ for all n large enough and so, for any $d = (s, b) \in \mathcal{D}$, we obtain

$$\tau_d(\xi) = \lim_{n\to+\infty} \frac{\#\{1 \le \ell \le \tau_s(\omega, n) : \xi_m^{(s)} = d\}}{\tau_s(\omega, n)} \cdot \frac{\tau_s(\omega, n)}{n} = \frac{\alpha_{(s,b)}}{\alpha_s} \sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\xi) = \alpha_{(s,b)}.$$

If $\alpha_s = 0$, then $\alpha_{(s,b)} = 0$ for each $b \in \mathcal{B}_s$ and by (2.19),

$$0 \le \tau_{(s,b)}(\xi) \le \lim_{n\to+\infty} \frac{\tau_s(\omega, n)}{n} = \alpha_s = 0.$$

This gives the first part of the statement.

As $\left\{\xi \in \mathcal{D}^{\mathbb{N}} : \tau_d(\xi) = \alpha_d \text{ for all } d \in \mathcal{D}\right\} \subseteq \pi_2^{-1}(Y(\boldsymbol{\alpha}))$, it follows from the definition of $\mu_{\boldsymbol{\alpha}}$ that $\nu_{\boldsymbol{\alpha}}(Y(\boldsymbol{\alpha})) = 1$. □

2

2.4.3 The Hausdorff Dimension of the Besicovitch-Eggleston Sets

In this section, we prove Theorem 2.2. The proof is similar to [BI09, Theorem 3.1] and [FLMW10, Theorem 1.1], which both treat digit frequencies for expansions with infinite digit sets that can be generated by an IFS on $\mathbb{R}$ as in (2.5). Their results do not apply to our setting because the IFS $\{h_d : [0, 1] \to [0, 1]\}_{d\in\mathcal{D}}$ on $\mathbb{R}$ is not of this type. Nevertheless, since we have a finite digit set, we can adapt the method of proof from [BI09, Theorem 3.1].

Fix $y \in Y(\boldsymbol{\alpha})$, and let $\omega \in \Omega$ be the unique sequence such that $\pi_{\mathcal{S}}(\omega) = y$ by Lemma 2.6. Recall the definition of the fibre fundamental intervals $\langle b_1 \cdots b_n\rangle_y$ from (2.18). Note that we obtain a semi-algebra of sets generating the Borel sigma-algebra $\mathcal{B}([0, 1])$ on $[0, 1]$ by taking the collection of all intervals (open, closed and half-open) that can be formed by the endpoints of the fibre fundamental intervals. Define the map $m_{y,\boldsymbol{\alpha}}$ on the collection of all fibre fundamental intervals by

$$m_{y,\boldsymbol{\alpha}}(\langle b_1 \cdots b_n\rangle_y) = \prod_{1\le\ell\le n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}}, \quad b_\ell \in \mathcal{B}_{\omega_\ell},\ 1 \le \ell \le n,\ n \in \mathbb{N},$$

and by the same quantity for any interval determined by the same endpoints. Put $m_{y,\boldsymbol{\alpha}}(\emptyset) := 0$.

Lemma 2.10. *$m_{y,\boldsymbol{\alpha}}$ extends to a probability measure on the measurable space* $([0, 1], \mathcal{B}([0, 1]))$.

Proof. It is easy to see that $m_{y,\boldsymbol{\alpha}}$ defines a pre-measure on the semi-algebra of fibre fundamental intervals, and it is then a straight-forward application of Carathéodory's extension theorem to extend $m_{y,\boldsymbol{\alpha}}$; see the proof of Lemma 3.14 in the next chapter for details of a similar result. □

We prove some properties of $m_{y,\boldsymbol{\alpha}}$ in certain situations in the next two lemmas.

Lemma 2.11. *Suppose that the frequency vector $\boldsymbol{\alpha}$ satisfies the additional property that there is a $\varsigma \in \mathcal{S}$ for which there exists a $b \in \mathcal{B}_\varsigma$ such that $0 < \alpha_{(\varsigma,b)} < \alpha_\varsigma$. Then $m_{y,\boldsymbol{\alpha}}(\{x\}) = 0$ for all $x \in [0, 1]$.*

Proof. By definition of $m_{y,\boldsymbol{\alpha}}$, any endpoint $x \in [0,1]$ of a fibre fundamental interval has $m_{y,\boldsymbol{\alpha}}(\{x\}) = 0$. If $x \in [0,1]$ is not an endpoint of a fibre fundamental interval, then there is a sequence $(b_n)_{n\in\mathbb{N}}$ such that

$$\bigcap_{n\in\mathbb{N}} \langle b_1 \cdots b_n \rangle_y = \{x\},$$

which implies that

$$m_{y,\boldsymbol{\alpha}}(\{x\}) = \lim_{n\to+\infty} \prod_{1\le\ell\le n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}} \le \lim_{n\to+\infty} \left(\frac{\max\{\alpha_{(\varsigma,b)} : b \in \mathcal{B}_\varsigma\}}{\alpha_\varsigma} \right)^{\tau_\varsigma(\omega,n)}.$$

Since $\alpha_s > 0$ for all $s \in \mathcal{S}$ by assumption on $\boldsymbol{\alpha}$, we have that $\tau_\varsigma(\omega) = \alpha_\varsigma > 0$ by Proposition 2.9, using that $y \in Y(\boldsymbol{\alpha})$, so we must have $\tau_\varsigma(\omega,n) \to +\infty$ as $n \to +\infty$. By the additional assumption on $\boldsymbol{\alpha}$, it follows that $\max\{\alpha_{(\varsigma,b)} : b \in \mathcal{B}_\varsigma\} < \alpha_\varsigma$ and so we have $m_{y,\boldsymbol{\alpha}}(\{x\}) = 0$ by the inequality above. □

2

Lemma 2.12. *For $\nu_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$, it holds that $m_{y,\boldsymbol{\alpha}}(F_y(\boldsymbol{\alpha})) = 1$.*

Proof. By Proposition 2.9, we have for each $(s,b) \in \mathcal{D}$ that

$$\begin{aligned}
&\int_{[0,1]} \int_{[0,1]} \mathbb{1}_{\pi([(s,b)])}(x,y) \,\mathrm{d}m_{y,\boldsymbol{\alpha}}(x) \,\mathrm{d}\nu_{\boldsymbol{\alpha}}(y) \\
&\quad = \int_{[\sum_{i=1}^{s-1} p_i, \sum_{i=0}^{s} p_i]} \int_{[0,1]} \mathbb{1}_{\langle b\rangle_y)}(x) \,\mathrm{d}m_{y,\boldsymbol{\alpha}}(x) \,\mathrm{d}\nu_{\boldsymbol{\alpha}}(y) \\
&\quad = \frac{\alpha_{(s,b)}}{\alpha_s} \nu_{\boldsymbol{\alpha}}\left(\left[\sum_{i=1}^{s-1} p_i, \sum_{i=1}^{s} p_i \right] \right) = \alpha_{(s,b)} = \int_{[0,1]^2} \mathbb{1}_{\pi([(s,b)])} \,\mathrm{d}\mu_{\boldsymbol{\alpha}} \circ \pi^{-1}.
\end{aligned} \tag{2.20}$$

Since the collection $\{\pi([d_1 \cdots d_n]) : d_\ell \in \mathcal{D},\ 1 \le \ell \le n\}$ generates the Borel sigma-algebra $\mathcal{B}([0,1]^2)$, we can conclude from (2.20) that

$$\int_{[0,1]} \int_{[0,1]} f \,\mathrm{d}m_{y,\boldsymbol{\alpha}} \,\mathrm{d}\nu_{\boldsymbol{\alpha}} = \int_{[0,1]^2} f \,\mathrm{d}\mu_{\boldsymbol{\alpha}} \circ \pi^{-1} \tag{2.21}$$

for all $f \in \mathcal{L}^1([0,1]^2, \mathcal{B}([0,1]^2), \mu_{\boldsymbol{\alpha}} \circ \pi^{-1})$.

Set $E \coloneqq \{y \in Y(\boldsymbol{\alpha}) : m_{y,\boldsymbol{\alpha}}(F_y(\boldsymbol{\alpha})) < 1\}$, and suppose so as to obtain a contradiction that $\nu_{\boldsymbol{\alpha}}(E) > 0$. By (2.21) with $f = \mathbb{1}_{F(\boldsymbol{\alpha})}$ together with Proposition 2.9, we find that

$$\begin{aligned}
\mu_{\boldsymbol{\alpha}} \circ \pi^{-1}(F(\boldsymbol{\alpha})) &= \int_{Y(\boldsymbol{\alpha})} \int_{[0,1]} \mathbb{1}_{F_y(\boldsymbol{\alpha})}(x) \,\mathrm{d}m_{y,\boldsymbol{\alpha}}(x) \,\mathrm{d}\nu_{\boldsymbol{\alpha}}(y) \\
&= \int_{Y(\boldsymbol{\alpha})\setminus E} 1 \,\mathrm{d}\nu_{\boldsymbol{\alpha}} + \int_E \int_{[0,1]} \mathbb{1}_{F_y(\boldsymbol{\alpha})}(x) \,\mathrm{d}m_{y,\boldsymbol{\alpha}}(x) \,\mathrm{d}\nu_{\boldsymbol{\alpha}}(y) \\
&< \int_{Y(\boldsymbol{\alpha})} 1 \,\mathrm{d}\nu_{\boldsymbol{\alpha}} = 1.
\end{aligned}$$

On the other hand, recall that $E_{\mathbb{1}}(\boldsymbol{\alpha})$ is the symbolic Besicovitch-Eggleston set containing all sequences $\xi \in \mathcal{D}^{\mathbb{N}}$ with $\tau_d(\xi) = \alpha_d$ for each $d \in \mathcal{D}$. Therefore, by the definition of $\mu_{\boldsymbol{\alpha}}$,

$$1 = \mu_{\boldsymbol{\alpha}}(E_{\mathbb{1}}(\boldsymbol{\alpha})) \le \mu_{\boldsymbol{\alpha}} \circ \pi^{-1}(F(\boldsymbol{\alpha})) < 1,$$

which is a contradiction. It follows that $m_{y,\boldsymbol{\alpha}}(F_y(\boldsymbol{\alpha})) = 1$ for $\nu_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$. □

Before we move to the proof of Theorem 2.2, to simplify notation, we put $p_{(s,b)} = p_s$ for all $(s,b) \in \mathcal{D}$. Also, let $\mathbb{P}_{\boldsymbol{\alpha}}$ be the $(\alpha_s)_{s\in\mathcal{S}}$-Bernoulli measure on Ω.

Proof of Theorem 2.2. Fix a $y \in Y(\boldsymbol{\alpha})$. Recall that the lower pointwise dimension of $m_{y,\boldsymbol{\alpha}}$ at the point $x \in [0,1]$ is defined by

$$\underline{d}_{m_{y,\boldsymbol{\alpha}}}(x) = \liminf_{r\to 0} \frac{\log m_{y,\boldsymbol{\alpha}}(B(x,r))}{\log r},$$

where $B(x,r)$ is the open interval of length $2r$ centred at x. One can verify that the collection $\{\langle b_1 \cdots b_n\rangle_y : n \in \mathbb{N}\}$ satisfies conditions (CB1)–(CB3) of the Moran-type construction from [Pes97, §15]. Moreover, for any intervals $\langle b_1 \cdots b_n\rangle_y$, $\langle b_1 \cdots b_n b_{n+1}\rangle_y$, it holds that

$$|\langle b_1 \cdots b_n\rangle_y| \le (\max_{d\in\mathcal{D}} q_d)^n, \qquad \min_{d\in\mathcal{D}} q_d \cdot |\langle b_1 \cdots b_n\rangle_y| \le |\langle b_1 \cdots b_n b_{n+1}\rangle_y|.$$

2

Therefore, by e.g. [Pes97, Theorem 15.3(1)], we can replace the balls $B(x,r)$ in the definition of $\underline{d}_{m_{y,\boldsymbol{\alpha}}}$ with the fibre fundamental intervals $\langle b_1 \cdots b_n\rangle_y$ to obtain an upper bound for $\underline{d}_{m_{y,\boldsymbol{\alpha}}}(x)$ for all $x \in F_y(\boldsymbol{\alpha})$ and a lower bound for $m_{y,\boldsymbol{\alpha}}$-a.e. $x \in [0,1]$ in the case that $m_{y,\boldsymbol{\alpha}}(F_y(\boldsymbol{\alpha})) = 1$. To be more precise, for $x \in F_y(\boldsymbol{\alpha})$ with $\xi = ((\omega_\ell, b_\ell))_{\ell\in\mathbb{N}} \in \pi^{-1}\{(x,y)\}$ that has $\tau_d(\xi) = \alpha_d$ for each $d \in \mathcal{D}$, we find that

$$\begin{aligned}\underline{d}_{m_{y,\boldsymbol{\alpha}}}(x) &\le \lim_{n\to+\infty} \frac{\log m_{y,\boldsymbol{\alpha}}(\langle b_1 \cdots b_n\rangle_y)}{\log|\langle b_1 \cdots b_n\rangle_y|} = \lim_{n\to+\infty} \frac{\log \prod_{1\le\ell\le n} \alpha_{(\omega_\ell,b_\ell)}/\alpha_{\omega_\ell}}{\log \prod_{1\le\ell\le n} q_{(\omega_\ell,b_\ell)}} \\ &= \lim_{n\to+\infty} \frac{\frac{1}{n}\sum_{1\le\ell\le n} \log \alpha_{(\omega_\ell,b_\ell)} - \frac{1}{n}\sum_{1\le\ell\le n} \log \alpha_{\omega_\ell}}{\frac{1}{n}\sum_{1\le\ell\le n} \log q_{(\omega_\ell,b_\ell)}}.\end{aligned}$$

By collecting like terms, we find that

$$\begin{aligned}\sum_{1\le\ell\le n} \log \alpha_{(\omega_\ell,b_\ell)} &= \sum_{d\in\mathcal{D}} \#\{1 \le \ell \le n : \xi_\ell = d\} \log \alpha_d, \\ \sum_{1\le\ell\le n} \log \alpha_{\omega_\ell} &= \sum_{s\in\mathcal{S}} \#\{1 \le \ell \le n : \omega_\ell = s\} \log \alpha_s, \\ \sum_{1\le\ell\le n} \log q_{(\omega_\ell,b_\ell)} &= \sum_{d\in\mathcal{D}} \#\{1 \le \ell \le n : \xi_\ell = d\} \log q_d.\end{aligned}$$

Since $x \in F_y(\boldsymbol{\alpha})$, we have for each $d \in \mathcal{D}$ and $s \in \mathcal{S}$ that

$$\begin{aligned}&\lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : \xi_\ell = d\}}{n} = \tau_d(\xi) = \alpha_d, \\ &\lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : \omega_\ell = s\}}{n} = \lim_{n\to+\infty} \sum_{b\in\mathcal{B}_s} \frac{\#\{1 \le \ell \le n : \xi_{(\omega_\ell,b)} = (s,b)\}}{n} = \alpha_s.\end{aligned}$$

Thus, recalling the definition of measure-theoretic entropy from (1.6), we find that

$$\underline{d}_{m_{y,\boldsymbol{\alpha}}}(x) \le \frac{\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d - \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{\sum_{d\in\mathcal{D}} \alpha_d \log q_d} = \frac{h_{\mu_{\boldsymbol{\alpha}}}(\sigma) - h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{d\in\mathcal{D}} \alpha_d \log q_d}$$

for all $x \in F_y(\boldsymbol{\alpha})$. Therefore, it follows from Lemma 1.5(i) that

$$\dim_H F_y(\boldsymbol{\alpha}) \le \frac{h_{\mu_{\boldsymbol{\alpha}}}(\sigma) - h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{d\in\mathcal{D}} \alpha_d \log q_d}.$$

To prove the second statement, we split into two cases. First, suppose that there is no $\varsigma \in \mathcal{S}$ such that $0 < \alpha_{(\varsigma,b)} < \alpha_\varsigma$ for some $b \in \mathcal{B}_\varsigma$. Then for each $s \in \mathcal{S}$, we must have that there is a unique $b_s \in \mathcal{B}_s$ such that $\alpha_{(s,b_s)} = \alpha_s$. Note then that $\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d = \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s$ and so

$$\frac{\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d - \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{\sum_{d\in\mathcal{D}} \alpha_d \log q_d} = 0,$$

which is trivially a lower bound for $\dim_H F_y(\boldsymbol{\alpha})$.

In the case that there is a $\varsigma \in \mathcal{S}$ such that $0 < \alpha_{(\varsigma,b)} < \alpha_\varsigma$ for some $b \in \mathcal{B}_\varsigma$, fix $y \in Y(\boldsymbol{\alpha})$ such that $m_{y,\boldsymbol{\alpha}}(F_y(\boldsymbol{\alpha})) = 1$, which holds for $\nu_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$ by Lemma 2.12. Let $\omega \in \pi_{\mathcal{S}}^{-1}\{y\}$ be the unique sequence with $\tau_s(\omega) = \alpha_s$ for each $s \in \mathcal{S}$ (recall Lemma 2.6). By the above computations for the upper bound of $\dim_H F_y(\boldsymbol{\alpha})$ together with [Pes97, Theorem 15.3(2)], we have for $m_{y,\boldsymbol{\alpha}}$-a.e. $x \in [0,1]$ that

$$\underline{d}_{m_{y,\boldsymbol{\alpha}}}(x) \geq \inf \lim_{n\to+\infty} \frac{\log m_{y,\boldsymbol{\alpha}}(\langle b_1 \cdots b_n\rangle_y)}{\log|\langle b_1 \cdots b_n\rangle_y|}, \tag{2.22}$$

where the infimum is taken over all sequences $(b_n)_{n\in\mathbb{N}}$ such that $(\omega_n, b_n)_{n\in\mathbb{N}} \in \pi^{-1}\{(x,y)\}$. We have seen that the set of x for which $\#\pi^{-1}\{(x,y)\} > 1$ is countable so is therefore a $m_{y,\boldsymbol{\alpha}}$-null set by Lemma 2.11, which may be applied in this case. Consequently, the infimum on the right-hand side of (2.22) is over a single sequence for $m_{y,\boldsymbol{\alpha}}$-a.e. $x \in [0,1]$. Fix $x \in F_y(\boldsymbol{\alpha})$ such that (2.22) and $\#\pi^{-1}\{(x,y)\} = 1$ both hold. Then

$$\begin{aligned}\underline{d}_{m_{y,\boldsymbol{\alpha}}}(x) &\geq \lim_{n\to+\infty} \frac{\log m_{y,\boldsymbol{\alpha}}(\langle b_1 \cdots b_n\rangle_y)}{\log \operatorname{diam}(\langle b_1 \cdots b_n\rangle_y)} \\ &= \lim_{n\to+\infty} \frac{\frac{1}{n}\sum_{d\in\mathcal{D}} \#\{1 \leq \ell \leq n : \xi_\ell = d\} \log \alpha_d - \frac{1}{n}\sum_{s\in\mathcal{S}} \#\{1 \leq \ell \leq n : \omega_\ell = s\} \log \alpha_s}{\frac{1}{n}\sum_{d\in\mathcal{D}} \#\{1 \leq \ell \leq n : \xi_\ell = d\} \log q_d}.\end{aligned}$$

Since $x \in F_y(\boldsymbol{\alpha})$, we find for each $d \in \mathcal{D}$ and $s \in \mathcal{S}$ that

$$\tau_d(\xi) = \lim_{n\to+\infty} \frac{\#\{1 \leq \ell \leq n : \xi_\ell = d\}}{n} = \alpha_d \quad \text{and}$$
$$\tau_s(\omega) = \lim_{n\to+\infty} \frac{\#\{1 \leq \ell \leq n : \omega_\ell = s\}}{n} = \alpha_s.$$

Therefore,

$$\underline{d}_{m_{y,\boldsymbol{\alpha}}}(x) \geq \frac{\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d - \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{\sum_{d\in\mathcal{D}} \alpha_d \log q_d} = \frac{h_{\mu_{\boldsymbol{\alpha}}}(\sigma) - h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{d\in\mathcal{D}} \alpha_d \log q_d}.$$

Since this holds for $m_{y,\boldsymbol{\alpha}}$-a.e. $x \in F_y(\boldsymbol{\alpha})$ and $m_{y,\boldsymbol{\alpha}}(F_y(\boldsymbol{\alpha})) = 1$, it follows from Lemma 2.12 and Lemma 1.5(ii) that

$$\dim_H F_y(\boldsymbol{\alpha}) \geq \dim_H m_{y,\boldsymbol{\alpha}} \geq \frac{h_{\mu_{\boldsymbol{\alpha}}}(\sigma) - h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{d\in\mathcal{D}} \alpha_d \log q_d}. \qquad \square$$

2.5 Additional Remarks & Examples

Remark 2.13. The bounds in Theorem 2.2 give a precise formula for the Hausdorff dimension for $\mathbb{P}_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$, and one may wonder if there exist $y \in Y(\boldsymbol{\alpha})$ for which this formula does not hold. We will see in Lemma 3.14 of Chapter 3 that, if we were to instead construct the measure $m_{y,\boldsymbol{\alpha}}$ by first starting on the sigma-algebra of sets that intersect $F_y(\boldsymbol{\alpha})$ and then extending to $\mathcal{B}([0,1])$ by setting $m_{y,\boldsymbol{\alpha}}(A) = m_{y,\boldsymbol{\alpha}}(A \cap F_y(\boldsymbol{\alpha}))$ for each $A \in \mathcal{B}([0,1])$ (as opposed to defining $m_{y,\boldsymbol{\alpha}}$ on $\mathcal{B}([0,1])$ immediately), then it would follow directly from the construction that $m_{y,\boldsymbol{\alpha}}(F_y(\boldsymbol{\alpha})) = 1$ for *all* $y \in Y(\boldsymbol{\alpha})$, which would allow us to extend the result in Theorem 2.2 to hold for *all* $y \in Y(\boldsymbol{\alpha})$.

Remark 2.14. As a corollary of Theorem 2.2, we can obtain a lower bound in terms of the Ledrappier-Young formula for the Bernoulli measure $\mu_{\boldsymbol{\alpha}}$ for the Hausdorff dimension of the Besicovitch-Eggleston sets $F(\boldsymbol{\alpha})$ from (2.10):

$$\dim_H F(\boldsymbol{\alpha}) \geq \frac{h_{\mu_{\boldsymbol{\alpha}}}(\sigma) - h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{d \in \mathcal{D}} \alpha_d \log q_d} + \frac{h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{s \in \mathcal{S}} \alpha_s \log p_s}.$$

To see this, set

$$t_0 := \frac{\sum_{d \in \mathcal{D}} \alpha_d \log \alpha_d - \sum_{s \in \mathcal{S}} \alpha_s \log \alpha_s}{\sum_{d \in \mathcal{D}} \alpha_d \log q_d}.$$

Put $A := \{y \in Y(\boldsymbol{\alpha}) : \dim_H F_y(\boldsymbol{\alpha}) = t_0\}$, and fix $t \geq 0$ such that $\mathcal{H}^t\big(F_y(\boldsymbol{\alpha})\big) > c$ for all $y \in A$ and some constant $c > 0$. The Marstrand Cartesian product theorem [Mar54] states that, for any $r \geq 0$, there is a constant $k > 0$ such that

$$\mathcal{H}^{r+t}\big(F(\boldsymbol{\alpha})\big) \geq ck \cdot \mathcal{H}^r(A).$$

By definition, $\mathcal{H}^r(A) = +\infty$ for all $0 \leq r < \dim_H A$ so we see that $\dim_H F(\boldsymbol{\alpha}) \geq r + t$. Since this holds for all $0 \leq t < \dim_H F_y(\boldsymbol{\alpha})$, we find for any $y \in A$ that

$$\dim_H F(\boldsymbol{\alpha}) \geq \dim_H F_y(\boldsymbol{\alpha}) + \dim_H A.$$

By definition of A and Theorem 2.2, we know that $\dim_H F_y(\boldsymbol{\alpha}) = t_0$ so it only remains to compute $\dim_H A$. By Proposition 2.9 and Theorem 2.2, we see that $\nu_{\boldsymbol{\alpha}}(A) = 1$ so we have

$$\dim_H A \geq \inf\{\dim_H F : F \subset [0,1] \text{ s.t. } \nu_{\boldsymbol{\alpha}}(F) = 1\} = \dim_H \nu_{\boldsymbol{\alpha}}.$$

We show that

$$\dim_H \nu_{\boldsymbol{\alpha}} \geq \frac{-\sum_{s \in \mathcal{S}} \alpha_s \log \alpha_s}{-\sum_{s \in \mathcal{S}} \alpha_s \log p_s}.$$

By Lemma 1.5(ii), it suffices to show that for $\nu_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$,

$$\underline{d}_{\nu_{\boldsymbol{\alpha}}}(y) \geq \frac{-\sum_{s \in \mathcal{S}} \alpha_s \log \alpha_s}{-\sum_{s \in \mathcal{S}} \alpha_s \log p_s}.$$

So, fix $y \in Y(\boldsymbol{\alpha})$, and write $\omega \in \Omega$ for the unique digit sequence associated to y (recall Lemma 2.6). Note that the measure $\nu_{\boldsymbol{\alpha}} = \mu_{\boldsymbol{\alpha}} \circ \pi_2^{-1}$ can be seen as the push-back of the Bernoulli measure on Ω determined by the probability vector $(\alpha_s)_{s \in \mathcal{S}}$. Therefore, we are in a situation very similar to that of [BI09, Theorem 3.1] but with finitely many digits. Indeed, we can define $T : [0,1] \to [0,1]$ by $T(y) = f_s^{-1}(y)$ if $y \in \big(\sum_{i=1}^{s-1} p_i, \sum_{i=0}^{s} p_i)\big)$ and $T(y) = 0$ otherwise. Then $\nu_{\boldsymbol{\alpha}}$ is an ergodic T-invariant measure, and we may therefore use the same methods as in the proof of [BI09, Theorem 3.1] to find that

$$\underline{d}_{\nu_{\boldsymbol{\alpha}}}(y) \geq \frac{-\sum_{s \in \mathcal{S}} \alpha_s \log \alpha_s}{-\sum_{s \in \mathcal{S}} \alpha_s \log p_s}$$

for $\nu_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$ as desired. Hence,

$$\begin{aligned}\dim_H F(\boldsymbol{\alpha}) &\geq \frac{\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d - \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{\sum_{d\in\mathcal{D}} \alpha_d \log q_d} + \frac{-\sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{-\sum_{s\in\mathcal{S}} \alpha_s \log p_s} \\ &= \frac{h_{\mu_{\boldsymbol{\alpha}}}(\sigma) - h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{d\in\mathcal{D}} \alpha_d \log q_d} + \frac{h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{s\in\mathcal{S}} \alpha_s \log p_s}.\end{aligned}$$

Example 2.15. For a concrete example, take $\mathcal{S} = \{1, 2\}$, $\mathcal{B}_1 = \{1, 2\}$, $\mathcal{B}_2 = \{1, 2, 3\}$ so that $\mathcal{D} = \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3)\}$. Let

$$\begin{aligned} h_{(1,b)}(x) &= \frac{x+b}{2}, \quad b \in \mathcal{B}_1, \\ h_{(2,b)}(x) &= \frac{x+b}{3}, \quad b \in \mathcal{B}_2. \end{aligned}$$

So, $\varepsilon_d = 0$ for all $d \in \mathcal{D}$, and $r_{(1,1)} = \frac{1}{2}$, $r_{(2,1)} = \frac{1}{3}$ and $r_{(2,2)} = \frac{2}{3}$. Take $p \in (0, 1)$ arbitrary, and let $p_1 = p$. Then $f_1(y) = py$ and $f_2(y) = (1 - p)y + p$. This gives

$$\mathbf{A}_{(1,b)} = \begin{bmatrix} 1/2 & 0 \\ 0 & p \end{bmatrix}, \quad \mathbf{A}_{(2,b)} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1-p \end{bmatrix}$$

and

$$\mathbf{v}_{(1,1)} = \mathbf{v}_{(2,1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_{(1,2)} = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_{(2,2)} = \begin{bmatrix} 1/3 \\ p \end{bmatrix}, \quad \mathbf{v}_{(2,3)} = \begin{bmatrix} 2/3 \\ p \end{bmatrix}.$$

See Figure 2.1(b) for an illustration of how this GLS IFS $\{\mathbf{A}_d + \mathbf{v}_d\}_{d\in\mathcal{D}}$ acts on $[0, 1]^2$. For the number expansions, if $(d_m)_{m\in\mathbb{N}} \in \mathcal{D}^{\mathbb{N}}$, then, for each $m \in \mathbb{N}$, we get $k_m = 0$, $N_m = 2$ if $\omega_m = 0$ and $N_m = 3$ if $\omega_m = 1$, and $t_m \in \{0, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}\}$. So for each $(d_m)_{m\in\mathbb{N}} \in \mathcal{D}^{\mathbb{N}}$, if we set $\kappa(n) = \#\{1 \leq \ell \leq n : \omega_\ell = 0\}$, then (2.7) becomes

$$\lim_{n\to+\infty} h_{d_1} \circ \cdots \circ h_{d_n}(0) = \sum_{n\in\mathbb{N}} \frac{t_n}{2^{\kappa(n)} 3^{n-\kappa(n)}}.$$

Hence, this GLS IFS produces number expansions in mixed base 2 and 3 for each $x \in [0, 1]$. Note that for this IFS, $h_d \neq h_{d'}$ if $d \neq d'$ so from Proposition 2.7, it follows that each $x \in [0, 1]$ has uncountably many different expansions with mixed bases 2 and 3. We have $\alpha_1 = \alpha_{(1,1)} + \alpha_{(1,2)}$ and $\alpha_2 = \alpha_{(2,1)} + \alpha_{(2,2)} + \alpha_{(2,3)}$ so Theorem 2.2 yields for $\nu_{\boldsymbol{\alpha}}$-a.e. $y \in Y(\boldsymbol{\alpha})$ that

$$\dim_H F_y(\boldsymbol{\alpha}) = \frac{-\sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d - (-\alpha_1 \log \alpha_1 - \alpha_2 \log \alpha_2)}{\alpha_1 \log 2 + \alpha_2 \log 3}.$$

Note that if we consider the IFS $\{\mathbf{A}_{(1,1)} + \mathbf{v}_{(1,1)}, \mathbf{A}_{(2,3)} + \mathbf{v}_{(2,3)}\}$ with $\mathbf{A}_{(s,b)}$ and $\mathbf{v}_{(s,b)}$ as in the example and we take $\frac{1}{2} < p < \frac{2}{3}$, then we obtain a diagonally affine IFS that satisfies condition (D')(a). Hence, we can apply Theorem 2.1 to this IFS to obtain for each $\alpha \in (0, 1)$ an expression for the Hausdorff dimension of the set of points $(x, y) \in [0, 1]^2$ that have a GLS expansion containing only the digits $(1, 1)$ and $(2, 3)$ and in which $(1, 1)$ occurs with frequency α (and thus $(2, 3)$ with frequency $1 - \alpha$). Of course, here we can take any other combination of a digit from $\{(1, 1), (1, 2)\}$ and a digit from $\{(2, 1), (2, 2), (2, 3)\}$ to obtain a similar result.

We can extend this example in the following sense. Fix some finite set $\mathcal{S} = \{1, \ldots, S\} \subset \mathbb{N}$ and a finite branch-indexing set $\mathcal{B}_s = \{1, \ldots, B_s\}$ for each $s \in \mathcal{S}$. So,

$$\mathcal{D} = \{(s, b) : b \in \mathcal{B}_s,\ s \in \mathcal{S}\}.$$

Also fix some probability vector $(p_s)_{s\in\mathcal{S}}$. For $(s,b)\in\mathcal{D}$, set

$$\mathbf{A}_{(s,b)} = \begin{bmatrix} 1/B_s & 0 \\ 0 & p_s \end{bmatrix}, \quad \mathbf{v}_{(s,b)} = \begin{bmatrix} b/B_s \\ \sum_{i=1}^{s-1} p_i \end{bmatrix}.$$

For each $x\in[0,1]$ and sequence $\omega\in\Omega$, the GLS expansion produced by this system has the form

$$x = \sum_{n\in\mathbb{N}} \frac{b_n}{B_1^{c_{1,n}} B_2^{c_{2,n}} \cdots B_S^{c_{S,n}}},$$

with $b_n \in \mathcal{B}_{\omega_n}$ and $c_{s,n} = \#\{1\le \ell \le n : \omega_\ell = s\}$. In other words, the system produces for each $x\in[0,1]$ uncountably many different mixed base expansions with bases $B_1,\dots,B_S$. Here, we need to remark that we consider two GLS expansions produced by the system different if the two corresponding sequences in

2

$$\mathcal{A} = \bigcup_{(s,b)\in\mathcal{D}} \left\{ \left(0, B_s, \frac{b}{B_s}\right) \right\}$$

are different. For the point 0, for example, this means that the GLS expansions generated by the system are all of the expansions of the form

$$0 = \sum_{n\in\mathbb{N}} \frac{0}{B_1^{c_{1,n}} \cdots B_S^{c_{S,n}}},$$

with $(c_{1,n},\dots,c_{S,n})\in\mathbb{N}^S$ satisfying $\sum_{s\in\mathcal{S}} c_{s,n} = n$ for each $n\in\mathbb{N}$.

Example 2.16. Fix an $N\in\mathbb{N}$ with $N\ge 3$ and a $0<p<1$. Set $\mathcal{B}=\{1,\dots,N\}$ and $\mathcal{D} := \{(s,b) : s\in\{0,1\},\, b\in\mathcal{B}\}$. For $b\in\mathcal{B}$, set

$$\mathbf{A}_{(1,b)} = \begin{bmatrix} 1/N & 0 \\ 0 & p \end{bmatrix}, \quad \mathbf{A}_{(2,b)} = \begin{bmatrix} -1/N & 0 \\ 0 & 1-p \end{bmatrix}$$

and

$$\mathbf{v}_{(1,b)} = \begin{bmatrix} b/N \\ 0 \end{bmatrix}, \quad \mathbf{v}_{(2,b)} = \begin{bmatrix} (b+1)/N \\ p \end{bmatrix}.$$

Then for any $x\in[0,1]$ and any $\omega\in\{0,1\}^{\mathbb{N}}$, the number expansion of x produced by this system has the form

$$x = \sum_{n\in\mathbb{N}} (-1)^{\omega_n} \frac{b_n}{N^n},$$

for some sequence $(b_n)_{n\in\mathbb{N}}\in\mathcal{B}^{\mathbb{N}}$. So, the system produces for each x a signed base N-expansion in which the signs of the terms correspond to a preset sequence of signs ω.

Additionally, this system satisfies $h_d \ne h_{d'}$ whenever $d\ne d'$ so, for any frequency vector $\boldsymbol{\alpha}\in(0,1)^{2N}$, the Hausdorff dimension of the Besicovitch-Eggleston set $F_y(\boldsymbol{\alpha})$ for $\nu_{\boldsymbol{\alpha}}$-a.e. $y\in Y(\boldsymbol{\alpha})$ is given by Theorem 2.2. For $\frac{1}{N} < p < \frac{N-1}{N}$ and any $b,c\in\mathcal{B}$, the system $\{\mathbf{A}_{(1,b)}+\mathbf{v}_{(1,b)}, \mathbf{A}_{(2,c)}+\mathbf{v}_{(2,c)}\}$ satisfies (D')(a) so Theorem 2.1 also applies.

Chapter 3

Besicovitch-Eggleston Sets for Non-autonomous GLS Maps

Abstract

We derive a formula for the Hausdorff dimension of Besicovitch-Eggleston level sets associated with non-autonomous dynamics constructed from families $\mathcal{T}$ of GLS maps, which we refer to as Non-autonomous Generalised Lüroth Series (NGLS) maps. The formula obtained shows that the universal-lower-bound phenomenon present in the autonomous case studied in [FLMW10] persists in this non-autonomous setting.

3

3.1 Introduction

Recall from §1.3.1 that the Lüroth map is the piecewise linear map $L : [0,1] \to [0,1]$ defined by $L(0) = 0$ and

$$L(x) = k(k+1)x - k \quad \text{whenever } x \in \left(\frac{1}{k+1}, \frac{1}{k}\right], \; k \in \mathbb{N};$$

see Figure 3.1(c). For each $k \in \mathbb{N}$, write $N_{L,k} := k(k+1)$ for the slope of the branch of L supported on the interval $(\frac{1}{k+1}, \frac{1}{k}]$. Every $x \in [0,1] \setminus \mathbb{Q}$ has a unique associated *Lüroth series expansion*

$$x = \sum_{n \in \mathbb{N}} \frac{b_n(x)}{\prod_{1 \leq \ell \leq n} N_{L,b_\ell(x)}},$$

where $b_n(x) := k$ whenever $L^{n-1}(x) \in (\frac{1}{k+1}, \frac{1}{k}]$ for $n \in \mathbb{N}$. Such expansions were first studied by Lüroth in [Lür83], and the map was first introduced in [JV69]. We refer to the numbers $b_n(x)$ as the *digits* of the Lüroth series expansion of x. In [FLMW10], Fan et al. obtained a formula for the Hausdorff dimension of the *Besicovitch-Eggleston* level sets for Lüroth series expansions. More precisely, for any frequency vector $\boldsymbol{\alpha} = (\alpha_k)_{k \in \mathbb{N}}$, the authors of [FLMW10] obtained the following result for the set $F_L(\boldsymbol{\alpha})$ of numbers in $[0,1] \setminus \mathbb{Q}$ that have each digit $k \in \mathbb{N}$ appear in their Lüroth series expansion with frequency α_k.

Theorem 3.1 ([FLMW10])**.** *Fix a frequency vector $\boldsymbol{\alpha} = (\alpha_k)_{k \in \mathbb{N}}$. Then the Hausdorff dimension of $F_L(\boldsymbol{\alpha})$ is*

$$\dim_H F_L(\boldsymbol{\alpha}) = \max\left\{\frac{1}{2}, \liminf_{m \to +\infty} \frac{-\sum_{1 \leq k \leq m} \alpha_k \log \alpha_k}{\sum_{1 \leq k \leq m} \alpha_k \log N_{L,k}}\right\}.$$

One of the important observations made in [FLMW10] is the universal lower bound of $\frac{1}{2}$ in the formula from Theorem 3.1, which occurs due to the fact that the digit set of the map L is infinite. This $\frac{1}{2}$ is the *exponent of convergence* of the sequence $(N_{L,k}^{-1})_{k\in\mathbb{N}}$ of lengths of the supports $(\frac{1}{k+1}, \frac{1}{k}]$ of the branches of the Lüroth map:

$$\frac{1}{2} = \inf\left\{ t \geq 0 : \sum_{n\in\mathbb{N}} N_{L,n}^{-t} < +\infty \right\}.$$

In the second term in the formula from Theorem 3.1, we recognise the entropy and the Lyapunov exponent for the $\boldsymbol{\alpha}$-Bernoulli measure on $\mathbb{N}$.

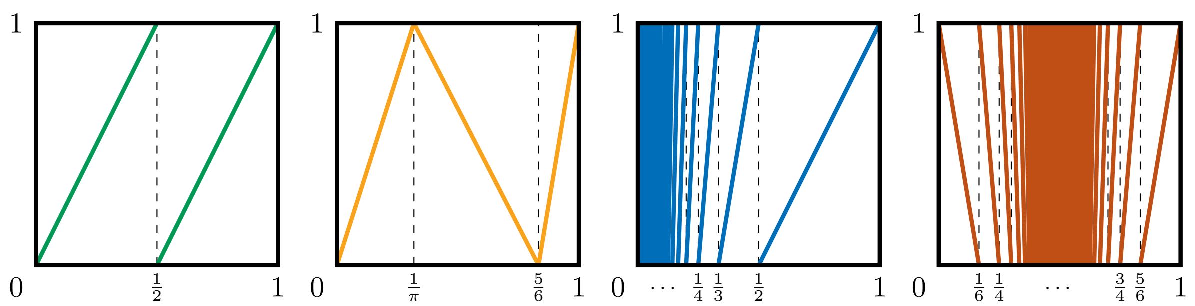
The main result from [FLMW10] is actually stronger than stated in Theorem 3.1 in that it also holds if the Lüroth map is replaced by any other piecewise linear interval map that can be modelled by a full shift on countably many symbols. In particular, this includes GLS maps; recall their definition from Chapter 2 and see Figure 3.1 for examples.

In this paper, we obtain a result similar to Theorem 3.1 but for the Besicovitch-Eggleston level sets $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ of non-autonomous dynamical systems $(\mathcal{T}, \omega)$, where $\mathcal{T} := \{T_s : [0,1] \to [0,1]\}_{s\in\mathcal{S}}$ is a collection of GLS maps indexed by a finite set $\mathcal{S}$ and $\omega \in \Omega := \mathcal{S}^{\mathbb{N}}$. In other words, the dynamics of the system are given non-autonomously by

$$T_\omega^n := T_{\omega_n} \circ \cdots \circ T_{\omega_2} \circ T_{\omega_1}, \quad n \in \mathbb{N}. \tag{3.1}$$

We refer to the pair $(\mathcal{T}, \omega)$ as a *Non-autonomous Generalised Lüroth Series* (NGLS) map. Note that these maps are essentially the same as the fibres of GLS maps with redundancy studied in Chapter 2 except that the fibre here will be kept as a sequence on $\mathcal{S}$ rather than be mapped to a point in $[0,1]$ as was done in Chapter 2. This is because we will not explore the associated two-dimensional skew product or related IFSs on the unit square in this chapter so there is no reason to map the fibre off of the symbolic space. It is also worth emphasising that, in contrast with the GLS IFSs treated in Chapter 2, the collections $\mathcal{T}$ of GLS maps considered in this chapter will *not* have the restriction that their maps each have only finitely many branches.

Similar to the expansions obtained from GLS IFSs in Chapter 2 (recall (2.9) from §2.1 and Proposition 2.7), we can obtain number expansions from an NGLS map $(\mathcal{T}, \omega)$. For a fixed point, these two expansions (with respect to the fibre ω) will be the same, but we shall nevertheless reconstruct them here in this slightly altered setting. For each $s \in \mathcal{S}$, write $\{I_{s,b} : b \in \mathcal{B}_s\}$ for the interval partition of the GLS map $T_s \in \mathcal{T}$. Define $b_n = b_n(\omega, x) := b$ whenever



(a) The doubling map (b) Finite-branched GLS (c) The Lüroth map (d) Infinite-branched GLS

Figure 3.1: Examples of GLS maps with finitely or infinitely many branches. The Lüroth map has infinitely many branches accumulating at zero. The GLS map in (d) has infinitely many branches accumulating at 1/2.

$T_{\omega}^{n-1}(x) \in I_{\omega_n,b}$, $n \in \mathbb{N}$. When this sequence exists, i.e. when it holds for each $n \in \mathbb{N}$ that $T_{\omega}^{n-1}(x) \in I_{\omega_n,b}$ for some $b \in \mathcal{B}_{\omega_n}$, then x can be expressed as

$$x = \sum_{n\in\mathbb{N}} \frac{(-1)^{\varepsilon_{\omega_1,b_1}+\cdots+\varepsilon_{\omega_{n-1},b_{n-1}}} t_{\omega_n,b_n}}{\prod_{1\le \ell\le n} N_{\omega_\ell,b_\ell}}, \tag{3.2}$$

for some quantities $\varepsilon_{\omega_n,b_n} \in \{0,1\}$, $N_{\omega_n,b_n} > 1$ and $t_{\omega_n,b_n} \in (0,1)$ that relate to the properties of the map T_{ω_n}, which we will define explicitly later in §3.2.

From (3.2), we see that the NGLS expansion of x is completely determined by the sequence $(\omega_n, b_n(\omega,x))_{n\in\mathbb{N}}$ so we refer to $\mathcal{D} := \{(s,b) : b \in \mathcal{B}_s, s \in \mathcal{S}\}$ as the *digit set* of $(\mathcal{T},\omega)$ and its elements as *digits* of NGLS expansions. We refer to an expansion as in (3.2) as an *NGLS expansion* and to the sequence $(b_n)_{n\in\mathbb{N}}$ as the $(\mathcal{T},\omega)$-*expansion* of x.

Our main results are on spectra for digit frequency level sets generated by NGLS maps $(\mathcal{T},\omega)$. For each $d \in \mathcal{D}$ and each frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$, these sets are the Besicovitch-Eggleston $\boldsymbol{\alpha}$-level sets for $(\mathcal{T},\omega)$ given by

$$F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) := \left\{ x \in [0,1] : \lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : (\omega_\ell, b_\ell(\omega,x)) = d\}}{n} = \alpha_d \text{ for all } d \in \mathcal{D} \right\}.$$

The *dimension spectrum* is $\boldsymbol{\alpha} \mapsto \dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$. Set $\Omega_{\mathcal{T}}(\boldsymbol{\alpha}) := \{\omega \in \Omega : F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \neq \emptyset\}$, which we call the *fibre spectrum of* $\boldsymbol{\alpha}$. Throughout the chapter, we will assume the following non-degeneracy condition on the frequency vectors $\boldsymbol{\alpha}$:

$$\text{For each } s \in \mathcal{S}, \text{ there exists a } b \in \mathcal{B}_s \text{ such that } \alpha_{(s,b)} > 0. \tag{$*$}$$

3

The first result is on the form of the frequency spectra $\Omega_{\mathcal{T}}(\boldsymbol{\alpha})$. Put $\alpha_s := \sum_{b\in\mathcal{B}_s} \alpha_{(s,b)}$ for each $s \in \mathcal{S}$.

Theorem 3.2. (Fibre Spectra) *Fix a finite collection* $\mathcal{T} := \{T_s : [0,1] \to [0,1]\}_{s\in\mathcal{S}}$ *of GLS maps and a frequency vector* $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ *satisfying* $(*)$. *Then*

$$\Omega_{\mathcal{T}}(\boldsymbol{\alpha}) = \left\{ \omega \in \Omega : \lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : \omega_\ell = s\}}{n} = \alpha_s \text{ for all } s \in \mathcal{S} \right\}.$$

This result is an analogue of Proposition 2.9 from Chapter 2 for NGLS maps. In fact, the statement of Theorem 3.2 is more general as we do not impose that the maps of $\mathcal{T}$ have only finitely many branches.

Before we state the second main result, we set some further notation. For each $s \in \mathcal{S}$, define the *exponent of convergence* of the map T_s by

$$\eta(T_s) := \inf\left\{ t \ge 0 : \sum_{b\in\mathcal{B}_s} N_{s,b}^{-t} < +\infty \right\} \in [0,1],$$

and put $\eta_{\mathcal{T}} := \max\{\eta(T_s) : s \in \mathcal{S}\}$. In addition, define the *fibre dimension* with respect to $\mathcal{T}$ and $\boldsymbol{\alpha}$ by

$$\beta_{\mathcal{T}}(\boldsymbol{\alpha}) := \liminf_{m\to+\infty} \frac{\sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s - \sum_{d\in\mathcal{D}_m} \alpha_d \log \alpha_d}{\sum_{d\in\mathcal{D}_m} \alpha_d \log N_d},$$

where $\mathcal{D}_m := \{(s,b) \in \mathcal{D} : b \le m\}$. Put $\mathcal{S}_{\mathbb{N}} := \{s \in \mathcal{S} : \mathcal{B}_s = \mathbb{N}\}$.

Theorem 3.3. (Dimension Spectra) *Fix a finite collection* $\mathcal{T} := \{T_s : [0,1] \to [0,1]\}_{s \in \mathcal{S}}$ *of GLS maps, a frequency vector* $\boldsymbol{\alpha} = (\alpha_d)_{d \in \mathcal{D}}$ *satisfying* $(*)$ *and an* $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$*. Assume that the limit*

$$\lim_{n \to +\infty} \frac{\log n}{\log N_{s,n}} \tag{3.3}$$

exists for all $s \in \mathcal{S}_{\mathbb{N}}$*. Then*

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = \max\{\eta_{\mathcal{T}}, \beta_{\mathcal{T}}(\boldsymbol{\alpha})\}. \tag{3.4}$$

Furthermore, if $\sum_{d \in \mathcal{D}} \alpha_d \log N_d = +\infty$*, then* $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = \eta_{\mathcal{T}}$*.*

Comparing (3.4) with the statement from Theorem 3.1, one can see the effect of the non-autonomous setting in both terms. The first term is now given by the maximal exponent of convergence over all maps in $\mathcal{T}$. In the second term, the measure-theoretic entropy of the $\boldsymbol{\alpha}$-Bernoulli measure appearing in the numerator of the expression from Theorem 3.1 is now replaced by the fibre entropy of that measure for the non-autonomous system; see §3.2.

The limit in condition (3.3), when it exists, is in fact an expression for $\eta(T_s)$ – see (3.19) in §3.5. The condition is quite unrestrictive and holds, for example, if $n \mapsto N_{s,n}$ is a regularly varying function (in the sense of [BGT87]) for each $s \in \mathcal{S}_{\mathbb{N}}$.

We can obtain Theorem 3.1 as a subcase of Theorem 3.3 by taking $\mathcal{T} = \{T_L\}$ to contain only the Lüroth map (so $\mathcal{D} = \{L\} \times \mathbb{N}$). Then any frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d \in \mathcal{D}}$ automatically satisfies $(*)$ since $\alpha_L := \sum_{n \in \mathbb{N}} \alpha_{(L,n)} = 1 > 0$. In addition, $n \mapsto N_{L,n} := n(n+1)$ is regularly varying with index 2 so it follows that (3.3) holds. Theorem 3.3 also includes the case where $\mathcal{T}$ contains only GLS maps with finitely many branches, in which case the digit set $\mathcal{D}$ is finite and $\eta(T_s) = 0$ for each $s \in \mathcal{S}$ so (3.3) holds and (3.4) becomes

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = \frac{\sum_{s \in \mathcal{S}} \alpha_s \log \alpha_s - \sum_{d \in \mathcal{D}} \alpha_d \log \alpha_d}{\sum_{d \in \mathcal{D}} \alpha_d \log N_d},$$

which is an analogue of the statement in Theorem 2.2. As the result of Theorem 2.2 holds only for almost all fibres, Theorem 3.3 is a generalisation of this.

Other generalisations of [FLMW10, Theorem 1.1] (and Theorem 3.1) include [GL16] to group frequencies of digits in Lüroth series expansions and [Rus23] to countable symbolic dynamical systems modelled by subshifts of finite type (both in the autonomous setting). A non-autonomous result similar in flavour to Theorem 3.3 is that of [NT24], which treats irrational numbers whose semi-regular continued fraction expansions satisfy certain growth conditions. Dynamics of the form (3.1) are also considered in the setting of random dynamics and the associated skew products. There are many results in this direction for systems related to number expansions including the study of random continued fraction expansions in [KKV17; DO18; KMTV22; BDKK+24], random binary expansions in [DK20], random β-expansions in [DK03; DV05; DK07; DV07; Kem14; BD17; Suz19; Tie23] and random Lüroth series expansions in [KM22a; KM22b; GKKS25].

The chapter is organised as follows. In §3.2, we introduce some further notation concerning NGLS maps and Besicovitch-Eggleston sets. The proof of Theorem 3.2 is given in §3.3. §3.4 proves the upper bound of Theorem 3.3 in Proposition 3.4. The proof of the lower bound $\eta_{\mathcal{T}}$ is given in §3.5 – see Proposition 3.9. The remainder of the proof of Theorem 3.3 is given in §3.6. In particular, the proof of the lower bound $\beta_{\mathcal{T}}(\boldsymbol{\alpha})$ is given in Proposition 3.17.

3.2 Preliminaries

In this section, we fix some notation and collect some preliminaries that are used throughout the chapter. Recall that for any subset $U \subseteq \mathbb{R}$, the notation $|U| \coloneqq \sup\{|x-y| : x, y \in U\}$ denotes the diameter of U. In addition, recall for any $u \in \mathcal{U}$, $n \in \mathbb{N}$ and sequence $\xi = (\xi_n)_{n\in\mathbb{N}} \in \mathcal{U}^{\mathbb{N}}$ that $\tau_u(\xi, n) = \#\{1 \leq \ell \leq n : \xi_\ell = u\}$ denotes the number of times that the symbol u occurs in the first n elements of ξ and

$$\tau_u(\zeta) = \lim_{n\to+\infty} \frac{\tau_u(\zeta, n)}{n}$$

denotes the frequency of the symbol u in the sequence ξ if the limit exists.

3.2.1 Definition of Non-autonomous GLS Maps

Throughout the chapter, we fix the following system. Let $\mathcal{S}$ be a finite set and let $\mathcal{T} = \{T_s\}_{s\in\mathcal{S}}$ be a family of GLS maps $T_s : [0,1] \to [0,1]$. Recall that this means that each map T_s has an associated countable index set $\mathcal{B}_s \subseteq \mathbb{N}$ and is given by two pieces of data:

- a collection of disjoint non-degenerate open intervals $\{I_{s,b} : b \in \mathcal{B}_s\}$ with $\sum_{b\in\mathcal{B}_s} |I_{s,b}| = 1$
- a sequence $(\varepsilon_{s,b})_{b\in\mathcal{B}_s}$ on $\{0, 1\}$ that specifies for each interval $I_{s,b}$ the orientation of T_s in the sense that the sign of T_s' on $I_{s,b}$ equals $(-1)^{\varepsilon_{s,b}}$

We call the index set $\mathcal{B}_s$ of the interval partition of the map T_s the *branch-indexing set* of T_s. We let $\mathcal{B}_s \coloneqq \mathbb{N}$ in the case that $\{I_{s,b} : b \in \mathcal{B}_s\}$ is an infinite set, and, otherwise, we let $\mathcal{B}_s \coloneqq \{1, 2, \ldots, B_s\}$ for some $B_s \in \mathbb{N}$. We assume that the partition elements are ordered by decreasing length so that, for $b, b' \in \mathcal{B}_s$ with $b < b'$, it holds that $|I_{s,b}| \geq |I_{s,b'}|$. Write $\ell_{s,b}$ for the left endpoints of the interval $I_{s,b}$, and set $N_{s,b} \coloneqq |I_{s,b}|^{-1}$ and $t_{s,b} \coloneqq \varepsilon_{s,b} + \ell_{s,b} N_{s,b}$. Then the map T_s is given on $I_{s,b}$ by 3

$$T_s(x) = (-1)^{\varepsilon_{s,b}}(N_{s,b}x - t_{s,b}).$$

We set $T_s(x) \coloneqq 0$ for any $x \notin \bigcup_{b\in\mathcal{B}_s} I_{s,b}$.

Set $\Omega \coloneqq \mathcal{S}^{\mathbb{N}}$, and, for each $\omega \in \Omega$, let

$$X_\omega \coloneqq \left\{ x \in [0,1] : T_\omega^n(x) \in \bigcup_{b\in\mathcal{B}_{\omega_{n+1}}} I_{\omega_{n+1},b} \ \text{ for all } n \in \mathbb{N} \right\}. \tag{3.5}$$

Notice that $[0,1] \setminus X_\omega$ is a countable set. For any $n \in \mathbb{N}$, let $\mathcal{B}_\omega^n \coloneqq \prod_{1\leq\ell\leq n} \mathcal{B}_{\omega_\ell}$ denote the collection of all blocks $b_1 \cdots b_n$ with the property that $b_\ell \in \mathcal{B}_{\omega_\ell}$ for all $1 \leq \ell \leq n$. The map T_ω^n is piecewise affine on what we call the *level-n fibre fundamental intervals* (FFIs), which are the intervals

$$\begin{aligned}\langle b_1 \cdots b_n \rangle_\omega &\coloneqq I_{\omega_1,b_1} \cap T_{\omega_1}^{-1} I_{\omega_2,b_2} \cap \cdots \cap (T_\omega^{n-1})^{-1} I_{\omega_n,b_n} \\ &= \{x \in [0,1] : T_\omega^k(x) \in I_{\omega_{k+1},b_{k+1}} \quad \text{for all } 0 \leq k < n\},\end{aligned}$$

for $b_1 \cdots b_n \in \mathcal{B}_\omega^n$. For $x \in X_\omega$, let the sequence $(b_n(\omega, x))_{n\in\mathbb{N}}$ be given as in the introduction, i.e. for $n \in \mathbb{N}$, we set $b_n = b_n(\omega, x) = b$ if $T_\omega^{n-1}(x) \in I_{\omega_n,b}$. Then for each $x \in X_\omega$ and $n \in \mathbb{N}$,

$$x = \frac{t_{\omega_1,b_1}}{N_{\omega_1,b_1}} + (-1)^{\varepsilon_{\omega_1,b_1}} \frac{T_{\omega_1}(x)}{N_{\omega_1,b_1}} = \cdots = \sum_{k=1}^{n} \frac{(-1)^{\sum_{1\leq\ell<k} \varepsilon_{\omega_\ell,b_\ell}} t_{\omega_k,b_k}}{\prod_{1\leq\ell\leq k} N_{\omega_\ell,b_\ell}} + \frac{(-1)^{\sum_{1\leq\ell\leq n} \varepsilon_{\omega_\ell,b_\ell}} T_\omega^n(x)}{\prod_{1\leq\ell\leq n} N_{\omega_\ell,b_\ell}},$$

and, as $T_\omega^n(x) \in [0,1]$ and $N_{\omega_n,b} > 1$ for each n and b, the last term converges as $n \to +\infty$ and so we obtain the expression from (3.2). The sequence $(b_n(\omega, x))_{n\in\mathbb{N}}$ satisfies $b_n(\omega, x) \in \mathcal{B}_{\omega_n}$ for each $n \in \mathbb{N}$. Denote by $\mathcal{B}_\omega^{\mathbb{N}} := \prod_{n\in\mathbb{N}} \mathcal{B}_{\omega_n}$ the set of such sequences, and define the projection $\pi_\omega : \mathcal{B}_\omega^{\mathbb{N}} \to [0,1]$ by

$$\pi_\omega\big((b_n)_{n\in\mathbb{N}}\big) = \sum_{n\in\mathbb{N}} \frac{(-1)^{\sum_{1\le \ell<n} \varepsilon_{\omega_\ell,b_\ell}} t_{\omega_n,b_n}}{\prod_{1\le\ell\le n} N_{\omega_\ell,b_\ell}}. \tag{3.6}$$

By the definition of X_ω, for any $x \in X_\omega$, there exists a unique sequence $(b_n)_{n\in\mathbb{N}}$ such that $x = \pi_\omega((b_n)_{n\in\mathbb{N}})$, i.e. the $(\mathcal{T},\omega)$-expansion of x is unique. The map π_ω is not injective: the set $(0,1)\setminus X_\omega$ contains both the points that have multiple sequences mapped to them, i.e. that are mapped to the common boundary of two FFIs after some iterations, and the points that have no sequences mapped to them, i.e. that are mapped to some point that does not lie in the closure of any FFI after some iterations. Note that the points 0 and 1 have at most one $(\mathcal{T},\omega)$-expansion.

Let $\mathcal{F}(n)$ denote the collection of all level-n FFIs. Note that for each $n \in \mathbb{N}$, $\langle b_1\cdots b_n\rangle_\omega \cap \langle c_1\cdots c_n\rangle_\omega = \emptyset$ whenever $b_1\cdots b_n \neq c_1\cdots c_n$ and that

$$|\langle b_1\cdots b_n\rangle_\omega| = \prod_{1\le\ell\le n} \frac{1}{N_{\omega_\ell,b_\ell}} \quad\text{and}\quad \Bigg|\bigcup_{b_1\cdots b_n\in\mathcal{B}_\omega^n} \langle b_1\cdots b_n\rangle_\omega\Bigg| = 1.$$

Since $\langle b_1\cdots b_n\rangle_\omega \subseteq \langle b_1\cdots b_{n-1}\rangle_\omega$ for each n and $\lim_{n\to+\infty} |\langle b_1\cdots b_n\rangle_\omega| = 0$, the collection

$$\mathcal{F} := \bigcup_{n\in\mathbb{N}} \mathcal{F}(n) = \big\{\langle b_1\cdots b_n\rangle_\omega : b_1\cdots b_n \in \mathcal{B}_\omega^n,\ n\in\mathbb{N}\big\}$$

3

generates the Borel σ-algebra $\mathcal{B}([0,1])$ on $[0,1]$.

3.3 Fibre Spectra

Let $\mathcal{T} = \{T_s\}_{s\in\mathcal{S}}$ be any finite collection of GLS maps and let $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ be a frequency vector satisfying $(*)$. Recall that the fibre spectrum $\Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ of $\boldsymbol{\alpha}$ is the set of $\omega\in\Omega$ for which $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ is non-empty, and recall that $\alpha_s := \sum_{b\in\mathcal{B}_s} \alpha_{(s,b)}$. In this section, we shall prove Theorem 3.2, which asserts that $\Omega_{\mathcal{T}}(\boldsymbol{\alpha}) = \{\omega\in\Omega : \tau_s(\omega) = \alpha_s\ \forall s\in\mathcal{S}\}$.

Proof of Theorem 3.2. First, suppose that $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$, and take $x \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$. Then

$$\tau_s(\omega) = \sum_{b\in\mathcal{B}_s} \tau_{(s,b)}(\omega,x) = \sum_{b\in\mathcal{B}_s} \alpha_{(s,b)} = \alpha_s$$

for all $s\in\mathcal{S}$, which proves the forward direction.

To prove the converse statement, take $\omega\in\Omega$ such that $\tau_s(\omega) = \alpha_s$ for all $s\in\mathcal{S}$. For each $s\in\mathcal{S}$, it follows from the proof of Proposition 2.9 when $|\mathcal{B}_s| < +\infty$ and from e.g. [FLMW10, Lemma 2.1] or [BK24, Theorem 3.1] when $\mathcal{B}_s = \mathbb{N}$ that there exists a sequence $(d_{s,n})_{n\in\mathbb{N}}$ of digits in $\{s\}\times\mathcal{B}_s$ whose digit frequencies are given by the frequency vector $(\alpha_{(s,b)}/\alpha_s)_{b\in\mathcal{B}_s}$:

$$\lim_{n\to+\infty} \frac{\#\{1\le\ell\le n : d_{s,\ell} = (s,b)\}}{n} = \frac{\alpha_{(s,b)}}{\alpha_s}, \qquad \forall b\in\mathcal{B}_s. \tag{3.7}$$

We weave these sequences together to define a sequence $(d_n)_{n\in\mathbb{N}}$ with $d_n = (\omega_n, b_n)$ for some $b_n \in \mathcal{B}_{\omega_n}$, $n\in\mathbb{N}$, and with digit frequencies given by $\boldsymbol{\alpha}$. Recall that $\tau_s(\omega,n)$ denotes the number of times that the symbol $s\in\mathcal{S}$ appears in the first block of $n\in\mathbb{N}$ symbols of ω. For each $n \in \mathbb{N}$, set $d_n := d_{\omega_n, \tau_{\omega_n}(\omega,n)}$. Then for each $(s,b) \in \mathcal{D}$, it follows from the assumption $\tau_s(\omega) = \alpha_s$ and (3.7) that

$$\lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : d_\ell = (s,b)\}}{n} = \lim_{n\to+\infty} \frac{\#\{1 \le \ell \le \tau_s(\omega,n) : d_{s,\ell} = (s,b)\}}{\tau_s(\omega,n)} \cdot \frac{\tau_s(\omega,n)}{n}$$
$$= \frac{\alpha_{(s,b)}}{\alpha_s} \cdot \alpha_s(\omega) = \alpha_{(s,b)},$$

as desired.

It remains to show that $(d_n)_{n\in\mathbb{N}}$ implies the existence of a point $x \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$. Put $x := \pi_\omega((b_n)_{n\in\mathbb{N}})$, where π_ω is the projection defined in (3.6). If $(d_n)_{n\in\mathbb{N}}$ gives the $(\mathcal{T},\omega)$-expansion of x, then $x \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ and we are done. Therefore, suppose that the $(\mathcal{T},\omega)$-expansion of x is not given by the sequence $(d_n)_{n\in\mathbb{N}}$. This means that x is an endpoint of the FFI $\langle b_1 \cdots b_\ell \rangle_\omega$ for some $\ell \in \mathbb{N}$, which implies that $\pi_{\sigma^k(\omega)}((b_{n+k})_{n\in\mathbb{N}}) \in \{0,1\}$ for all $k \ge \ell$. For any $k \ge \ell$ there is at most one sequence $(c_n^{(k)})_{n\in\mathbb{N}} \in \mathcal{B}^{\mathbb{N}}_{\sigma^k(\omega)}$ with $\pi_{\sigma^k(\omega)}((c_{k+n})_{n\in\mathbb{N}}) = 0$ and at most one sequence $(e_n^{(k)})_{n\in\mathbb{N}} \in \mathcal{B}^{\mathbb{N}}_{\sigma^k(\omega)}$ with $\pi_{\sigma^k(\omega)}((e_{k+n})_{n\in\mathbb{N}}) = 1$ and one of these sequences equals $(b_{n+k})_{n\in\mathbb{N}}$.

To find a point $y \in F_{\mathcal{T},\omega}(\alpha)$, we change the sequence $(b_n)_{n\in\mathbb{N}}$ to another sequence $(\tilde{b}_n)_{n\in\mathbb{N}} \in \mathcal{B}^{\mathbb{N}}_\omega$ that has the same digit frequencies but for which

$$\pi_{\sigma^k(\omega)}((\tilde{b}_{k+n})_{n\in\mathbb{N}}) \notin \{0,1\}$$

for any $k \in \mathbb{N}$ so that $y = \pi_\omega((\tilde{b}_n)_{n\in\mathbb{N}})$ has a unique $(\mathcal{T},\omega)$-expansion. We obtain the sequence $(\tilde{b}_n)_{n\in\mathbb{N}}$ by setting

$$\tilde{b}_n = \begin{cases} \min\left(\mathcal{B}_{\omega_n} \setminus \{b_n\}\right), & \text{if } n \ge \ell \text{ and } \sqrt{n} \in \mathbb{N}, \\ b_n, & \text{otherwise.} \end{cases}$$

Obviously $(\tilde{b}_{k+n})_{n\in\mathbb{N}} \neq (b_{k+n})_{n\in\mathbb{N}}$ for any $k \ge \ell$ so

$$\pi_{\sigma^k(\omega)}((\tilde{b}_{k+n})_{n\in\mathbb{N}}) \neq \pi_{\sigma^k(\omega)}((b_{k+n})_{n\in\mathbb{N}}) \tag{3.8}$$

for any $k \ge \ell$. Suppose without loss of generality that $(b_{k+n})_{n\in\mathbb{N}} = (e_n^{(k)})_{n\in\mathbb{N}}$. If the sequence $(c_n^{(k)})_{n\in\mathbb{N}}$ also exists, then either there is a j such that $c_{j+n}^{(k)} = e_{j+n}^{(k)}$ for all $n \in \mathbb{N}$, or $c_n^{(k)} \neq e_n^{(k)}$ for all $n \in \mathbb{N}$. This is due to the fact that for any $s \in S$ 0 and 1 are the boundary points of at most one interval $I_{s,b}$. In both cases $(\tilde{b}_{k+n})_{n\in\mathbb{N}} \neq (c_n^{(k)})_{n\in\mathbb{N}}$ and we obtain that for any $k \ge \ell$,

$$\pi_{\sigma^k(\omega)}((\tilde{b}_{k+n})_{n\in\mathbb{N}}) \neq \pi_{\sigma^k(\omega)}((c_n^{(k)})_{n\in\mathbb{N}}),$$

so together with (3.8) we see that $\pi_{\sigma^k(\omega)}((\tilde{b}_{k+n})_{n\in\mathbb{N}}) \notin \{0,1\}$. Hence, the sequence $(\tilde{b}_n)_{n\in\mathbb{N}}$ is the $(\mathcal{T},\omega)$-expansion of the point $\pi_\omega((\tilde{b}_n)_{n\in\mathbb{N}})$. Moreover, for each $d \in \mathcal{D}$ and each $n \in \mathbb{N}$,

$$\frac{\#\{1 \le \ell \le n : (\omega_n, \tilde{b}_n) = d\}}{n} \ge \frac{\#\{1 \le \ell \le n : (\omega_n, b_n) = d\} - \sqrt{n} - 1}{n},$$

and similarly

$$\frac{\#\{1 \le \ell \le n : (\omega_n, \tilde{b}_n) = d\}}{n} \le \frac{\#\{1 \le \ell \le n : (\omega_n, b_n) = d\} + \sqrt{n} + 1}{n}.$$

Therefore,

$$\lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : (\omega_n, \tilde{b}_n) = d\}}{n} = \alpha_d$$

so $y \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ and thus $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$. □

3

3.4 Upper Bound for $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$

Recall that

$$\eta_{\mathcal{T}} = \max_{s\in\mathcal{S}} \inf\left\{ t \geq 0 : \sum_{b\in\mathcal{B}_s} N_{s,b}^{-t} < +\infty \right\},$$
$$\beta_{\mathcal{T}}(\boldsymbol{\alpha}) = \liminf_{m\to+\infty} \frac{\sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s - \sum_{d\in\mathcal{D}_m} \alpha_d \log \alpha_d}{\sum_{d\in\mathcal{D}_m} \alpha_d \log N_d},$$

where $\mathcal{D}_m := \{(s,b) \in \mathcal{D} : b \leq m\}$. In this section, we prove the following proposition, which is precisely the upper bound in Theorem 3.3.

Proposition 3.4. *Let $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ be a frequency vector satisfying $(*)$, and fix $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$. Then $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \leq \max\{\eta_{\mathcal{T}}, \beta_{\mathcal{T}}(\boldsymbol{\alpha})\}$.*

Throughout this section, we fix a frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ satisfying $(*)$ and an $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$. We first introduce some notation and then prove four auxiliary lemmas.

For each $m \in \mathbb{N}$, put $\mathcal{D}_m^c := \{(s,b) \in \mathcal{D} : b > m\}$, and set $\mathcal{B}_{s,m} := \{b \in \mathcal{B}_s : b \leq m\}$ and $\mathcal{B}_{s,m}^c := \{b \in \mathcal{B}_s : b > m\}$ for each $s \in \mathcal{S}$. For $b_1 \cdots b_n \in \mathcal{B}_\omega^n$ set

$$\tau_d(\omega, b_1 \cdots b_n) := \#\{1 \leq \ell \leq n : (\omega_\ell, b_\ell) = d\}.$$

For each $n, m \in \mathbb{N}$ and $\varepsilon > 0$, let $\mathcal{N}_n = \mathcal{N}_n(m,\varepsilon)$ be the set of vectors $(n_d)_{d\in\mathcal{D}_m} \in \mathbb{N}_0^{\#\mathcal{D}_m}$ that satisfy the following two properties:

- $\left|\frac{n_d}{n} - \alpha_d\right| < \varepsilon$ for all $d \in \mathcal{D}_m$;
- $\sum_{b\in\mathcal{B}_{s,m}} n_{(s,b)} \leq \tau_s(\omega, n)$ for all $s \in \mathcal{S}$.

Lemma 3.5. *Let $\varepsilon > 0$. Then there is an $N_\varepsilon \in \mathbb{N}$ such that $\mathcal{N}_n(m,\varepsilon) \neq \emptyset$ for all $n \geq N_\varepsilon$ and $m \in \mathbb{N}$.*

Proof. By $(*)$, there is an $N_1 \in \mathbb{N}$ such that, for each $n \geq N_1$ and $s \in \mathcal{S}$, we have

$$\tau_s(\omega, n) > n \sum_{b\in\mathcal{B}_s} \alpha_{(s,b)} - \frac{n\varepsilon}{2}.$$

Put $N_\varepsilon := \max\{N_1, \lceil\frac{2}{\varepsilon}\rceil\}$, and, for each $m \in \mathbb{N}$, $n \geq N_\varepsilon$ and $d \in \mathcal{D}_m$, set $n_d = \lfloor n\alpha_d - \frac{n\varepsilon}{2}\rfloor$. Then for $n \geq N_\varepsilon$ and $d \in \mathcal{D}_m$,

$$n\alpha_d - \frac{n\varepsilon}{2} - 1 < \left\lfloor n\alpha_d - \frac{n\varepsilon}{2} \right\rfloor < n\alpha_d$$

and hence

$$\alpha_d - \varepsilon < \alpha_d - \frac{\varepsilon}{2} - \frac{1}{n} < \frac{n_d}{n} < \alpha_d.$$

Moreover, for any $s \in \mathcal{S}$, since $\#\mathcal{B}_{s,m} \geq 1$,

$$\sum_{b\in\mathcal{B}_{s,m}} n_{(s,b)} \leq \sum_{b\in\mathcal{B}_{s,m}} \left(n\alpha_{(s,b)} - \frac{n\varepsilon}{2}\right) \leq n \sum_{b\in\mathcal{B}_{s,m}} \alpha_{(s,b)} - \frac{n\varepsilon}{2} < \tau_s(\omega, n).$$

Therefore, $(n_d)_{d\in\mathcal{D}_m} \in \mathcal{N}_n(m,\varepsilon)$. □

The next lemma concerns the sets

$$H_n = H_n(\boldsymbol{\alpha}, m, \varepsilon) := \left\{ x \in [0,1] : \left| \frac{\tau_d(\omega, x, n)}{n} - \alpha_d \right| < \varepsilon \, \forall d \in \mathcal{D}_m \right\},$$

which will be used to cover $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$. Note that $\mathcal{N}_n \neq \emptyset$ if $H_n \neq \emptyset$.

Lemma 3.6. *Fix $t \geq 0$, $\varepsilon > 0$ and $m, k_0 \in \mathbb{N}$. Then*

$$\mathcal{H}^t\left(\bigcap_{k \geq k_0} H_k\right) \leq \liminf_{n \to +\infty} \sum_{(n_d) \in \mathcal{N}_n} \left(\prod_{d \in \mathcal{D}_m} N_d^{-t n_d} \right) \sum_{\substack{b_1 \cdots b_n \in \mathcal{B}_\omega^n \\ \tau_d(\omega, b_1 \cdots b_n) = n_d}} \prod_{\substack{1 \leq \ell \leq n \\ b_\ell > m}} N_{\omega_\ell, b_\ell}^{-t}. \tag{3.9}$$

Proof. Fix $\delta > 0$ and let $n(\delta)$ be the smallest integer such that $|\langle b_1 \cdots b_{n(\delta)} \rangle_\omega|^t < \delta$ for all $b_1 \cdots b_{n(\delta)} \in \mathcal{B}_\omega^{n(\delta)}$. Then for any $n > \max\{k_0, n(\delta)\}$, we have that

$$\begin{aligned}
\mathcal{H}^t_\delta\left(\bigcap_{k \geq k_0} H_k\right) &= \inf\left\{ \sum_{U \in \mathcal{U}} |U|^t \; : \mathcal{U} \text{ is a } \delta\text{-cover of } \bigcap_{k \geq k_0} H_k \right\} \\
&\leq \inf_{k \geq n(\delta)} \left\{ \sum_{\substack{b_1 \cdots b_k \in \mathcal{B}_\omega^k \\ \langle b_1 \cdots b_k \rangle_\omega \cap H_n \neq \emptyset}} |\langle b_1 \cdots b_k \rangle_\omega|^t \right\} \leq \sum_{\substack{b_1 \cdots b_n \in \mathcal{B}_\omega^n \\ \langle b_1 \cdots b_n \rangle_\omega \cap H_n \neq \emptyset}} |\langle b_1 \cdots b_n \rangle_\omega|^t \\
&\leq \sum_{(n_d) \in \mathcal{N}_n} \sum_{\substack{b_1 \cdots b_n \in \mathcal{B}_\omega^n \\ \tau_d(\omega, b_1 \cdots b_n) = n_d, d \in \mathcal{D}_m}} \prod_{1 \leq \ell \leq n} N_{\omega_\ell, b_\ell}^{-t} \\
&= \sum_{(n_d) \in \mathcal{N}_n} \left(\prod_{d \in \mathcal{D}_m} N_d^{-t n_d} \right) \sum_{\substack{b_1 \cdots b_n \in \mathcal{B}_\omega^n \\ \tau_d(\omega, b_1 \cdots b_n) = n_d, d \in \mathcal{D}_m}} \prod_{\substack{1 \leq \ell \leq n \\ b_\ell > m}} N_{\omega_\ell, b_\ell}^{-t}.
\end{aligned}$$

Since this holds for all $n > \max\{k_0, n(\delta)\}$, we find that

$$\mathcal{H}^t_\delta\left(\bigcap_{k \geq k_0} H_k\right) \leq \liminf_{n \to +\infty} \sum_{(n_d) \in \mathcal{N}_n} \left(\prod_{d \in \mathcal{D}_m} N_d^{-t n_d} \right) \sum_{\substack{b_1 \cdots b_n \in \mathcal{B}_\omega^n \\ \tau_d(\omega, b_1 \cdots b_n) = n_d}} \prod_{\substack{1 \leq \ell \leq n \\ b_\ell > m}} N_{\omega_\ell, b_\ell}^{-t}.$$

The result follows by taking $\delta \to 0$. □

In the next lemma, we compute the inner sum of the right-hand side of (3.9). Given a vector $(n_d) \in \mathcal{N}_n$, for each $s \in \mathcal{S}$, we set $n_s := \sum_{b \in \mathcal{B}_{s,m}} n_{(s,b)}$.

Lemma 3.7. *Fix $t \geq 0$, $\varepsilon > 0$ and $m \in \mathbb{N}$. Let n be large enough so that $\mathcal{N}_n(m, \varepsilon) \neq \emptyset$, and let $(n_d) \in \mathcal{N}_n$. Then*

$$\begin{aligned}
&\sum_{\substack{b_1 \cdots b_n \in \mathcal{B}_\omega^n \\ \tau_d(\omega, b_1 \cdots b_n) = n_d, \, d \in \mathcal{D}_m}} \prod_{\substack{1 \leq \ell \leq n \\ b_\ell > m}} N_{\omega_\ell, b_\ell}^{-t} \\
&\qquad = \left(\prod_{s \in \mathcal{S}} \left(\sum_{b \in \mathcal{B}_{s,m}^c} N_{s,b}^{-t} \right)^{\tau_s(\omega, n) - n_s} \right) \frac{\prod_{s \in \mathcal{S}} \tau_s(\omega, n)!}{(\prod_{d \in \mathcal{D}_m} n_d!)(\prod_{s \in \mathcal{S}} (\tau_s(\omega, n) - n_s)!)}.
\end{aligned}$$

Proof. Set $\kappa := n - \sum_{d\in\mathcal{D}_m} n_d$. Any string $b_1\cdots b_n \in \mathcal{B}^n_\omega$ for which $\tau_d(\omega, b_1\cdots b_n) = n_d$ for all $d \in \mathcal{D}_m$ has exactly $\sum_{d\in\mathcal{D}_m} n_d$ terms that are at most m and κ terms that are larger than m. The value of the product

$$\prod_{\substack{1\le\ell\le n\\ b_\ell>m}} N^{-t}_{\omega_\ell,b_\ell}$$

depends only on the index and value of the κ entries in $b_1\cdots b_n$ that are larger than m.

Fix indices $1 \le v_1 < v_2 < \cdots < v_\kappa \le n$ where the digits $b > m$ can occur. This means that for each $1 \le \ell \le \kappa$, we have $\#\mathcal{B}_{\omega_{v_\ell}} > m$ and $n_{\omega_{v_\ell}} < \tau_{\omega_{v_\ell}}(\omega, n)$. In addition, fix values $c_1, \ldots, c_\kappa$ that these digits can take; so, $c_\ell \in \mathcal{B}^c_{\omega_{v_\ell},m}$ for each $1 \le \ell \le \kappa$. Then for each $b_1\cdots b_n \in \mathcal{B}^n_\omega$ that has $b_{v_\ell} = c_\ell$ for each $1 \le \ell \le \kappa$ and $\tau_d(\omega, b_1\cdots b_n) = n_d$ for each $d \in \mathcal{D}_m$, we have

$$\prod_{\substack{1\le\ell\le n\\ b_\ell>m}} N^{-t}_{\omega_\ell,b_\ell} = \prod_{1\le\ell\le\kappa} N^{-t}_{\omega_{v_\ell},c_\ell}.$$

The number of words $b_1\cdots b_n$ with these properties is given by

$$\prod_{s\in\mathcal{S}} \frac{n_s!}{\prod_{b\in\mathcal{B}_{s,m}} n_{(s,b)}!} = \frac{\prod_{s\in\mathcal{S}} n_s!}{\prod_{d\in\mathcal{D}_m} n_d!}.$$

Therefore, if we let $\mathcal{V}_n$ be the collection of sets $\{v_1, \ldots, v_\kappa\}$ of indices with $1 \le v_1 < \cdots < v_\kappa \le n$, $\#\mathcal{B}_{\omega_{v_\ell}} > m$ and $n_{\omega_{v_\ell}} < \tau_{\omega_{v_\ell}}(\omega, n)$ for each $1 \le \ell \le \kappa$, then

$$\sum_{\substack{b_1\cdots b_n\in\mathcal{B}^n_\omega\\ \tau_d(\omega,b_1\cdots b_n)=n_d, d\in\mathcal{D}_m}} \prod_{\substack{1\le\ell\le n\\ b_\ell>m}} N^{-t}_{\omega_\ell,b_\ell} = \frac{\prod_{s\in\mathcal{S}} n_s!}{\prod_{d\in\mathcal{D}_m} n_d!} \sum_{\{v_1,\ldots,v_\kappa\}\in\mathcal{V}_n} \sum_{\substack{c_1,\ldots,c_\kappa\\ c_\ell\in\mathcal{B}^c_{\omega_{v_\ell},m}}} \prod_{1\le\ell\le\kappa} N^{-t}_{\omega_{v_\ell},c_\ell}. \tag{3.10}$$

Next, we focus on the inner sum on the right-hand side of (3.10) and show that this is also independent of the choice of positions $\{v_1, \ldots, v_\kappa\} \in \mathcal{V}_n$; so, fix $\{v_1, \ldots, v_\kappa\} \in \mathcal{V}_n$. Observe that

$$\begin{aligned}\sum_{\substack{c_1,\ldots,c_\kappa\\ c_\ell\in\mathcal{B}^c_{\omega_{v_\ell},m}}} \prod_{1\le\ell\le\kappa} N^{-t}_{\omega_{v_\ell},b_{c_\ell}} &= \sum_{c_1\in\mathcal{B}^c_{\omega_{v_1},m}} N^{-t}_{\omega_{v_1},c_1}\left(\sum_{c_2\in\mathcal{B}^c_{\omega_{v_2},m}} N^{-t}_{\omega_{v_2},c_2}\left(\cdots\left(\sum_{c_\kappa\in\mathcal{B}^c_{\omega_{v_\kappa},m}} N^{-t}_{\omega_{v_\kappa},c_\kappa}\right)\cdots\right)\right)\\ &= \prod_{1\le\ell\le\kappa}\sum_{b\in\mathcal{B}^c_{\omega_{v_\ell},m}} N^{-t}_{\omega_{v_\ell},b},\end{aligned} \tag{3.11}$$

using independence of the κ sums in the second equality. By collecting like-terms, we get

$$\prod_{1\le\ell\le\kappa}\sum_{b\in\mathcal{B}^c_{\omega_{v_\ell},m}} N^{-t}_{\omega_{v_\ell},b} = \prod_{s\in\mathcal{S}}\prod_{\substack{1\le\ell\le\kappa\\ \omega_{v_\ell}=s}}\sum_{b\in\mathcal{B}^c_{s,m}} N^{-t}_{s,b} = \prod_{s\in\mathcal{S}}\left(\sum_{b\in\mathcal{B}^c_{s,m}} N^{-t}_{s,b}\right)^{\tau_s(\omega,n)-n_s}, \tag{3.12}$$

where the right-hand side is independent of $\{v_1, \ldots, v_\kappa\}$. Therefore, putting (3.10), (3.11) and (3.12) together yields

$$\sum_{\substack{b_1\cdots b_n\in\mathcal{B}^n_\omega\\ \tau_d(\omega,b_1\cdots b_n)=n_d, d\in\mathcal{D}_m}} \prod_{\substack{1\le\ell\le n\\ b_\ell>m}} N^{-t}_{\omega_\ell,b_\ell} = \left(\prod_{s\in\mathcal{S}}\left(\sum_{b\in\mathcal{B}^c_{s,m}} N^{-t}_{s,b}\right)^{\tau_s(\omega,n)-n_s}\right)\frac{\prod_{s\in\mathcal{S}} n_s!}{\prod_{d\in\mathcal{D}_m} n_d!}\#\mathcal{V}_n. \tag{3.13}$$

It only remains to compute $\#\mathcal{V}_n$. For each $s \in \mathcal{S}$, there are $\tau_s(\omega, n)$ indices $1 \leq \ell \leq n$ with $\omega_\ell = s$ out of which $\tau_s(\omega, n) - n_s$ have $(\omega_\ell, b_\ell) = (s, b)$ for some $b > m$. There are $\binom{\tau_s(\omega,n)}{n_s}$ possible arrangements of this kind and so

$$\#\mathcal{V}_n = \prod_{s\in\mathcal{S}} \binom{\tau_s(\omega, n)}{n_s} = \frac{\prod_{s\in\mathcal{S}} \tau_s(\omega, n)!}{\prod_{s\in\mathcal{S}} n_s! \prod_{s\in\mathcal{S}} (\tau_s(\omega, n) - n_s)!}.$$

Substituting this into (3.13) gives the result. □

In order to find recognisable quantities later, we shall find the exponential behaviour of some of the terms appearing in Lemmas 3.6 and 3.7. Set

$$f\big(t; m, \varepsilon, n, (n_d)\big) := \frac{1}{n} \log \left[\left(\prod_{d\in\mathcal{D}_m} N_d^{-tn_d} \right) \left(\prod_{s\in\mathcal{S}} \left(\sum_{b\in\mathcal{B}_{s,m}^c} N_{s,b}^{-t} \right)^{\tau_s(\omega,n)-n_s} \right) \frac{\prod_{s\in\mathcal{S}} \tau_s(\omega, n)!}{\prod_{d\in\mathcal{D}_m} n_d! \prod_{s\in\mathcal{S}} (\tau_s(\omega, n) - n_s)!} \right].$$

Then $f\big(t; m, \varepsilon, n, (n_d)\big)$ can be written as the sum of five terms. We consider each of these separately. First, set

$$A := t \sum_{d\in\mathcal{D}_m} \left(\alpha_d - \frac{n_d}{n} \right) \log N_d$$

so that

$$\frac{1}{n} \log \left(\prod_{d\in\mathcal{D}_m} N_d^{-tn_d} \right) = -t \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d + A.$$

Note for each $c > 0$ that, by taking $\varepsilon < c/(t \sum_{d\in\mathcal{D}_m} \log N_d)$ and N_ε to be the value given by Lemma 3.5, we have for any $n \geq N_\varepsilon$ and $(n_d)_{d\in\mathcal{D}_m} \in \mathcal{N}_n(m, \varepsilon)$ that $|A| < c$. Similarly, for the second part, set

$$B := \sum_{s\in\mathcal{S}} \left(\frac{\tau_s(\omega, n) - n_s}{n} - \sum_{b\in\mathcal{B}_{s,m}^c} \alpha_{(s,b)} \right) \log \sum_{b\in\mathcal{B}_{s,m}^c} N_{s,b}^{-t}$$

so that

$$\frac{1}{n} \log \prod_{s\in\mathcal{S}} \left(\sum_{b\in\mathcal{B}_{s,m}^c} N_{s,b}^{-t} \right)^{\tau_s(\omega,n)-n_s} = \sum_{s\in\mathcal{S}} \left(\sum_{b\in\mathcal{B}_{s,m}^c} \alpha_{(s,b)} \right) \log \sum_{b\in\mathcal{B}_{s,m}^c} N_{s,b}^{-t} + B.$$

Let $c > 0$. Take $\varepsilon < c/\sum_{s\in\mathcal{S}} \log \sum_{b\in\mathcal{B}_{s,m}^c} N_{s,b}^{-t}$. Then using $(*)$ and Lemma 3.5, we can find an $N \in \mathbb{N}$ such that, for each $n \geq N$ and $(n_d)_{d\in\mathcal{D}_m} \in \mathcal{N}_n(m, \varepsilon)$,

$$\left| \frac{\tau_s(\omega, n) - n_s}{n} - \sum_{b\in\mathcal{B}_{s,m}^c} \alpha_{(s,b)} \right| \leq \left| \frac{\tau_s(\omega, n)}{n} - \sum_{b\in\mathcal{B}_s} \alpha_{(s,b)} \right| + \sum_{b\in\mathcal{B}_{s,m}} \left| \alpha_{(s,b)} - \frac{n_{(s,b)}}{n} \right| < \varepsilon \tag{3.14}$$

and thus $|B| < c$. For the last three parts, we can write

$$\frac{1}{n} \log \frac{\prod_{s\in\mathcal{S}} \tau_s(\omega, n)!}{\prod_{d\in\mathcal{D}_m} n_d! \prod_{s\in\mathcal{S}} (\tau_s(\omega, n) - n_s)!} = \frac{1}{n} \sum_{s\in\mathcal{S}} \log \tau_s(\omega, n)! - \frac{1}{n} \sum_{d\in\mathcal{D}_m} \log n_d! - \frac{1}{n} \sum_{s\in\mathcal{S}} \log(\tau_s(\omega, n) - n_s)!.$$

3

Stirling's approximation $\log x! = x\log x - x + \mathcal{O}(\log x)$, $x \to +\infty$, and the fact that

$$\sum_{d\in\mathcal{D}_m} n_d + \sum_{s\in\mathcal{S}}(\tau_s(\omega,n) - n_s) = n = \sum_{s\in\mathcal{S}} \tau_s(\omega,n)$$

together yield

$$\begin{aligned}\frac{1}{n}\log\frac{\prod_{s\in\mathcal{S}}\tau_s(\omega,n)!}{\prod_{d\in\mathcal{D}_m} n_d!\prod_{s\in\mathcal{S}}(\tau_s(\omega,n)-n_s)!} &= \sum_{s\in\mathcal{S}}\frac{\tau_s(\omega,n)}{n}\log\tau_s(\omega,n) - \sum_{d\in\mathcal{D}_m}\frac{n_d}{n}\log n_d \\ &\quad - \sum_{s\in\mathcal{S}}\frac{\tau_s(\omega,n)-n_s}{n}\log(\tau_s(\omega,n)-n_s) + \mathcal{O}\Big(\frac{\log n}{n}\Big) \\ &= \sum_{s\in\mathcal{S}}\frac{\tau_s(\omega,n)}{n}\log\frac{\tau_s(\omega,n)}{n} - \sum_{d\in\mathcal{D}_m}\frac{n_d}{n}\log\frac{n_d}{n} \\ &\quad - \sum_{s\in\mathcal{S}}\frac{\tau_s(\omega,n)-n_s}{n}\log\frac{\tau_s(\omega,n)-n_s}{n} + \mathcal{O}\Big(\frac{\log n}{n}\Big).\end{aligned}$$

Set

$$C := \sum_{s\in\mathcal{S}}\frac{\tau_s(\omega,n)}{n}\log\frac{\tau_s(\omega,n)}{n} - \sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s,$$

and note that $|C| \to 0$ as $n \to +\infty$. Set

$$D := \sum_{d\in\mathcal{D}_m}\alpha_d\log\alpha_d - \sum_{d\in\mathcal{D}_m}\frac{n_d}{n}\log\frac{n_d}{n}.$$

By the uniform continuity of the map $x \mapsto x\log x$ on $[0,1]$ and by Lemma 3.5, for any $c > 0$, there is an $\varepsilon > 0$ and an $N \in \mathbb{N}$ such that, for all $n \geq N$ and all $(n_d)_{d\in\mathcal{D}_m} \in \mathcal{N}_n(m,\varepsilon)$, we have $|D| < c$. Finally, set

$$E := \sum_{s\in\mathcal{S}}\Bigg(\sum_{b\in\mathcal{B}^c_{s,m}}\alpha_{(s,b)}\Bigg)\log\Bigg(\sum_{b\in\mathcal{B}^c_{s,m}}\alpha_{(s,b)}\Bigg) - \sum_{s\in\mathcal{S}}\frac{\tau_s(\omega,n)-n_s}{n}\log\frac{\tau_s(\omega,n)-n_s}{n}.$$

Using (3.14) and the uniform continuity of the map $x \mapsto x\log x$ on $[0,1]$, we find again by Lemma 3.5 that, for any $c > 0$, there is an $\varepsilon > 0$ and an $N \in \mathbb{N}$ such that $|E| < c$ for all $n \geq N$ and all $(n_d)_{d\in\mathcal{D}_m} \in \mathcal{N}_n(m,\varepsilon)$. The above observations together lead to the following lemma.

Lemma 3.8. *Fix $t \geq 0$. Then for any $m \in \mathbb{N}$, $\varepsilon > 0$, $n \geq N_\varepsilon$ and $(n_d) \in \mathcal{N}_n(\alpha,m,\varepsilon)$,*

$$\begin{aligned} f_n\big(t;m,\varepsilon,n,(n_d)\big) &= -t\sum_{d\in\mathcal{D}_m}\alpha_d\log N_d + \sum_{s\in\mathcal{S}}\Bigg(\sum_{b\in\mathcal{B}^c_{s,m}}\alpha_{(s,b)}\Bigg)\log\sum_{b\in\mathcal{B}^c_{s,m}}N_{s,b}^{-t} \\ &\quad + \sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s - \sum_{d\in\mathcal{D}_m}\alpha_d\log\alpha_d - \sum_{s\in\mathcal{S}}\Bigg(\sum_{b\in\mathcal{B}^c_{s,m}}\alpha_{(s,b)}\Bigg)\log\Bigg(\sum_{b\in\mathcal{B}^c_{s,m}}\alpha_{(s,b)}\Bigg) \\ &\quad + A + B + C + D + E + \mathcal{O}\Big(\frac{\log n}{n}\Big).\end{aligned}$$

Moreover, for any $c > 0$ and $m \in \mathbb{N}$ there are $\varepsilon > 0$ and $\mathfrak{n} = \mathfrak{n}(m,\varepsilon) \in \mathbb{N}$ such that, for each $n \geq \mathfrak{n}$, the set $\mathcal{N}_n(m,\varepsilon)$ is non-empty and, for each $(n_d)_{d\in\mathcal{D}_m} \in \mathcal{N}_n(m,\varepsilon)$,

$$|A + B + C + D + E| < c.$$

We are now ready to prove the upper bound in Theorem 3.3.

Proof of Proposition 3.4. Put $t_0 := \max\{\eta_{\mathcal{T}}, \beta_{\mathcal{T}}(\boldsymbol{\alpha})\}$. To show that $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \leq t_0$, it suffices to show that $\mathcal{H}^{t_0}\big(F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = 0$. Observe that for any fixed $\varepsilon > 0$ and $m \in \mathbb{N}$,

$$F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset \bigcup_{k_0 \in \mathbb{N}} \bigcap_{k \geq k_0} H_k(\boldsymbol{\alpha}, m, \varepsilon).$$

We first show, using the lemmas above, that for any $k_0 \in \mathbb{N}$ and appropriate $m \in \mathbb{N}$ and $\varepsilon > 0$,

$$\mathcal{H}^{t_0}\left(\bigcap_{k \geq k_0} H_k(\boldsymbol{\alpha}, m, \varepsilon)\right) = 0. \tag{3.15}$$

So, fix $k_0 \in \mathbb{N}$. Let $\delta > 0$ and set $t(\delta) := t_0 + 5\delta$. Then for any $m \in \mathbb{N}$ and $\varepsilon > 0$, the Lemmas 3.6, 3.7 and 3.8 give

$$\mathcal{H}^{t(\delta)}\left(\bigcap_{k \geq k_0} H_k\right) \leq \liminf_{n \to +\infty} \sum_{(n_d) \in \mathcal{N}_n} \exp n f_n\big(t(\delta); m, \varepsilon, n, (n_d)\big). \tag{3.16}$$

To prove (3.15), it suffices to find $\varepsilon > 0$ and $m \in \mathbb{N}$ for which the right-hand side of (3.16) equals 0. We start by, for suitable n, bounding $f_n(t(\delta); m, \varepsilon, n, (n_d))$ above by a negative quantity independent of the choice of $(n_d) \in \mathcal{N}_n(m, \varepsilon)$. By definition of $\beta_{\mathcal{T}}(\boldsymbol{\alpha})$, there are infinitely many $m \in \mathbb{N}$ such that

$$\left| \frac{\sum_{s \in \mathcal{S}} \alpha_s \log \alpha_s - \sum_{d \in \mathcal{D}_m} \alpha_d \log \alpha_d}{\sum_{d \in \mathcal{D}_m} \alpha_d \log N_d} - \beta_{\mathcal{T}}(\boldsymbol{\alpha}) \right| < \delta.$$

For any such m,

$$\sum_{s \in \mathcal{S}} \alpha_s \log \alpha_s - \sum_{d \in \mathcal{D}_m} \alpha_d \log \alpha_d \leq (\beta_{\mathcal{T}}(\boldsymbol{\alpha}) + \delta) \sum_{d \in \mathcal{D}_m} \alpha_d \log N_d \leq (t_0 + \delta) \sum_{d \in \mathcal{D}_m} \alpha_d \log N_d.$$

Next, $\mathcal{B}^c_{s,m} \supseteq \mathcal{B}^c_{s,m+1}$ for each m and $\bigcap \mathcal{B}^c_{s,m} = \emptyset$, so for each $s \in \mathcal{S}$, $\sum_{b \in \mathcal{B}^c_{s,m}} \alpha_{(s,b)} \to 0$ as $m \to +\infty$. Since $-x \log x \to 0$ as $x \to 0$, and since $\mathcal{D}_m \subseteq \mathcal{D}_{m+1}$, we thus find for any large enough $m \in \mathbb{N}$ that

$$-\sum_{s \in \mathcal{S}} \left(\sum_{b \in \mathcal{B}^c_{s,m}} \alpha_{(s,b)} \right) \log \left(\sum_{b \in \mathcal{B}^c_{s,m}} \alpha_{(s,b)} \right) < \delta \sum_{d \in \mathcal{D}_m} \alpha_d \log N_d.$$

In addition, since $t(\delta) > \eta_{\mathcal{T}}$, it follows that for any $s \in \mathcal{S}$ and $m \in \mathbb{N}$,

$$\sum_{b \in \mathcal{B}^c_{s,m}} N_{s,b}^{-t(\delta)} < +\infty.$$

Hence, for any large enough $m \in \mathbb{N}$ we have

$$\sum_{s \in \mathcal{S}} \left(\sum_{b \in \mathcal{B}^c_{s,m}} \alpha_{(s,b)} \right) \log \sum_{b \in \mathcal{B}^c_{s,m}} N_{s,b}^{-t(\delta)} < \delta \sum_{d \in \mathcal{D}_m} \alpha_d \log N_d.$$

This means that there are infinitely many $m \in \mathbb{N}$, such that

$$
\begin{aligned}
&-t(\delta) \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d + \sum_{s\in\mathcal{S}} \left(\sum_{b\in\mathcal{B}^c_{s,m}} \alpha_{(s,b)} \right) \log \sum_{b\in\mathcal{B}^c_{s,m}} N_{s,b}^{-t(\delta)} + \sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s - \sum_{d\in\mathcal{D}_m} \alpha_d \log \alpha_d \\
&\quad - \sum_{s\in\mathcal{S}} \left(\sum_{b\in\mathcal{B}^c_{s,m}} \alpha_{(s,b)} \right) \log \left(\sum_{b\in\mathcal{B}^c_{s,m}} \alpha_{(s,b)} \right) \\
&\quad \le (-t(\delta) + t_0 + \delta + \delta + \delta) \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d = -2\delta \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d.
\end{aligned}
$$

Fix such an m. Then by Lemma 3.8 there is an $\varepsilon > 0$ and an $\mathfrak{n} = \mathfrak{n}(m, \varepsilon) \in \mathbb{N}$ such that for all $n \ge \mathfrak{n}$ the set $\mathcal{N}_n(m, \varepsilon)$ is non-empty and for all $(n_d) \in \mathcal{N}_n(m, \varepsilon)$,

$$
A + B + C + D + E < \delta \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d.
$$

Thus for all $n \ge \mathfrak{n}$,

$$
f_n\big(t(\delta); m, \varepsilon, n, (n_d)\big) \le -\delta \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d + \mathcal{O}\left(\frac{\log n}{n}\right),
$$

which is negative for large enough $n \in \mathbb{N}$. Substituting this into (3.16) and noting that

$$
\#\mathcal{N}_n(m, \varepsilon) \le \sum_{d\in\mathcal{D}_m} \big(n(\alpha_d + \varepsilon) - n(\alpha_d - \varepsilon)\big) = \#\mathcal{D}_m \cdot 2\varepsilon n
$$

for all $n \in \mathbb{N}$ yields

$$
\begin{aligned}
0 \le \mathcal{H}^{t(\delta)}\left(\bigcap_{k\ge k_0} H_k \right) &\le \liminf_{n\to+\infty} \#\mathcal{N}_n \cdot \exp\left(-\delta n \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d + \mathcal{O}(\log n) \right) \\
&\le \lim_{n\to+\infty} \#\mathcal{D}_m \cdot 2\varepsilon n \exp\left(-\delta n \sum_{d\in\mathcal{D}_m} \alpha_d \log N_d + \mathcal{O}(\log n) \right) = 0.
\end{aligned}
$$

Since this holds for all $k_0 \in \mathbb{N}$ we find that

$$
\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \le \dim_H \left(\bigcup_{k_0\in\mathbb{N}} \bigcap_{k\ge k_0} H_k \right) = \sup_{k_0\in\mathbb{N}} \dim_H \left(\bigcap_{k\ge k_0} H_k \right) \le t_0 + 5\delta.
$$

As this holds for any $\delta > 0$, we may take $\delta \to 0$ to obtain $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \le \max\{\eta_{\mathcal{T}}, \beta_{\mathcal{T}}(\boldsymbol{\alpha})\}$. □

3.5 Lower Bound: Exponent of Convergence

Let $\mathcal{T}$ be a collection of GLS maps indexed by a finite set $\mathcal{S}$ that satisfies (3.3), i.e. that the limit

$$
\lim_{n\to+\infty} \frac{\log n}{\log N_{s,n}}
$$

exists for all $s \in \mathcal{S}_{\mathbb{N}}$. In this section, we will prove that $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \ge \eta_{\mathcal{T}}$ for any frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ satisfying (∗) and any $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$. If $\eta_{\mathcal{T}} = 0$, then this holds trivially; therefore, we assume hereafter that $\eta_{\mathcal{T}} > 0$, which implies that $\mathcal{S}_{\mathbb{N}} \ne \emptyset$, i.e. that $\mathcal{T}$ contains at least one map T_s with $\mathcal{B}_s = \mathbb{N}$. We prove the following proposition.

Proposition 3.9. *Let $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ be a frequency vector satisfying* $(*)$*, and fix* $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$*. Assume that* $\eta_{\mathcal{T}} > 0$ *and that* (3.3) *holds. Then* $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \geq \eta_{\mathcal{T}}$*.*

Throughout this section, we fix a frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ satisfying $(*)$, an $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ and $\varsigma \in \mathcal{S}_{\mathbb{N}}$ such that

$$\eta_{\mathcal{T}} = \eta(T_\varsigma) > 0 \tag{3.17}$$

holds. To prove Proposition 3.9, we will need to prove several auxiliary lemmas.

Observe that $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ together with the non-redundancy condition $(*)$ on $\boldsymbol{\alpha}$ implies that $\tau_\varsigma(\omega) = \alpha_\varsigma > 0$ so there must be infinitely many $\ell \in \mathbb{N}$ for which $\omega_\ell = \varsigma$. Order these indices $j_1 < j_2 < \cdots < j_k < \cdots$. Fix $\gamma \in (1, 2)$, and define a (strictly increasing) function $\theta : \mathbb{N} \to \mathbb{N}$ by $\theta(k) := j_{\lceil k^\gamma \rceil}$. Recall that $\tau_\varsigma(\omega, k)$ denotes the number of occurrences of the symbol ς in the first block of length k of symbols in the sequence ω. Since $\tau_\varsigma(\omega) > 0$, we have

$$j_k = \frac{\tau_\varsigma(\omega, j_k)}{\tau_\varsigma(\omega)}\big(1 + o(1)\big) = \frac{k}{\tau_\varsigma(\omega)}\big(1 + o(1)\big), \qquad k \to +\infty.$$

Therefore, θ has the following properties:

(Θ1) $\omega_{\theta(k)} = \varsigma$ for all $k \in \mathbb{N}$.

(Θ2) $\theta(k) = \frac{k^\gamma}{\tau_\varsigma(\omega)}\big(1 + o(1)\big)$ as $k \to +\infty$ with $1 < \gamma < 2$.

During the proof of Proposition 3.9, we will obtain a lower bound for the Hausdorff dimension of $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ by altering the proof in [FLMW10, Theorem 1.2] (the part involving the exponent of convergence) to work in our setting. More precisely, the strategy will be to fix a $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ and consider a subset of points $x \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ such that the digit sequence $(b_n(\omega, x))_{n\in\mathbb{N}}$ of x is the same as the digit sequence $(b_n(\omega, z))_{n\in\mathbb{N}}$ of z except along the subsequence $(b_{\theta(k)}(\omega, x))_{k\in\mathbb{N}}$; see (3.21) below. The digits along this subsequence will instead be chosen from intervals whose size is governed by the sequence of numbers $(2^k)_{k\in\mathbb{N}}$. The freedom of choice of these digits will allow us to construct in Lemma 3.10 a measure μ_z on this subset that has desirable measure-theoretic entropy, which can then be used to deduce that $\eta_{\mathcal{T}}$ is a lower bound for $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ via an analogue of Billingsley's Lemma (see [Bil65], [BP17, Lemma 1.4.1]), namely Lemma 3.12 below.

Throughout this section, we shall fix $\varepsilon, \delta \in (0, 1)$. There is a $\kappa_1(\varepsilon, \delta) \in \mathbb{N}$ such that

$$2^k - 2^{\delta k} > 2^{(1-\varepsilon)k}, \qquad \forall k \geq \kappa_1(\varepsilon, \delta). \tag{3.18}$$

Additionally, since we have ordered for each $s \in \mathcal{S}$ the (unsigned) slopes $N_{s,b}$, $b \in \mathcal{B}_s$, of T_s so that $N_{s,b}$ increases as b increases, we have by [PS72, p26] and (3.3) that

$$\eta(T_s) = \limsup_{n\to+\infty} \frac{\log n}{\log N_{s,n}} = \lim_{n\to+\infty} \frac{\log n}{\log N_{s,n}}. \tag{3.19}$$

Thus, since $\mathcal{S}_{\mathbb{N}}$ is a finite set, there is a $\kappa_2(\varepsilon, \delta) \in \mathbb{N}$ such that

$$N_{s,n}^{\eta(T_s)} \leq n^{1+\varepsilon}, \qquad \forall s \in \mathcal{S}_{\mathbb{N}},\ n \geq \kappa_2(\varepsilon, \delta). \tag{3.20}$$

Fix $\kappa \geq \max\{\kappa_1(\varepsilon, \delta), \kappa_2(\varepsilon, \delta)\}$. Next, take $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$, which exists since $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$, and define the set

$$F_z := \big\{x \in X_\omega : b_{\theta(k)}(\omega, x) \in \mathbb{N} \cap (2^k - 2^{\delta k}, 2^k]\ \forall k \geq \kappa,\ b_\ell(\omega, x) = b_\ell(\omega, z) \text{ otherwise}\big\}. \tag{3.21}$$

Note that $F_z \neq \emptyset$, since by (Θ1), $\mathbb{N} \cap (2^k - 2^{\delta k}, 2^k] \subset \mathcal{B}_{\theta(k)}$ for all $k \in \mathbb{N}$. Moreover, we have $\theta(k)/k \to +\infty$ by (Θ2) so the digit frequencies of $x \in F_z$ are the same as the digit frequencies of z. Thus, it follows that F_z is a subset of $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$.

To obtain the desired lower bound in the proof of Proposition 3.9, we will apply a non-autonomous analogue (Lemma 3.12 below) of Billingsley's Lemma to F_z and a measure μ_z supported on F_z. We will deduce the existence of μ_z via the Ionescu-Tulcea theorem. For each $n \in \mathbb{N}$, define $\kappa(n) := \min\{k \geq \kappa : \theta(k) > n\}$ and note that $\kappa(n) > \kappa$ if and only if $n \geq \theta(\kappa)$.

Lemma 3.10. *For any $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ there is a probability measure μ_z on $\big([0,1], \mathcal{B}([0,1])\big)$ that has $\mu_z(F_z) = 1$ and satisfies*

$$\mu_z\big(\langle b_1 \cdots b_n \rangle_\omega\big) = \begin{cases} \prod_{\kappa \leq k < \kappa(n)} \lceil 2^{\delta k} \rceil^{-1}, & \text{if } \langle b_1 \cdots b_n \rangle_\omega \cap F_z \neq \emptyset, \\ 0, & \text{otherwise}, \end{cases} \tag{3.22}$$

for each $b_1 \cdots b_n \in \mathcal{B}_\omega^n$, $n \in \mathbb{N}$, where $\prod_{\kappa \leq k < \kappa(n)} \lceil 2^{\delta k} \rceil^{-1} := 1$ if there is no $\kappa \leq k < \kappa(n)$.

Proof. Let $\widetilde{\mu}_1 = \delta_{b_1(\omega,z)}$ be the Dirac measure at $b_1(\omega, z)$ on the measurable space $(\mathcal{B}_{\omega_1}, \mathcal{P}(\mathcal{B}_{\omega_1}))$. For each $n \in \mathbb{N} \setminus \{1\}$, define the stochastic kernel $K_n : (\mathcal{B}_\omega^{n-1}, \mathcal{P}(\mathcal{B}_\omega^{n-1})) \to (\mathcal{B}_{\omega_n}, \mathcal{P}(\mathcal{B}_{\omega_n}))$ by setting for each $b_1 \cdots b_{n-1} \in \mathcal{B}_\omega^{n-1}$ and each $A \in \mathcal{P}(\mathcal{B}_{\omega_n})$

$$K_n(b_1 \cdots b_{n-1}, A) := \begin{cases} \delta_{b_n(\omega,z)}, & \text{if } n \neq \theta(k),\ k \geq \kappa, \\ \lceil 2^{k\delta} \rceil^{-1} \sum_{j \in (2^k - 2^{\delta k}, 2^k] \cap \mathbb{N}} \delta_j, & \text{if } n = \theta(k),\ k \geq \kappa. \end{cases}$$

For each $n \in \mathbb{N}$, let $\widetilde{\mu}_n$ be the probability measure $\widetilde{\mu}_1 \otimes \bigotimes_{k=2}^n K_k$ on $(\mathcal{B}_\omega^n, \mathcal{P}(\mathcal{B}_\omega^n))$. Also, let $\mathcal{C}$ denote the σ-algebra on $\mathcal{B}_\omega^{\mathbb{N}}$ generated by the cylinder sets

$$[b_1 \cdots b_n]_\omega := \{\xi = (\xi_i)_{i \in \mathbb{N}} \in \mathcal{B}_\omega^{\mathbb{N}} : \xi_i = b_i,\ 1 \leq i \leq n\}.$$

By the Ionescu-Tulcea Theorem (see e.g. [Kle20, Theorem 14.35]), there exists a unique probability measure $\widetilde{\mu}_z$ on $(\mathcal{B}_\omega^{\mathbb{N}}, \mathcal{C})$ with the property that, for each $A \in \mathcal{P}(\mathcal{B}_\omega^n)$,

$$\widetilde{\mu}_z(A \times \mathcal{B}_{\omega_{n+1}} \times \mathcal{B}_{\omega_{n+2}} \times \cdots) = \widetilde{\mu}_n(A). \tag{3.23}$$

Recall the definition of π_ω from (3.6), and let μ_z be the probability measure on $([0,1], \mathcal{B}([0,1]))$ given by $\mu_z = \widetilde{\mu}_z \circ \pi_\omega^{-1}$. Fix $n \in \mathbb{N}$ and $b_1 \cdots b_n \in \mathcal{B}_\omega^n$. Then $[b_1 \cdots b_n]_\omega = \pi_\omega^{-1}\big(\overline{\langle b_1 \cdots b_n \rangle_\omega}\big)$. The inverse image $\pi_\omega^{-1}\big(\overline{\langle b_1 \cdots b_n \rangle_\omega} \setminus \langle b_1 \cdots b_n \rangle_\omega\big)$ consists of at most four sequences in $\mathcal{B}_\omega^{\mathbb{N}}$ corresponding to at most the two endpoints of $\langle b_1 \cdots b_n \rangle_\omega$. Hence, we find by (3.23) that

$$\mu_z\big(\langle b_1 \cdots b_n \rangle_\omega\big) = \mu_z\big(\overline{\langle b_1 \cdots b_n \rangle_\omega}\big) = \widetilde{\mu}_z([b_1 \cdots b_n]_\omega) = \widetilde{\mu}_n(\{b_1 \cdots b_n\}).$$

Then (3.22) follows from the definition of $\widetilde{\mu}_n$.

Finally, we verify that $\mu_z(F_z) = 1$. Put

$$\mathcal{E}_n = \{a_1 \cdots a_n \in \mathcal{B}_\omega^n : \widetilde{\mu}_n(\{a_1 \cdots a_n\}) > 0\}$$

and note that

$$F_z = \bigcap_{n \in \mathbb{N}} \bigcup_{a_1 \cdots a_n \in \mathcal{E}_n} \langle a_1 \cdots a_n \rangle_\omega \cap X_\omega.$$

Since for each $m \in \mathbb{N}$

$$\bigcap_{n=1}^{m} \bigcup_{a_1\cdots a_n \in \mathcal{E}_n} \langle a_1 \cdots a_n \rangle_\omega = \bigcup_{a_1\cdots a_m \in \mathcal{E}_m} \langle a_1 \cdots a_m \rangle_\omega \supseteq \bigcup_{a_1\cdots a_{m+1} \in \mathcal{E}_{m+1}} \langle a_1 \cdots a_{m+1} \rangle_\omega,$$

and $[0,1] \setminus X_\omega$ consists of at most countably many endpoints of FFIs, it holds that

$$\begin{aligned}\mu_z(F_z) &= \lim_{n\to\infty} \mu_z\left(\bigcup_{a_1\cdots a_n \in \mathcal{E}_n} \langle a_1 \cdots a_n \rangle_\omega \right) = \lim_{n\to\infty} \sum_{a_1\cdots a_n \in \mathcal{E}_n} \mu_z(\langle a_1 \cdots a_n \rangle_\omega) \\ &= \lim_{n\to\infty} \sum_{a_1\cdots a_n \in \mathcal{E}_n} \widetilde{\mu}_z \circ \pi_\omega^{-1}\left(\overline{\langle a_1 \cdots a_n \rangle_\omega} \right) \\ &= \lim_{n\to+\infty} \sum_{a_1\cdots a_n \in \mathcal{B}_\omega^n} \widetilde{\mu}_n(\{a_1 \cdots a_n\}) = \lim_{n\to+\infty} 1 = 1.\end{aligned}$$

This gives the result. □

Next, we will construct a point $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ such that the first block of n digits of z corresponds to branches of maps of $\mathcal{T}$ with sufficiently small slopes. Write $\eta := \min\{\eta(T_s) : s \in \mathcal{S}_\mathbb{N}\} \in (0,1]$.

Lemma 3.11. *There is a $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ for which there exists a constant $N_z \in \mathbb{N}$ (depending on ε and δ) such that $N_{\omega_n, b_n(\omega,z)} \le n^{(1+\varepsilon)/\eta}$ for all $n \ge N_z$. Furthermore, as $n \to +\infty$,*

$$\sum_{1\le \ell \le n} \log N_{\omega_\ell, b_\ell(\omega,z)} = o(\kappa(n)^2).$$

Proof. [FLMW10, Lemma 2.1] yields for each $s \in \mathcal{S}$ a sequence $(b_n^{(s)})_{n\in\mathbb{N}}$ on $\mathcal{B}_s \subseteq \mathbb{N}$ such that

$$\lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : b_\ell^{(s)} = b\}}{n} = \frac{\alpha_{(s,b)}}{\alpha_s}, \qquad \forall\, b \in \mathcal{B}_s,$$

and $b^{(s)}_{\tau_s(\omega,n)} \le n$ for all $n \in \mathbb{N}$. By weaving the sequences $(b_n^{(s)})_{n\in\mathbb{N}}$, $s \in \mathcal{S}$, together according to ω as we did in the proof of Theorem 3.2, we can obtain an ω-compatible sequence $(b_n)_{n\in\mathbb{N}}$ on $\mathbb{N}$ such that

$$\lim_{n\to+\infty} \frac{\#\{1 \le \ell \le n : (\omega_\ell, b_\ell) = (s,b)\}}{n} = \frac{\alpha_{(s,b)}}{\alpha_s}, \qquad \forall\, (s,b) \in \mathcal{D},$$

and $b_n := b^{\omega_n}_{\tau_{\omega_n}(\omega,n)} \le n$ for all $n \in \mathbb{N}$. Put $d_n := (\omega_n, b_n)$ for each $n \in \mathbb{N}$, and, if necessary, alter d_n, $n > 1$, as we did at the end of the proof of Theorem 3.2 so that $\pi((d_n)_{n\in\mathbb{N}}) \in X_\omega$ gives a valid $(\mathcal{T},\omega)$-expansion. Note that altering d_n, $n > 1$, in this way will change some of the b_n to 1 or to 2 and thus preserves the inequality $b_n \le n$. If we set $z := \pi((d_n)_{n\in\mathbb{N}}) \in X_\omega$, then $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ and is such that $b_n(\omega,z) \le n$ for all $n \in \mathbb{N}$.

Now, since $\#\mathcal{B}_s < +\infty$ for each $s \in \mathcal{S} \setminus \mathcal{S}_\mathbb{N}$, we may define $M := \max\{N_{s,b} : b \in \mathcal{B}_s, s \in \mathcal{S} \setminus \mathcal{S}_\mathbb{N}\}$. Fix $n \ge N_z := \max\{M, \kappa\} \in \mathbb{N}$. If $\omega_n \in \mathcal{S}_\mathbb{N}$, then $\log N_{\omega_n, b_n(\omega,z)} \le \log N_{\omega_n, n} \le \frac{1+\varepsilon}{\eta(T_{\omega_n})} \log n$ by construction of z and (3.20), using that $n \ge \kappa$. If $\omega_n \in \mathcal{S} \setminus \mathcal{S}_\mathbb{N}$ instead, then we have that $\log N_{\omega_n, b_n(\omega,z)} \le \log M \le \log n$ by definition of M and N_z. The proof of the first statement then follows.

To prove the second statement, we may apply the first statement to obtain

$$\begin{aligned}\sum_{1\le\ell\le n} \log N_{\omega_\ell, b_\ell(\omega,z)} &\le \sum_{1\le\ell\le N_z} \log N_{\omega_\ell, b_\ell(\omega,z)} + c_z \sum_{N_z\le\ell\le n} \log \ell \le C + c_z(n - N_z)\log n \\ &\le (C + c_z) n \log n,\end{aligned} \tag{3.24}$$

for any $n \geq N_z$, where we have set $c_z := (1+\varepsilon)/\eta$ and $C := \sum_{1\leq \ell \leq N_z} \log N_{\omega_\ell, b_\ell(\omega,z)}$, which are both independent of n. By (Θ2), we have (for $1 < \gamma < 2$) that, for any $\varepsilon_0 > 0$, there is a $K = K(\varepsilon_0) \in \mathbb{N}$ such that $|\tau_\varsigma(\omega)\theta(k)/k^\gamma - 1| < \varepsilon_0$ for all $k \geq K$, which implies that

$$\theta(k) < \frac{k^\gamma}{\tau_\varsigma(\omega)}(1+\varepsilon_0).$$

Recall for each $n \in \mathbb{N}$ that $\kappa(n) = \min\{k \geq \kappa : \theta(k) > n\}$ so $\theta(k) > n$ for any $k \geq \kappa(n)$. Let N be such that $\kappa(N) > K$. Then for all $n \geq N$ and $k \geq \kappa(n)$, $k^\gamma > \frac{\tau_\varsigma(\omega)}{1+\varepsilon_0} n$. Thus,

$$\kappa(n)^2 > \left(\frac{\tau_\varsigma(\omega)}{1+\varepsilon_0} n\right)^{2/\gamma}, \qquad \forall n \geq N. \tag{3.25}$$

Since $2/\gamma > 1$, we have $n \log n = o(n^{2/\gamma})$ and so it follows from (3.25) that $n \log n = o(\kappa(n)^2)$ as $n \to +\infty$. Substituting this into (3.24) completes the proof. □

3

Fix z as in Lemma 3.11. We shall use [Caj81, p36] (see also [Weg68, Satz 2]) to compute a lower bound for the Hausdorff dimension of F_z using δ_0-covers of FFIs as $\delta_0 \to 0$. Dimension defined using only covers of fundamental intervals (or, more generally, cylinder sets) is known as the Billingsley dimension; see [Bil65]. In our setting, we shall define the *Billingsley fibre dimension* $\dim_\omega A$ of a subset $A \subset [0,1]$ to be the value $t_0 \geq 0$ such that $\mathcal{H}^t_\omega(A) = 0$ for all $t > t_0$ and $\mathcal{H}^t_\omega(A) = +\infty$ for all $t < t_0$, where $\mathcal{H}^t_\omega$ is the *Billingsley fibre t-measure* given by

$$\mathcal{H}^t_\omega(A) := \lim_{\delta_0 \to 0} \inf \left\{ \sum_{i \in \mathcal{I}} |A_i|^t : \{A_i\}_{i\in\mathcal{I}} \text{ is a } \delta_0\text{-cover of } A \text{ with FFIs} \right\}.$$

The limit in δ_0 above exists and defines a measure by e.g. [Caj81, §2].

Lemma 3.12. *Fix $\omega \in \Omega_\mathcal{T}(\boldsymbol{\alpha})$. Let $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ be as in Lemma 3.11, and let μ_z be the measure constructed in Lemma 3.10 for z. If there is a $t_0 \geq 0$ for which*

$$\liminf_{n\to+\infty} \frac{\log \mu_z(\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega)}{\log|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega|} \geq t_0, \quad \forall x \in F_z, \tag{3.26}$$

then $\dim_H F_z \geq t_0$.

Proof. The method of proof is to first show that (3.26) implies $\dim_\omega F_z \geq t_0$ and then to show that $\dim_H F_z = \dim_\omega F_z$. The first part of the method can be proved analogously to Billingsley's Lemma for autonomous systems (see e.g. [BP17, Lemma 1.4.1]) by using the FFIs $\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega$ in place of the usual fundamental intervals. For completeness, we reproduce the proof here in our setting.

Fix $0 \leq t < t_0$. Given (3.26), we have for all $x \in F_z$ that

$$\limsup_{n\to+\infty} \frac{\mu_z\big(\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega\big)}{\big|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega\big|^t} \leq 1.$$

For a fixed constant $C > 1$ and $k \in \mathbb{N}$, put

$$F_{z,k} := \Big\{ x \in F_z : \mu_z\big(\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega\big) < C\big|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega\big|^t \ \forall n > k \Big\}.$$

Note that $F_z = \bigcup_{k\in\mathbb{N}} F_{z,k}$ and $F_{z,k} \subset F_{z,k+1}$ for all $k \in \mathbb{N}$ so $\mu_z(F_{z,k}) \to \mu_z(F_z) = 1$ as $k \to +\infty$. Take $\delta_0 \in (0,1)$ and consider any δ_0-cover $\{\mathcal{C}_j\}_{j\in\mathcal{J}}$ of F_z by FFIs. Then $\{\mathcal{C}_j\}_{j\in\mathcal{J}}$ is a cover of $F_{z,k}$ for every k and so we find that

$$\sum_{j\in\mathcal{J}} |\mathcal{C}_j|^t \geq \sum_{\substack{j\in\mathcal{J} \\ \mathcal{C}_j\cap F_{z,k}\neq\emptyset}} |\mathcal{C}_j|^t \geq C^{-1} \sum_{\substack{j\in\mathcal{J} \\ \mathcal{C}_j\cap F_{z,k}\neq\emptyset}} \mu_z(\mathcal{C}_j) \geq C^{-1}\mu_z(F_{z,k})$$

for all $k \in \mathbb{N}$. Thus, we have $\mathcal{H}^t_\omega(F_z) \geq C^{-1}\mu_z(F_{z,k})$ for all $k \in \mathbb{N}$ and taking $k \to +\infty$ gives $\mathcal{H}^t_\omega(F_z) \geq C^{-1} > 0$ for any $0 \leq t < t_0$. Therefore, $\dim_\omega F_z \geq t_0$.

We now show that $\dim_H F_z = \dim_\omega F_z$. Since the length of FFIs decreases to 0 as their level increases, it suffices by [Caj81, p36] to show for each $x \in F_z$ that

$$\lim_{n\to+\infty} \frac{\log|\langle b_1(\omega,x)\cdots b_{n+1}(\omega,x)\rangle_\omega|}{\log|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega|} = 1. \tag{3.27}$$

To prove this, first recall from (3.25) that there is an $N \in \mathbb{N}$ and a constant $c > 0$ such that $\kappa(n) > cn^{1/\gamma}$ for all $n \geq N$, where $\gamma \in (1,2)$, so we thus have for some $N' \geq N$ that $n < 2^{\kappa(n)}$ for all $n \geq N'$. Fix $n \geq \max\{N', N_z, \theta(\kappa)\}$ so that (3.18), (3.20) and $n < 2^{\kappa(n)}$ all hold, and fix $x \in F_z$. The chosen ordering on the digits of T_ς implies that $N_{\varsigma,1} \leq \cdots \leq N_{\varsigma,b}$ for all $b \in \mathbb{N}$. Since $\sum_{b\in\mathcal{B}_\varsigma} N_{\varsigma,b}^{-1} = 1$, we must have $N_{\varsigma,b} \geq b$ for all $b \in \mathbb{N}$, else $\sum_{1\leq\ell\leq b} N_{\varsigma,\ell}^{-1} > b\cdot b^{-1} = 1$. Then (3.18) implies for any $b \in (2^k - 2^{\delta k}, 2^k]$, $k \geq \kappa$, that

$$\log N_{\varsigma,b} \geq \log b > \log(2^k - 2^{\delta k}) > (1-\varepsilon)k\log 2$$

and so it follows from (Θ1) and the fact that $x \in F_z$ that

$$\begin{aligned} 0 \leq \frac{\log N_{\omega_{n+1},b_{n+1}(\omega,x)}}{\sum_{1\leq\ell\leq n}\log N_{\omega_\ell,b_\ell(\omega,x)}} &\leq \frac{\log N_{\omega_{n+1},b_{n+1}(\omega,x)}}{\sum_{\kappa\leq k<\kappa(n)}\log N_{\varsigma,b_{\theta(k)}(\omega,x)}} \\ &\leq \frac{1}{(1-\varepsilon)\log 2}\frac{\log N_{\omega_{n+1},b_{n+1}(\omega,x)}}{\sum_{\kappa\leq k<\kappa(n)} k} \\ &= \frac{2}{(1-\varepsilon)\log 2}\frac{\log N_{\omega_{n+1},b_{n+1}(\omega,x)}}{\kappa(n)(\kappa(n)-1)-\kappa(\kappa-1)}. \end{aligned}$$

We show that the quantity on the right-hand side of the above equation vanishes as $n \to +\infty$, which we do by splitting into two cases. First, suppose that $n+1 = \theta(\kappa(n))$. Then we have $b_{n+1}(\omega,x) \in \left(2^{\kappa(n)} - 2^{\delta\kappa(n)}, 2^{\kappa(n)}\right]$ since $x \in F_z$ so (Θ1) and (3.20) yield

$$\log N_{\omega_{n+1},b_{n+1}(\omega,x)} = \log N_{\varsigma,b_{n+1}(\omega,x)} \leq \frac{1+\varepsilon}{\eta_{\mathcal{T}}}\log b_{n+1}(\omega,x) \leq \frac{1+\varepsilon}{\eta_{\mathcal{T}}}\kappa(n)\log 2. \tag{3.28}$$

Suppose instead that $n+1 \neq \theta(\kappa(n))$. Then $n+1 \neq \theta(k)$ for all $k \geq \kappa$ so it follows from the fact that $x \in F_z$ and Lemma 3.11 that

$$\log N_{\omega_{n+1},b_{n+1}(\omega,x)} = \log N_{\omega_{n+1},b_{n+1}(\omega,z)} \leq \frac{1+\varepsilon}{\eta}\log(n+1) \leq \frac{1+\varepsilon}{\eta}\kappa(n)\log 2,$$

using that $n \geq N'$. Since $\eta_{\mathcal{T}} \geq \eta$, we have that

$$0 \leq \lim_{n\to+\infty}\frac{\log N_{\omega_{n+1},b_{n+1}(\omega,x)}}{\kappa(n)(\kappa(n)-1)-\kappa(\kappa-1)} \leq \frac{1+\varepsilon}{\eta}\lim_{n\to+\infty}\frac{\kappa(n)\log 2}{\kappa(n)(\kappa(n)-1)-\kappa(\kappa-1)} = 0.$$

Consequently,

$$\lim_{n\to+\infty} \frac{\log|\langle b_1(\omega,x)\cdots b_{n+1}(\omega,x)\rangle_\omega|}{\log|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega|} = \lim_{n\to+\infty} \frac{\sum_{1\le\ell\le n+1}\log N_{\omega_\ell,b_\ell(\omega,x)}}{\sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,x)}}$$
$$= 1 + \lim_{n\to+\infty} \frac{\log N_{\omega_{n+1},b_{n+1}(\omega,x)}}{\sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,x)}} = 1$$

and so (3.27) holds. Therefore, $\dim_H F_z = \dim_\omega F_z \ge t_0$. □

We now have everything that we need to prove Proposition 3.9.

Proof of Proposition 3.9. Recall that $\varsigma \in \mathcal{S}$ satisfies (3.17). Fix $\varepsilon, \delta \in (0,1)$, and take $\kappa \in \mathbb{N}$ (depending on ε, δ) so that (3.18) and (3.20) hold. Fix $z \in F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ as in Lemma 3.11. Let $F_z = F_z(\varepsilon,\delta)$ be the subset of $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ defined in (3.21), and let μ_z be the measure constructed in Lemma 3.10 that is supported on F_z. To obtain a lower bound for $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$, by Lemma 3.12, it suffices to show that

$$\liminf_{n\to+\infty} \frac{\log \mu_z(\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega)}{\log|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega|} \ge \frac{\delta}{1+\varepsilon}\eta_{\mathcal{T}}, \quad \forall x \in F_z, \tag{3.29}$$

since then $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \ge \dim_H F_z \ge \frac{\delta}{1+\varepsilon}\eta_{\mathcal{T}}$ and we can take $\varepsilon \to 0$ and $\delta \to 1$ to get the desired lower bound. Take $x \in F_z$ and observe that

$$-\log|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega| = \sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,x)}$$
$$\le \sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,z)} + \sum_{\kappa\le k<\kappa(n)}\log N_{\varsigma,b_{\theta(k)}(\omega,x)}.$$

By the same reasoning as in the derivation of (3.28), we have

$$\log N_{\varsigma,b_{\theta(k)}(\omega,x)} \le \frac{1+\varepsilon}{\eta_{\mathcal{T}}} k\log 2, \qquad \forall k \ge \kappa,$$

so, for any $x \in F_z$,

$$\frac{\log \mu_z(\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega)}{\log|\langle b_1(\omega,x)\cdots b_n(\omega,x)\rangle_\omega|} \ge \frac{\sum_{\kappa\le k<\kappa(n)}\delta k\log 2}{\sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,z)} + \frac{1+\varepsilon}{\eta_{\mathcal{T}}}\sum_{\kappa\le k<\kappa(n)} k\log 2} = \frac{c_n\delta}{1+\varepsilon}\eta_{\mathcal{T}},$$

where

$$c_n := \frac{\sum_{\kappa\le k<\kappa(n)} k\log 2}{\frac{\eta_{\mathcal{T}}}{1+\varepsilon}\sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,z)} + \sum_{\kappa\le k<\kappa(n)} k\log 2}.$$

By Lemma 3.11, it follows that

$$\lim_{n\to+\infty}\frac{\sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,z)}}{\sum_{\kappa\le k<\kappa(n)} k\log 2} = \frac{2}{\log 2}\lim_{n\to+\infty}\frac{\sum_{1\le\ell\le n}\log N_{\omega_\ell,b_\ell(\omega,z)}}{\kappa(n)(\kappa(n)-1)-\kappa(\kappa-1)} = 0,$$

and thus $c_n \to 1$ as $n \to +\infty$, which gives (3.29). Therefore,

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \ge \frac{\delta}{1+\varepsilon}\eta_{\mathcal{T}}$$

by Lemma 3.12 and we may take $\varepsilon \to 0$ and $\delta \to 1$ to conclude the proof. □

3.6 Lower Bound: Fibre Dimension

Fix an NGLS map $(\mathcal{T},\omega)$. This section will be devoted to proving that $\beta_{\mathcal{T}}(\boldsymbol{\alpha})$ is a lower bound for $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$. In the autonomous setting of [FLMW10], this is achieved by applying the Birkhoff Ergodic Theorem, but this is not an option for us in our non-autonomous setting. Consequently, we must find an alternate method of proof for this bound. We do this by constructing in §3.6.1 a sequence $((\mathcal{T}^{(m)},\omega))_{m\in\mathbb{N}}$ of NGLS maps from $\mathcal{T}$, where $\mathcal{T}^{(m)} := \{T_s^{(m)}\}_{s\in\mathcal{S}}$ consists of unit interval maps with only finitely many branches; the first m branches of $T_s^{(m)}$, ordered by decreasing branch-support length, are taken to agree with the first m branches of T_s. The purpose of these maps will be to provide a finite-branched approximation of the maps of $\mathcal{T}$, becoming more accurate as $m\to+\infty$. We will then show in §3.6.2 that the lower bound $\beta_{\mathcal{T}}(\boldsymbol{\alpha})$ of $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ follows by relating certain dimension quantities associated with $(\mathcal{T},\omega)$ and $(\mathcal{T}^{(m)},\omega)$ as $m\to+\infty$, allowing us to apply methods that work on finite alphabets, e.g. from Theorem 2.2 and [BI09, Theorem 3.1]. More precisely, we will relate the pointwise dimensions of a measure $\mu_{\omega,\boldsymbol{\alpha}}$ supported on $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ and its counterpart measure $\mu_{\omega,\boldsymbol{\alpha}}^{(m)}$ associated with $(\mathcal{T}^{(m)},\omega)$ – see Lemma 3.14 and Corollary 3.15 below for the construction of the measures and (3.38) for the relation between their pointwise dimensions.

3.6.1 Approximation with Finite-branched NGLS Maps

In this section, we construct the NGLS maps $(\mathcal{T}^{(m)},\omega)$, $m\in\mathbb{N}$, and establish some of their main properties needed in the proof of Theorem 3.3.

Fix $m\in\mathbb{N}$ and $s\in\mathcal{S}$. Note that the set $[0,1]\setminus\bigcup_{b\in\mathcal{B}_s,\,b\le m} I_{s,b}$ is the union of at most finitely many disjoint (non-degenerate) closed intervals and at most finitely many singletons and can possibly be empty. We will construct $T_s^{(m)}$ from $T_s\in\mathcal{T}$ by keeping the first m branches of T_s, that is, those branches supported on intervals $I_{s,b}$ with $b\le m$, and then putting affine and positively oriented branches on the remainder of the interval. Instead of labelling the intervals $I_{s,b}^{(m)}$ of the GLS maps $T_s^{(m)}$ by consecutive integers, as we did for T_s, we now do the following. For $b\in\mathcal{B}_s\cap\{1,2,\dots,m\}$, we set $I_{s,b}^{(m)} := I_{s,b}$, i.e. we use the same labels as for T_s. In addition, we label each non-degenerate open interval I of maximal length in $[0,1]\setminus\bigcup_{b\in\mathcal{B}_s,\,b\le m} I_{s,b}$ by $I_{s,b}^{(m)}$, where $b=\min\{c\in\mathbb{N} : I_{s,c}\subseteq I\}>m$. Let $\mathcal{B}_s^{(m)}$ denote the corresponding branch-indexing set for $T_s^{(m)}$. Then $B_s^{(m)}=\{1,2,\dots,m+1\}\cup\widetilde{B}_s^{(m)}$ for some set $\widetilde{B}_s^{(m)}\subseteq\mathbb{N}_{>m+1}$. Let $\ell_{s,b}^{(m)}$ denote the left endpoint of the interval $I_{s,b}^{(m)}$. Then we define $T_s^{(m)}:[0,1]\to[0,1]$ by

$$T_s^{(m)}(x)=\begin{cases} T_s(x), & \text{if } x\in I_{s,b}^{(m)},\ b\le m,\\ \left|I_{s,b}^{(m)}\right|^{-1}\left(x-\ell_{s,b}^{(m)}\right), & \text{if } x\in I_{s,b}^{(m)},\ b>m,\\ 0, & \text{otherwise.}\end{cases}$$

See Figure 3.2 for examples of GLS maps and their first three associated approximations. Note that we have chosen the labels of the intervals such that the map $T_s^{(m+1)}$ equals $T_s^{(m)}$ on all intervals except $I_{s,m+1}^{(m)}$, which for $T_s^{(m+1)}$ is split into at most three intervals; these intervals are $I_{s,m+1}^{(m+1)}$ and at most two other intervals $I_{s,b'}^{(m+1)}$ for some $b'\in\mathbb{N}$ such that no interval $I_{s,b'}^{(m)}$ exists. For each $m\in\mathbb{N}$, put $\mathcal{T}^{(m)} := \left\{T_s^{(m)}\right\}_{s\in\mathcal{S}}$, and denote by $\mathcal{D}^{(m)} := \left\{(s,b)\in\mathcal{D} : s\in\mathcal{S}, b\in\mathcal{B}_s^{(m)}\right\}$ the associated digit set. Observe then that $\mathcal{D}_m$ is a subset of both $\mathcal{D}$ and $\mathcal{D}^{(m)}$. In fact, $\mathcal{D}_m$ is precisely the set of labels of branches that are shared by each T_s and $T_s^{(m)}$, $s\in\mathcal{S}$.

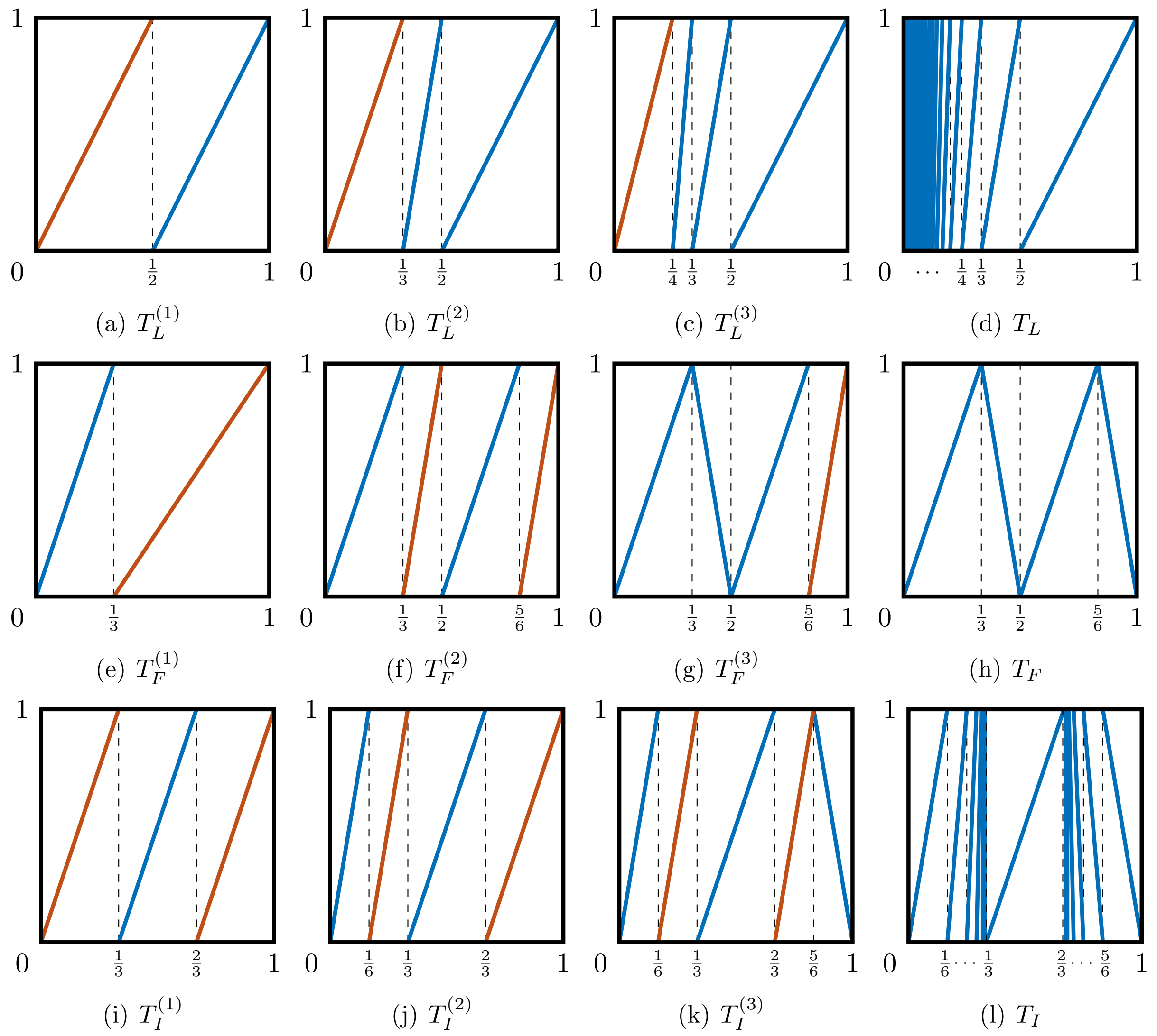


Figure 3.2: Three examples of GLS maps T_s with their first three approximations $T_s^{(m)}$ for $m = 1, 2, 3$. In (a)-(c), we see the first three approximations of the Lüroth map T_L shown in (d). In (e)-(g), we see the first three approximations of the GLS map T_F with finite set $\mathcal{B}_F$ shown in (h). In (i)-(k), we see the first three approximations of the GLS map T_I with infinite set $\mathcal{B}_I$ shown in (l).

Given a frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$, define the frequency vector $\boldsymbol{\alpha}^{(m)} = (\alpha_d^{(m)})_{d\in\mathcal{D}^{(m)}}$ by

$$\alpha_{(s,b)}^{(m)} := \begin{cases} \alpha_{(s,b)}, & \text{if } b \leq m, \\ \displaystyle\sum_{\substack{c\in\mathcal{B}_s \\ I_{s,c}\subset I_{s,b}^{(m)}}} \alpha_{(s,c)}, & \text{otherwise.} \end{cases} \tag{3.30}$$

In other words, $\boldsymbol{\alpha}^{(m)}$ is defined to be the same as $\boldsymbol{\alpha}$ on the intervals $I_{s,b}^{(m)}$, $b \leq m$, and is otherwise the sum of the entries of $\boldsymbol{\alpha}$ whose indices label intervals $I_{s,c}$ contained in $I_{s,b}^{(m)}$, $b > m$. For each $c \in \mathcal{B}_s^{(m)}$, set $N_{s,c}^{(m)} := \left|T_s^{(m)\prime}\right|$ on $I_{s,c}^{(m)}$. To distinguish the FFIs of an NGLS $(\mathcal{T}, \omega)$ and an approximation NGLS $(\mathcal{T}^{(m)}, \omega)$, $m \in \mathbb{N}$, we shall write $\langle b_1 \cdots b_n\rangle_\omega^{(m)}$ with $b_\ell \in \mathcal{B}_{\omega_\ell}^{(m)}$ for each

$1 \leq \ell \leq n$ and $n \in \mathbb{N}$ for the FFIs of the latter NGLS. Given the definition of the frequency vector $\boldsymbol{\alpha}^{(m)}$ in (3.30), define the Besicovitch-Eggleston $\boldsymbol{\alpha}^{(m)}$-level set for $(\mathcal{T}^{(m)}, \omega)$ by

$$F_{\mathcal{T}^{(m)},\omega}(\boldsymbol{\alpha}^{(m)}) := \left\{ x \in X_\omega^{(m)} : \lim_{n\to+\infty} \frac{\#\{1 \leq \ell \leq n : (\omega_\ell, b_\ell^{(m)}(\omega, x)) = d\}}{n} = \alpha_d^{(m)}\, \forall d \in \mathcal{D}^{(m)} \right\},$$

where $X_\omega^{(m)}$ is defined similarly to X_ω but instead with respect to $(\mathcal{T}^{(m)}, \omega)$ (recall (3.5)), and $b_\ell^{(m)}(\omega, x)$ is the ℓ^{th} digit in the $(\mathcal{T}^{(m)}, \omega)$-expansion of x.

In Lemma 3.14 below, we will construct a probability measure $\mu_{\omega,\boldsymbol{\alpha}}$ supported on $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$. The method used will allow us to obtain as a corollary the existence of probability measures $\mu_{\omega,\boldsymbol{\alpha}}^{(m)}$ supported on $F_{\mathcal{T}^{(m)},\omega}(\boldsymbol{\alpha}^{(m)})$, $m \in \mathbb{N}$. These measures will allow us to obtain the lower bound $\beta_{\mathcal{T}}(\boldsymbol{\alpha})$ for $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ in the next section. We will deduce the existence of $\mu_{\omega,\boldsymbol{\alpha}}$ via Carathéodory's extension theorem, but, before we can do this, we first need to construct a measure-theoretic ring of subsets $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$ of $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ and a candidate pre-measure $\mu^*_{\omega,\boldsymbol{\alpha}}$ on $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$. Put

$$\mathcal{F}_{\omega,\boldsymbol{\alpha}} := \big\{A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) : A \text{ is a countable union of FFIs of } (\mathcal{T}, \omega)\big\} \cup \{\emptyset\}.$$

For $A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \in \mathcal{F}_{\omega,\boldsymbol{\alpha}}$, we define the $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$-maximal partition to be the unique partition $\mathcal{P} \subseteq \mathcal{F}_{\omega,\boldsymbol{\alpha}}$ of $A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ into non-empty elements of $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$ such that, for each $P \in \mathcal{P}$, there does not exist an FFI B with $P \subsetneq B \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subsetneq A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$, and we denote it by $\mathcal{P}_A$. We then set $\mu^*_{\omega,\boldsymbol{\alpha}}(\emptyset) := 0$ and

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) := \sum_{\langle b_1\cdots b_n\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \in \mathcal{P}_A} \prod_{1\leq\ell\leq n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}}$$

for any countable union of FFIs A. Note that if $A = \langle b_1 \cdots b_n\rangle_\omega$ is an FFI, then it follows that $\mathcal{P}_A = \{A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\}$ and $\mu^*_{\omega,\boldsymbol{\alpha}}(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})) = \prod_{1\leq\ell\leq n} \alpha_{(\omega_\ell,b_\ell)}/\alpha_{\omega_\ell} < +\infty$. From this, one can deduce for each $m \geq n$ that

$$\begin{aligned}
&\sum_{\substack{c_1\cdots c_m \in \mathcal{B}_\omega^m \\ \langle c_1\cdots c_m\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})}} \mu^*_{\omega,\boldsymbol{\alpha}}\big(\langle c_1 \cdots c_m\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) \\
&\quad= \sum_{\substack{c_1\cdots c_m \in \mathcal{B}_\omega^m \\ \langle c_1\cdots c_m\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})}} \prod_{1\leq\ell\leq m} \frac{\alpha_{(\omega_\ell,c_\ell)}}{\alpha_{\omega_\ell}} \\
&\quad= \sum_{\substack{c_1\cdots c_{m-1} \in \mathcal{B}_\omega^{m-1} \\ \langle c_1\cdots c_{m-1}\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})}} \left(\prod_{1\leq\ell\leq m-1} \frac{\alpha_{(\omega_\ell,c_\ell)}}{\alpha_{\omega_\ell}}\right) \sum_{c_m \in \mathcal{B}_{\omega_m}} \frac{\alpha_{(\omega_m,c_m)}}{\alpha_{\omega_m}} \\
&\quad= \sum_{\substack{c_1\cdots c_{m-1} \in \mathcal{B}_\omega^{m-1} \\ \langle c_1\cdots c_{m-1}\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})}} \prod_{1\leq\ell\leq m-1} \frac{\alpha_{(\omega_\ell,c_\ell)}}{\alpha_{\omega_\ell}} = \cdots \\
&\quad= \sum_{\substack{c_1\cdots c_n \in \mathcal{B}_\omega^n \\ \langle c_1\cdots c_n\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})}} \prod_{1\leq\ell\leq n} \frac{\alpha_{(\omega_\ell,c_\ell)}}{\alpha_{\omega_\ell}} = \prod_{1\leq\ell\leq n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}} \\
&\quad= \mu^*_{\omega,\boldsymbol{\alpha}}\big(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big). \qquad (3.31)
\end{aligned}$$

We have the following lemma on the value of $\mu^*_{\omega,\boldsymbol{\alpha}}$ on FFIs.

Lemma 3.13. *Fix a frequency vector* $\boldsymbol{\alpha}$ *satisfying* $(*)$*, and suppose that there is an* $s \in \mathcal{S}$ *for which* $0 < \alpha_{(s,b)} < \alpha_s$ *for some* $b \in \mathcal{B}_s$*. Let* $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$*. Then for each* $\varepsilon > 0$ *there is a* $K \in \mathbb{N}$ *such that for all* $n \geq K$ *and for all* $b_1 \cdots b_n \in \mathcal{B}_\omega^n$*,*

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) < \varepsilon.$$

Proof. The assumption implies that $0 \leq \alpha_{(s,b)} < \alpha_s$ for all $b \in \mathcal{B}_s$ so, for any $b_1 \cdots b_n \in \mathcal{B}_\omega^n$, we have

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \prod_{1\leq \ell\leq n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}} \leq \left(\frac{\max\{\alpha_{(s,b)} : b \in \mathcal{B}_s\}}{\alpha_s}\right)^{\tau_s(\omega,n)}.$$

The result follows since $\tau_s(\omega, n) \to +\infty$ as $n \to +\infty$. □

The following lemma states that $\mu^*_{\omega,\boldsymbol{\alpha}}$ can be extended to a probability measure.

Lemma 3.14. *Fix a frequency vector* $\boldsymbol{\alpha}$ *satisfying* $(*)$*, and suppose that there is an* $s \in \mathcal{S}$ *for which* $0 < \alpha_{(s,b)} < \alpha_s$ *for some* $b \in \mathcal{B}_s$*. Then* $\mu^*_{\omega,\boldsymbol{\alpha}}$ *is a pre-measure over the measure-theoretic ring of subsets* $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$ *of* $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ *and extends to a probability measure* $\mu_{\omega,\boldsymbol{\alpha}}$ *on the measurable space* $\big([0,1], \mathcal{B}([0,1])\big)$*. Furthermore,* $\mu_{\omega,\boldsymbol{\alpha}}$ *has support* $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ *and satisfies*

$$\mu_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle_\omega\big) = \prod_{1\leq \ell\leq n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}}, \qquad \forall b_1 \cdots b_n \in \mathcal{B}_\omega^n,\ n \in \mathbb{N}. \tag{3.32}$$

Proof. To show that the map $\mu^*_{\omega,\boldsymbol{\alpha}}$ is a pre-measure on $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$ we need to prove for any countable union of FFIs A and any countable collection of countable unions of FFIs $(A_j)_{j\in J}$ so that $A = \bigcup_{j\in J} A_j$ that

$$\mu^*_{\omega,\boldsymbol{\alpha}}(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})) = \sum_{j\in J} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A_j \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big).$$

We split the proof that the map $\mu^*_{\omega,\boldsymbol{\alpha}}$ is a pre-measure on $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$ into a couple of steps.

Step 1: A bound on the level of the FFIs

The first step is to consider the situation where there is a bound on the level of the FFIs involved. Let $A = \langle b_1 \cdots b_n\rangle_\omega$, $A_j = \langle b_1^{(j)} \cdots b_{n_j}^{(j)}\rangle_\omega$ be FFIs and assume that $M := \sup\{n_j : j \in J\} < +\infty$. Then

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \prod_{1\leq \ell\leq n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}} \quad \text{and} \quad \mu^*_{\omega,\boldsymbol{\alpha}}\big(A_j \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \prod_{1\leq \ell\leq n_j} \frac{\alpha_{(\omega_\ell,b_\ell^{(j)})}}{\alpha_{\omega_\ell}} \tag{3.33}$$

for all $j \in J$. Then, using (3.31) for A and each A_j, $j \in J$, yields

$$\begin{aligned}
\mu^*_{\omega,\boldsymbol{\alpha}}\big(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) &= \sum_{\substack{c_1\cdots c_M \in \mathcal{B}_\omega^M \\ \langle c_1\cdots c_M\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})}} \prod_{1\leq \ell\leq M} \frac{\alpha_{(\omega_\ell,c_\ell)}}{\alpha_{\omega_\ell}} \\
&= \sum_{j\in J} \sum_{\substack{c_1\cdots c_M \in \mathcal{B}_\omega^M \\ \langle c_1\cdots c_M\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subset A_j \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})}} \prod_{1\leq \ell\leq M} \frac{\alpha_{(\omega_\ell,c_\ell)}}{\alpha_{\omega_\ell}} \\
&= \sum_{j\in J} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A_j \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big).
\end{aligned} \tag{3.34}$$

Now suppose that A and A_j, $j \in J$, are instead countable unions of FFIs. We shall denote by $\mathcal{P}_A = \{A^{(i)} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) : i \in I\}$ the $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$-maximal partition of A with countable index set I so that each $A^{(i)}$ is an FFI and

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \sum_{i\in I} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A^{(i)} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big).$$

Similarly, for each $j \in J$, write

$$\mathcal{P}_{A_j} = \left\{\left\langle b_1^{(j,k)} \cdots b_{n_{j,k}}^{(j,k)}\right\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) : k \in K_j\right\}$$

for some countable index set K_j so that

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(A_j \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \sum_{k \in K_j} \mu^*_{\omega,\boldsymbol{\alpha}}\big(\langle b_1^{(j,k)} \cdots b_{n_{j,k}}^{(j,k)}\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big).$$

Assume that $\sup\{n_{j,k} : j \in J,\, k \in K_j\} < +\infty$. Set $A_{j,k} := \langle b_1^{(j,k)} \cdots b_{n_{j,k}}^{(j,k)}\rangle_\omega$. Then for each $A_{j,k}$, there is a unique $A^{(i)}$ such that $A_{j,k} \subseteq A^{(i)}$ so define $J^{(i)} := \{(j,k) : A_{j,k} \subseteq A^{(i)},\, j \in J,\, k \in K_j\}$. Since

$$\bigcup_{i \in I} A^{(i)} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = \bigcup_{j \in J} \bigcup_{k \in K_j} A_{j,k} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}),$$

it must hold that $\bigcup_{i \in I} J^{(i)} = \{(j,k) : j \in J,\, k \in K_j\}$. Therefore,

$$\begin{aligned}\mu^*_{\omega,\boldsymbol{\alpha}}\big(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) &= \sum_{i \in I} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A^{(i)} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \sum_{i \in I} \sum_{(j,k) \in J^{(i)}} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A_{j,k} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) \\ &= \sum_{j \in J} \sum_{k \in K_j} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A_{j,k} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \sum_{j \in J} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A_j \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big),\end{aligned}$$

where we have used (3.34) in the second equality.

Step 2: No bound on the level of the FFIs

Suppose that A is an FFI and each A_j, $j \in J$, is a countable union of FFIs. As above, for each $j \in J$, write

$$\mathcal{P}_{A_j} = \left\{\left\langle b_1^{(j,k)} \cdots b_{n_{j,k}}^{(j,k)}\right\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) : k \in K_j\right\}$$

for some countable index set K_j and FFIs $A_{j,k} = \langle b_1^{(j,k)} \cdots b_{n_{j,k}}^{(j,k)}\rangle_\omega$, $k \in K_j$. Now assume that $\sup\{n_{j,k} : j \in J,\, k \in K_j\} = +\infty$. Enumerate the elements of $\{n_{j,k} : j \in J, k \in K_j\}$ by $p_1 < p_2 < p_3 < \cdots$ and for each m set $J_m = \{(j,k) : n_{j,k} = p_m,\, j \in J,\, k \in K_j\}$. For each $m \in \mathbb{N}$, we have by the previous step that

$$\mu^*_{\omega,\boldsymbol{\alpha}}\left(\bigcup_{1 \le \ell \le m} \bigcup_{(j,k) \in J_\ell} A_{j,k} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\right) = \sum_{1 \le \ell \le m} \sum_{(j,k) \in J_\ell} \mu^*_{\omega,\boldsymbol{\alpha}}\big(A_{j,k} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big),$$

and we see that the sequence $(\mu^*_{\omega,\boldsymbol{\alpha}}(\bigcup_{1 \le \ell \le m} \bigcup_{(j,k) \in J_\ell} A_{j,k} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})))_{m \in \mathbb{N}}$ is increasing and the limit as $m \to +\infty$ exists. Also for $m \in \mathbb{N}$, let

$$B_m := A \setminus \bigcup_{1 \le \ell \le m} \bigcup_{(j,k) \in J_\ell} A_{j,k}.$$

Note that B_m is not necessarily a countable union of FFIs as it might contain some of the endpoints of FFIs that are not contained in any FFI or some accumulation points of the branches of T^m_ω. However, since $F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \subseteq X_\omega$, we have that

$$B_m \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = \bigcup_{\ell \in L_m} B_{m,\ell} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$$

for some countable collection $(B_{m,\ell})_{\ell\in L_m}$ of FFIs of level p_m. Therefore, $\mu^*_{\omega,\boldsymbol{\alpha}}(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}))$ is well-defined and, by (3.34),

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)=\sum_{\ell\in L_m}\mu^*_{\omega,\boldsymbol{\alpha}}\big(B_{m,\ell}\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big).$$

Note that

$$B_m=B_{m+1}\cup\bigcup_{(j,k)\in J_{m+1}}A_{j,k},\tag{3.35}$$

so, by Step 1, $(\mu^*_{\omega,\boldsymbol{\alpha}}(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})))_{m\in\mathbb{N}}$ is a decreasing sequence and the limit as $m\to+\infty$ exists. For each $m\in\mathbb{N}$, we can write

$$A=\bigcup_{1\le\ell\le m}\bigcup_{(j,k)\in J_\ell}A_{j,k}\cup B_m$$

and thus we have by Step 1 that

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)=\sum_{1\le\ell\le m}\sum_{(j,k)\in J_\ell}\mu^*_{\omega,\boldsymbol{\alpha}}\big(A_{j,k}\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)+\mu^*_{\omega,\boldsymbol{\alpha}}\big(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big).$$

Taking the limit as $m\to+\infty$ yields

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)=\sum_{j\in J}\mu^*_{\omega,\boldsymbol{\alpha}}\big(A_j\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)+\lim_{m\to+\infty}\mu^*_{\omega,\boldsymbol{\alpha}}\big(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big).$$

3

What is left to show is that $\mu^*_{\omega,\boldsymbol{\alpha}}(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}))\to 0$ as $m\to+\infty$. Suppose to the contrary that $\mu^*_{\omega,\boldsymbol{\alpha}}(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}))\nrightarrow 0$ as $m\to+\infty$. This means that there exists an $\varepsilon>0$ such that, for all $m\in\mathbb{N}$, we have $\mu^*_{\omega,\boldsymbol{\alpha}}(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}))>2\varepsilon$. By Step 1, we know for each $m\in\mathbb{N}$ and $m'\ge p_m$ that

$$\begin{aligned}\mu^*_{\omega,\boldsymbol{\alpha}}\big(B_m\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)&=\sum_{\substack{d_1\cdots d_{m'}\in\mathcal{B}^{m'}_\omega\\ \langle d_1\cdots d_{m'}\rangle_\omega\subseteq B_m}}\prod_{1\le\ell\le m'}\frac{\alpha_{\omega_\ell,d_\ell}}{\alpha_{\omega_\ell}}\\&\le\sum_{\substack{d_1\cdots d_{m'}\in\mathcal{B}^{m'}_\omega\\ \langle d_1\cdots d_{m'}\rangle_\omega\subseteq A}}\prod_{1\le\ell\le m'}\frac{\alpha_{\omega_\ell,d_\ell}}{\alpha_{\omega_\ell}}=\mu^*_{\omega,\boldsymbol{\alpha}}\big(A\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)<+\infty.\end{aligned}$$

Therefore, by Lemma 3.13 there is a $\kappa_1\ge p_m$ such that for each $k\ge\kappa_1$ we can find a finite subset $\mathcal{E}^{(m,k)}\subseteq\{d_1\cdots d_k\in\mathcal{B}^k_\omega:\langle d_1\cdots d_k\rangle_\omega\subseteq B_m\}$ for which

$$\sum_{d_1\cdots d_k\in\mathcal{E}^{(m,k)}}\prod_{1\le\ell\le k}\frac{\alpha_{\omega_\ell,d_\ell}}{\alpha_{\omega_\ell}}\ge\sum_{\substack{d_1\cdots d_k\in\mathcal{B}^k_\omega\\ \langle d_1\cdots d_k\rangle_\omega\subseteq B_m}}\prod_{1\le\ell\le k}\frac{\alpha_{\omega_\ell,d_\ell}}{\alpha_{\omega_\ell}}-\frac{\varepsilon}{2^{m+1}}.$$

Let $E_{k,l}$ and $E_{k,r}$ denote the left and rightmost FFIs from $\{\langle d_1\cdots d_k\rangle_\omega : d_1\cdots d_k\in\mathcal{E}^{(m,k)}\}$ respectively. By Lemma 3.13, we can find a $\kappa=\kappa(m)\ge\kappa_1$ such that

$$\mu^*_{\omega,\boldsymbol{\alpha}}\big(E_{\kappa,l}\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)+\mu^*_{\omega,\boldsymbol{\alpha}}\big(E_{\kappa,r}\cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)<\frac{\varepsilon}{2^{m+1}}.$$

Let

$$C_m=\bigcup_{d_1\cdots d_\kappa\in\mathcal{E}^{(m,\kappa)}}\langle d_1\cdots d_\kappa\rangle_\omega\setminus(E_{\kappa,l}\cup E_{\kappa,r}).$$

Then $\overline{C_m} \subseteq B_m$ and moreover, by (3.34)

$$\begin{aligned}\mu^*_{\omega,\boldsymbol{\alpha}}\big(C_m \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) &= \sum_{d_1\cdots d_{\mathrm{K}} \in \mathcal{E}^{(m,\mathrm{K})}} \prod_{1\le \ell \le \mathrm{K}} \frac{\alpha_{\omega_\ell,d_\ell}}{\alpha_{\omega_\ell}} - \mu^*_{\omega,\boldsymbol{\alpha}}\big(E_{\mathrm{K},l} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) - \mu^*_{\omega,\boldsymbol{\alpha}}\big(E_{\mathrm{K},r} \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big)\\ &\ge \sum_{\substack{d_1\cdots d_{\mathrm{K}} \in \mathcal{B}^{\mathrm{K}}_\omega \\ \langle d_1\cdots d_{\mathrm{K}}\rangle_\omega \subseteq B_m}} \prod_{1\le \ell \le \mathrm{K}} \frac{\alpha_{\omega_\ell,d_\ell}}{\alpha_{\omega_\ell}} - \frac{\varepsilon}{2^{m+1}} - \frac{\varepsilon}{2^{m+1}}\\ &= \mu^*_{\omega,\boldsymbol{\alpha}}\big(B_m \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) - \frac{\varepsilon}{2^m}.\end{aligned}$$

For each $m \in \mathbb{N}$, the set $\bigcap_{1\le k\le m} C_m$ is a countable union of FFIs of level at most $\mathrm{K}(m)$. Therefore, we have for each $m \in \mathbb{N}$ by Step 1 that

$$\begin{aligned}&\mu^*_{\omega,\boldsymbol{\alpha}}\big(B_m \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) - \mu^*_{\omega,\boldsymbol{\alpha}}\Bigg(\bigcap_{1\le k\le m} C_k \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\Bigg)\\ &\quad= \mu^*_{\omega,\boldsymbol{\alpha}}\Bigg(\Bigg(B_m \setminus \bigcap_{1\le k\le m} C_k\Bigg) \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\Bigg) \le \mu^*_{\omega,\boldsymbol{\alpha}}\Bigg(\bigcup_{1\le k\le m} (B_k \setminus C_k) \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\Bigg)\\ &\quad\le \sum_{1\le k\le m} \mu^*_{\omega,\boldsymbol{\alpha}}\big((B_k \setminus C_k) \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) \le \sum_{1\le k\le m} \frac{\varepsilon}{2^k} \le \varepsilon.\end{aligned}$$

However, we also have $\mu^*_{\omega,\boldsymbol{\alpha}}(B_m \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})) \ge 2\varepsilon$ so we must have

$$\mu^*_{\omega,\boldsymbol{\alpha}}\Bigg(\bigcap_{1\le k\le m} C_k \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\Bigg) \ge \varepsilon, \qquad \forall m \in \mathbb{N}.$$

For each $m \in \mathbb{N}$, if we put $K_m := \bigcap_{1\le k\le m} \overline{C_k}$, then K_m is compact with $\mu^*_{\omega,\boldsymbol{\alpha}}(K_m \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})) \ge \varepsilon$, which means that $K_m \ne \emptyset$. Therefore, the fact that $K_m \subset \overline{C_m} \subset B_m$ for all $m \in \mathbb{N}$ together with (3.35) yields

$$\emptyset \ne \bigcap_{m\in\mathbb{N}} K_m \subset \bigcap_{m\in\mathbb{N}} B_m = \emptyset,$$

which is a contradiction. Hence,

$$\lim_{m\to+\infty} \mu^*_{\omega,\boldsymbol{\alpha}}\big(B_m \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = 0$$

and thus $\mu^*_{\omega,\boldsymbol{\alpha}}(A \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})) = \sum_{j\in J} \mu^*_{\omega,\boldsymbol{\alpha}}(A_j \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha}))$ also in this case.

To show countable additivity, it only remains to consider the case when A and A_j, $j \in J$, are both countable unions of FFIs. Having established the result for FFIs A, this now follows by the same reasoning used at the end of Step 1. Hence, $\mu^*_{\omega,\boldsymbol{\alpha}}$ is countably additive.

Step 3: Existence of the probability measure $\mu_{\omega,\boldsymbol{\alpha}}$ satisfying (3.32)

From the above, it follows that $\mu^*_{\omega,\boldsymbol{\alpha}}$ is a pre-measure on $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$ and so extends to a measure $\mu_{\omega,\boldsymbol{\alpha}}$ on $\big(F_{\mathcal{T},\omega}(\boldsymbol{\alpha}), \sigma(\mathcal{F}_{\omega,\boldsymbol{\alpha}})\big)$ by Carathéodory's extension theorem. We may further extend $\mu_{\omega,\boldsymbol{\alpha}}$ to the measurable space $\big([0,1], \mathcal{B}([0,1])\big)$ by setting

$$\mu_{\omega,\boldsymbol{\alpha}}(B) = \mu_{\omega,\boldsymbol{\alpha}}\big(B \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big), \qquad \forall B \in \mathcal{B}([0,1]).$$

Since $\mu_{\omega,\boldsymbol{\alpha}}$ agrees with $\mu^*_{\omega,\boldsymbol{\alpha}}$ on $\mathcal{F}_{\omega,\boldsymbol{\alpha}}$, (3.32) follows from (3.33). To see that $\mu_{\omega,\boldsymbol{\alpha}}$ is a probability measure, observe that

$$\mu_{\omega,\boldsymbol{\alpha}}\big(F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = \mu_{\omega,\boldsymbol{\alpha}}\Bigg(\bigcup_{b\in\mathcal{B}_{\omega_1}} \langle b\rangle_\omega \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\Bigg) = \sum_{b\in\mathcal{B}_{\omega_1}} \frac{\alpha_{(\omega_1,b)}}{\alpha_{\omega_1}} = 1.$$

Then, since $\mu_{\omega,\boldsymbol{\alpha}}(B) = \mu_{\omega,\boldsymbol{\alpha}}(B \cap F_{\mathcal{T},\omega}(\boldsymbol{\alpha})) \leq \mu_{\omega,\boldsymbol{\alpha}}(F_{\mathcal{T},\omega}(\boldsymbol{\alpha}))$ for all $B \in \mathcal{B}([0,1])$, it follows that $\mu_{\omega,\boldsymbol{\alpha}}$ is a probability measure with support $F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$. □

We obtain the following corollary of Lemma 3.14 for the NGLS maps $(\mathcal{T}^{(m)}, \omega)$, $m \in \mathbb{N}$.

Corollary 3.15. *Fix $m \in \mathbb{N}$, and suppose that there exists an $s \in \mathcal{S}$ such that $0 < \alpha^{(m)}_{(s,b)} < \alpha_s$ for some $b \in \mathcal{B}^{(m)}_s$. Then there is a probability measure $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$ defined on the measurable space $\big([0,1], \mathcal{B}([0,1])\big)$ that is supported on $F_{\mathcal{T}^{(m)},\omega}(\boldsymbol{\alpha}^{(m)})$ with*

$$\mu^{(m)}_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle^{(m)}_\omega\big) = \prod_{1\leq \ell\leq n} \frac{\alpha^{(m)}_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}}, \qquad \forall b_1 \cdots b_n \in \prod_{1\leq \ell \leq n} \mathcal{B}^{(m)}_{\omega_\ell}. \tag{3.36}$$

The next lemma relates the measures $\mu_{\omega,\boldsymbol{\alpha}}$ and $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$, $m \in \mathbb{N}$.

Lemma 3.16. *For every $B \in \mathcal{B}([0,1])$,*

$$\mu_{\omega,\boldsymbol{\alpha}}(B) = \lim_{m\to+\infty} \mu^{(m)}_{\omega,\boldsymbol{\alpha}}(B).$$

Proof. As $\mathcal{F} \cup \{\emptyset\}$ generates $\mathcal{B}([0,1])$ and $\mu_{\omega,\boldsymbol{\alpha}}(\emptyset) = \mu^{(m)}_{\omega,\boldsymbol{\alpha}}(\emptyset) = 0$ for all $m \in \mathbb{N}$, it suffices to show for each $b_1 \cdots b_n \in \mathcal{B}^n_\omega$ that there is an $m_0 \in \mathbb{N}$ such that

$$\mu^{(m)}_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle_\omega\big) = \mu_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle_\omega\big)$$

for all $m \geq m_0$. So, fix $n \in \mathbb{N}$ and $b_1 \cdots b_n \in \mathcal{B}^n_\omega$. Then $m_0 := \max\{b_\ell : 1 \leq \ell \leq n\} + 1$ is such that $(\omega_\ell, b_\ell) \in \mathcal{D}_m \subset \mathcal{D}^{(m)}$ for all $1 \leq \ell \leq n$ and all $m \geq m_0$. Thus, by (3.32), (3.36) and the definition of $\boldsymbol{\alpha}^{(m)}$, we have

$$\mu^{(m)}_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle_\omega\big) = \mu^{(m)}_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle^{(m)}_\omega\big) = \prod_{1\leq \ell\leq n} \frac{\alpha^{(m)}_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}} = \prod_{1\leq \ell\leq n} \frac{\alpha_{(\omega_\ell,b_\ell)}}{\alpha_{\omega_\ell}} = \mu_{\omega,\boldsymbol{\alpha}}\big(\langle b_1 \cdots b_n\rangle_\omega\big)$$

for all $m \geq m_0$. □

3.6.2 Proving Theorem 3.3

We first show that $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \geq \beta_{\mathcal{T}}(\boldsymbol{\alpha})$ for all $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ using the NGLS maps $(\mathcal{T}^{(m)}, \omega)$, $m \in \mathbb{N}$, constructed in §3.6.1, which will then allow us to prove Theorem 3.3 at the end of this section. To prove the lower bound, we begin with the case when $\sum_{d\in\mathcal{D}} \alpha_d \log N_d < +\infty$.

Proposition 3.17. *If $\sum_{d\in\mathcal{D}} \alpha_d \log N_d < +\infty$, then $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \geq \beta_{\mathcal{T}}(\boldsymbol{\alpha})$ for all $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$.*

Proof. Fix $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$. We first prove a trivial case. If there is no $s \in \mathcal{S}$ for which $0 < \alpha_{(s,b)} < \alpha_s$ for some $b \in \mathcal{B}_s$, then it follows for each $s \in \mathcal{S}$ that there is a $c_s \in \mathcal{B}_s$ such that $\alpha_{(s,c_s)} = \alpha_s$ and $\alpha_{(s,b)} = 0$ for all $b \in \mathcal{B}_s \setminus \{c_s\}$. Thus, we have

$$\sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s = \sum_{s\in\mathcal{S}} \alpha_{(s,c_s)} \log \alpha_{(s,c_s)} = \sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d,$$

using the convention that $0 \log 0 := 0$, which implies that

$$\beta_{\mathcal{T}}(\boldsymbol{\alpha}) = \liminf_{m\to+\infty} \frac{\sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s - \sum_{d\in\mathcal{D}} \alpha_d \log \alpha_d}{\sum_{d\in\mathcal{D}_m} \alpha_d \log N_d} = 0 \leq \dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}).$$

Therefore, suppose that there is an $s \in \mathcal{S}$ such that $0 < \alpha_{(s,b)} < \alpha_s$ for some $b \in \mathcal{B}_s$, and note that this implies that there is an $M \in \mathbb{N}$ such that, for each $m \geq M$, $0 < \alpha^{(m)}_{(s,b)} < \alpha_s$ for the same $s \in \mathcal{S}$ and for some $b \in \mathcal{B}^{(m)}_s$. Let $\mu_{\omega,\boldsymbol{\alpha}}$ be the measure defined in Lemma 3.14, and, for each $m \geq M$, let $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$ be the measure defined in Corollary 3.15. To obtain the desired lower bound, the aim is to relate $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})$ with the Hausdorff dimension of each of the measures $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$ for $m \in \mathbb{N}$ sufficiently large. Since these measures are defined on limit sets constructed using finite digit sets, we will then be able to use the method in [BI09, Theorem 3.1] to show that

$$\underline{d}_{\mu^{(m)}_{\omega,\boldsymbol{\alpha}}}(x) \geq c_m \frac{\sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s - \sum_{d\in\mathcal{D}_m} \alpha_d \log \alpha_d}{\sum_{d\in\mathcal{D}_m} \alpha_d \log N_d}, \qquad \mu^{(m)}_{\omega,\boldsymbol{\alpha}}\text{-a.e. } x \in [0,1], \tag{3.37}$$

where $(c_m)_{m\in\mathbb{N}}$ is a sequence of constants such that $c_m \to 1$ as $m \to +\infty$. To complete the proof, we will apply Lemma 1.5; so, fix $\delta > 0$. Since $\mu_{\omega,\boldsymbol{\alpha}}\big(F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) = 1$, Lemma 3.16 implies that there is an $M(\delta) \geq M$ such that $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}\big(F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\big) > 1-\delta$ for all $m \geq M(\delta)$. Fix $m \geq M(\delta)$. Then

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \geq \inf\big\{\dim_H A : A \in \mathcal{B}([0,1]),\ \mu^{(m)}_{\omega,\boldsymbol{\alpha}}(A) > 1-\delta\big\} \geq \dim_H \mu^{(m)}_{\omega,\boldsymbol{\alpha}} - E(\delta), \tag{3.38}$$

where

$$E(\delta) := \sup_{m'\geq M(\delta)} \Big|\dim_H \mu^{(m')}_{\omega,\boldsymbol{\alpha}} - \inf\big\{\dim_H A : A \in \mathcal{B}([0,1]),\ \mu^{(m')}_{\omega,\boldsymbol{\alpha}}(A) > 1-\delta\big\}\Big|.$$

From (1.7), we see for any $m' \geq M(\delta)$ that

$$\dim_H \mu^{(m')}_{\omega,\boldsymbol{\alpha}} = \lim_{\varepsilon\to 0} \inf\{\dim_H A : A \in \mathcal{B}([0,1]),\ \mu^{(m')}_{\omega,\boldsymbol{\alpha}}(A) > 1-\varepsilon\}.$$

Then

$$\lim_{\varepsilon\to 0} \Big|\dim_H \mu^{(m')}_{\omega,\boldsymbol{\alpha}} - \inf\big\{\dim_H A : A \in \mathcal{B}([0,1]),\ \mu^{(m')}_{\omega,\boldsymbol{\alpha}}(A) > 1-\varepsilon\big\}\Big| = 0$$

so, since $M(\delta)$ increases to $+\infty$ as δ decreases to 0, we see that $E(\delta)$ is monotone decreasing to 0 as $\delta \to 0$.

By Lemma 1.5, proving (3.37) will mean that the same lower bound will hold for $\dim_H \mu^{(m)}_{\omega,\boldsymbol{\alpha}}$. By a standard argument in dimension theory (see e.g. [Pes97, Theorem 15.3]), replacing the balls in the definition of the (lower) pointwise dimension of $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$ at x with FFIs of $(\mathcal{T}^{(m)},\omega)$ yields a lower bound for $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$-a.e. $x \in [0,1]$:

$$\begin{aligned}\underline{d}_{\mu^{(m)}_{\omega,\boldsymbol{\alpha}}}(x) &:= \liminf_{r\to 0} \frac{\log \mu^{(m)}_{\omega,\boldsymbol{\alpha}}\big(B(x,r)\big)}{\log r} \geq \liminf_{n\to+\infty} \frac{\log \mu^{(m)}_{\omega,\boldsymbol{\alpha}}\big(\langle b^{(m)}_1(\omega,x)\cdots b^{(m)}_n(\omega,x)\rangle^{(m)}_\omega\big)}{\log\big|\langle b^{(m)}_1(\omega,x)\cdots b^{(m)}_n(\omega,x)\rangle^{(m)}_\omega\big|} \\ &= \liminf_{n\to+\infty} \frac{\frac{1}{n}\sum_{1\leq\ell\leq n} \log(\alpha^{(m)}_{(\omega_\ell, b^{(m)}_\ell(\omega,x))}/\alpha_{\omega_\ell})}{-\frac{1}{n}\sum_{1\leq\ell\leq n} \log N^{(m)}_{\omega_\ell, b^{(m)}_\ell(\omega,x)}}.\end{aligned}$$

By Corollary 3.15, $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$-a.e. $x \in [0,1]$ belongs to $F_{\mathcal{T}^{(m)},\omega}(\boldsymbol{\alpha}^{(m)})$ so we may collect like terms to see that for such x,

$$\begin{aligned}
&\lim_{n\to+\infty}\frac{1}{n}\sum_{1\le\ell\le n}\log\frac{\alpha^{(m)}_{(\omega_\ell,b^{(m)}_\ell(\omega,x))}}{\alpha_{\omega_\ell}}\\
&\qquad=\lim_{n\to+\infty}\left(\sum_{d\in\mathcal{D}^{(m)}}\frac{\tau_d\big(\omega,b^{(m)}_1(\omega,x)\cdots b^{(m)}_n(\omega,x)\big)}{n}\log\alpha^{(m)}_d-\sum_{s\in\mathcal{S}}\frac{\tau_s(\omega,n)}{n}\log\alpha_s\right)\\
&\qquad=\sum_{d\in\mathcal{D}^{(m)}}\alpha^{(m)}_d\log\alpha^{(m)}_d-\sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s
\end{aligned}$$

and

$$\begin{aligned}
\lim_{n\to+\infty}\frac{1}{n}\sum_{1\le\ell\le n}\log N^{(m)}_{\omega_\ell,b^{(m)}_\ell(\omega,x)}&=\lim_{n\to+\infty}\sum_{d\in\mathcal{D}^{(m)}}\frac{\tau_d\big(\omega,b^{(m)}_1(\omega,x)\cdots b^{(m)}_n(\omega,x)\big)}{n}\log N^{(m)}_d\\
&=\sum_{d\in\mathcal{D}^{(m)}}\alpha^{(m)}_d\log N^{(m)}_d>0.
\end{aligned}$$

Since the sets $\mathcal{D}^{(m)}$ and $\mathcal{S}$ are finite, we therefore have for $\mu^{(m)}_{\omega,\boldsymbol{\alpha}}$-a.e. $x\in[0,1]$ that

$$\begin{aligned}
\underline{d}_{\mu^{(m)}_{\omega,\boldsymbol{\alpha}}}(x)&\ge\frac{\sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s-\sum_{d\in\mathcal{D}^{(m)}}\alpha^{(m)}_d\log\alpha^{(m)}_d}{\sum_{d\in\mathcal{D}^{(m)}}\alpha^{(m)}_d\log N^{(m)}_d}\\
&\ge\frac{\sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s-\sum_{d\in\mathcal{D}_m}\alpha_d\log\alpha_d}{\sum_{d\in\mathcal{D}_m}\alpha_d\log N_d+\sum_{d\in\mathcal{D}^{(m)}\setminus\mathcal{D}_m}\alpha^{(m)}_d\log N^{(m)}_d}\\
&=c_m\frac{\sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s-\sum_{d\in\mathcal{D}_m}\alpha_d\log\alpha_d}{\sum_{d\in\mathcal{D}_m}\alpha_d\log N_d},
\end{aligned}$$

where

$$c_m:=\frac{\sum_{d\in\mathcal{D}_m}\alpha_d\log N_d}{\sum_{d\in\mathcal{D}_m}\alpha_d\log N_d+\sum_{d\in\mathcal{D}^{(m)}\setminus\mathcal{D}_m}\alpha^{(m)}_d\log N^{(m)}_d}.$$

Therefore, (3.37) holds and so putting (3.37), (3.38) and Lemma 1.5 together yields

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\ge c_m\frac{\sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s-\sum_{d\in\mathcal{D}_m}\alpha_d\log\alpha_d}{\sum_{d\in\mathcal{D}_m}\alpha_d\log N_d}-E(\delta)$$

for all $m\ge M(\delta)$. Now, for any $m\ge M(\delta)$, observe that $N_{s,c}\ge N^{(m)}_{s,b}$ for any $(s,b)\in\mathcal{D}^{(m)}$ and $(s,c)\in\mathcal{D}$ such that $I_{s,c}\subset I^{(m)}_{s,b}$ so the definition of $\alpha^{(m)}$ yields

$$\begin{aligned}
\sum_{d\in\mathcal{D}^{(m)}\setminus\mathcal{D}_m}\alpha^{(m)}_d\log N^{(m)}_d&=\sum_{s\in\mathcal{S}}\sum_{\substack{b\in\mathcal{B}^{(m)}_s\\ b>m}}\sum_{\substack{c\in\mathcal{B}_s\\ I_{s,c}\subset I^{(m)}_{s,b}}}\alpha_{(s,c)}\log N^{(m)}_{s,b}\le\sum_{s\in\mathcal{S}}\sum_{\substack{b\in\mathcal{B}^{(m)}_s\\ b>m}}\sum_{\substack{c\in\mathcal{B}_s\\ I_{s,c}\subset I^{(m)}_{s,b}}}\alpha_{(s,c)}\log N_{s,c}\\
&=\sum_{d\in\mathcal{D}\setminus\mathcal{D}_m}\alpha_d\log N_d<+\infty
\end{aligned}$$

by the assumption. Thus, we may take $m\to+\infty$ to see that $\sum_{d\in\mathcal{D}^{(m)}\setminus\mathcal{D}_m}\alpha^{(m)}_d\log N^{(m)}_d\to 0$ and so $c_m\to 1$ as $m\to+\infty$, as desired. Hence, taking $m\to+\infty$ gives

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha})\ge\liminf_{m\to+\infty}\left(\frac{\sum_{s\in\mathcal{S}}\alpha_s\log\alpha_s-\sum_{d\in\mathcal{D}_m}\alpha_d\log\alpha_d}{\sum_{d\in\mathcal{D}_m}\alpha_d\log N_d}-E(\delta)\right)=\beta_{\mathcal{T}}(\boldsymbol{\alpha})-E(\delta).$$

Since $E(\delta)\to 0$ as $\delta\to 0$, we may take $\delta\to 0$ to obtain the desired lower bound. □

Proof of Theorem 3.3. The upper bound is precisely Proposition 3.4.

In the case that $\sum_{d\in\mathcal{D}} \alpha_d \log N_d < +\infty$, then the lower bound follows from Propositions 3.9 and 3.17. If $\sum_{d\in\mathcal{D}} \alpha_d \log N_d = +\infty$, then we have for all $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ that

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) \geq \eta_{\mathcal{T}} \geq \beta_{\mathcal{T}}(\boldsymbol{\alpha}). \tag{3.39}$$

The first inequality follows from Proposition 3.9 and the second inequality follows analogously to [FLMW10, Lemma 2.4]. In other words, we have $\max\{\eta_{\mathcal{T}}, \beta_{\mathcal{T}}(\boldsymbol{\alpha})\} = \eta_{\mathcal{T}}$ so Proposition 3.9 gives the desired lower bound in this case.

The proof of the final statement of Theorem 3.3 follows from Proposition 3.4 and (3.39). □

3.7 Additional Remarks & Examples

Remark 3.18. As for the fibres of GLS IFSs in Chapter 2, the NGLS maps $(\mathcal{T}, \omega)$, $\omega \in \Omega$, are fibre maps of a corresponding two-dimensional skew product map $\Omega \times [0,1] \to \Omega \times [0,1]$ given by $(\omega, x) \mapsto (\sigma_{\mathcal{S}}(\omega), T_{\omega_1}(x))$. We may therefore use Theorem 3.3 to deduce a lower bound for the Hausdorff dimension of the $\boldsymbol{\alpha}$-level sets $F(\boldsymbol{\alpha}) \coloneqq \{(\omega, x) \in \Omega \times [0,1] : \tau_d(\omega, x) = \alpha_d \, \forall d \in \mathcal{D}\}$. Note that the Hausdorff dimension for subsets of $\Omega \times [0,1]$ is defined with respect to the product metric on $\Omega \times [0,1]$ of the metric $\rho_{\mathbf{p}}$ on Ω and of the Euclidean metric on $[0,1]$, where $\rho_{\mathbf{p}}$ is defined for any choice of fixed probability vector $\mathbf{p} = (p_s)_{s\in\mathcal{S}}$ with positive entries by

$$\rho_{\mathbf{p}}(\omega, \omega') = \prod_{1\leq \ell \leq n(\omega,\omega')} p_{\omega_\ell}, \qquad \omega, \omega' \in \Omega,$$

where $n(\omega, \omega') \coloneqq \min\{n \in \mathbb{N} : \omega_{n+1} \neq \omega'_{n+1}\}$. The lower bound in question is in terms of the Ledrappier-Young formula for the $\boldsymbol{\alpha}$-Bernoulli measure $\mu_{\boldsymbol{\alpha}}$ on $\mathcal{D}^{\mathbb{N}}$:

$$\dim_H F(\boldsymbol{\alpha}) \geq \frac{h_{\mu_{\boldsymbol{\alpha}}}(\sigma) - h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{\sum_{d\in\mathcal{D}} \alpha_d \log N_d} + \frac{h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{s\in\mathcal{S}} \alpha_s \log p_s},$$

where $\mathbb{P}_{\boldsymbol{\alpha}}$ is the $(\alpha_s)_{s\in\mathcal{S}}$-Bernoulli measure on Ω. Set $t_0 \coloneqq \max\{\eta_{\mathcal{T}}, \beta_{\mathcal{T}}(\boldsymbol{\alpha})\}$. By Theorem 3.3, we have $\{\omega \in \Omega : \dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = t_0\} = \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ so we may use the same reasoning as in Remark 2.14 to find for any $\omega \in \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ that

$$\dim_H F(\boldsymbol{\alpha}) \geq \dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) + \dim_H \Omega_{\mathcal{T}}(\boldsymbol{\alpha}) = t_0 + \dim_H \Omega_{\mathcal{T}}(\boldsymbol{\alpha}).$$

It remains to compute $\dim_H \Omega_{\mathcal{T}}(\boldsymbol{\alpha})$. Since $\mathbb{P}_{\boldsymbol{\alpha}}$ is $\sigma_{\mathcal{S}}$-ergodic with $\mathbb{P}_{\boldsymbol{\alpha}}([s]) = \alpha_s$ for each $s \in \mathcal{S}$, it follows from Theorem 3.2 that $\Omega_{\mathcal{T}}(\boldsymbol{\alpha})$ is a symbolic Birkhoff level set on the finite alphabet $\mathcal{S}$ with respect to $\mathbb{P}_{\boldsymbol{\alpha}}$ so we may apply [FLMW10, Theorem 1.2] (the symbolic version of Theorem 3.1) to find that

$$\dim_H \Omega_{\mathcal{T}}(\boldsymbol{\alpha}) = \frac{-\sum_{s\in\mathcal{S}} \alpha_s \log \alpha_s}{-\sum_{s\in\mathcal{S}} \alpha_s \log p_s} = \frac{h_{\mathbb{P}_{\boldsymbol{\alpha}}}(\sigma_{\mathcal{S}})}{-\sum_{s\in\mathcal{S}} \alpha_s \log p_s}.$$

Example 3.19. Set $\mathcal{S} = \{1, 2\}$. Let T_1 be the Lüroth map and T_2 be the alternating Lüroth map defined by $T_2(x) = 1 - T_1(x)$, i.e. T_2 is the Lüroth map but the orientation of each of the branches is reversed; see Figure 3.3. Set $\mathcal{T} \coloneqq \{T_s\}_{s\in\mathcal{S}}$ and $\mathcal{D} = \mathcal{S} \times \mathbb{N}$. Since $N_{s,n} = n(n+1)$ for all $(s, n) \in \mathcal{D}$, (3.3) holds and $\eta(T_1) = \eta(T_2) = \frac{1}{2}$, which implies that $\eta_{\mathcal{T}} = \frac{1}{2}$. Fix $\omega \in \mathcal{S}^{\mathbb{N}}$. For

3

any frequency vector $\boldsymbol{\alpha} = (\alpha_d)_{d\in\mathcal{D}}$ satisfying $(*)$, the conclusions of Theorems 3.2 and 3.3 hold, i.e. we have that $\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = 0$ if there is an $s \in \mathcal{S}$ such that $\tau_s(\omega) \neq \alpha_s$ and, otherwise,

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = \max\left\{\frac{1}{2}, \liminf_{m\to+\infty} \frac{\alpha_1 \log \alpha_1 + \alpha_2 \log \alpha_2 - \sum_{s\in\mathcal{S}} \sum_{1\leq n\leq m} \alpha_{(s,n)} \log \alpha_{(s,n)}}{\sum_{1\leq n\leq m} (\alpha_{(1,n)} + \alpha_{(2,n)}) \log n(n+1)}\right\}.$$

In particular, if $\boldsymbol{\alpha}$ additionally satisfies $\sum_{n\in\mathbb{N}} \alpha_{(s,n)} \log n(n+1) < +\infty$ for each $s \in \mathcal{S}$, then the last equation becomes

$$\dim_H F_{\mathcal{T},\omega}(\boldsymbol{\alpha}) = \max\left\{\frac{1}{2}, \frac{\alpha_1 \log \alpha_1 + \alpha_2 \log \alpha_2 - \sum_{s\in\mathcal{S}} \sum_{n\in\mathbb{N}} \alpha_{(s,n)} \log \alpha_{(s,n)}}{\sum_{n\in\mathbb{N}} (\alpha_{(1,n)} + \alpha_{(2,n)}) \log n(n+1)}\right\}.$$

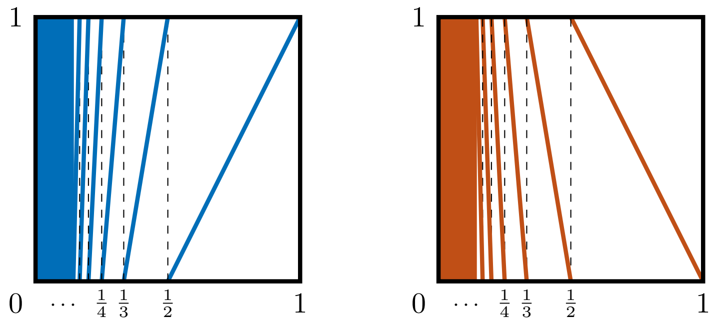


Figure 3.3: The Lüroth map (left) and the alternating Lüroth map (right).

3

Part II

Stable Large Deviations for Deterministic Dynamical Systems

Chapter 4

Non-Cauchy Stable Large Deviations for Deterministic Dynamical Systems

The results in this chapter are based on the article [IT24]:

Stable Large Deviations for Deterministic Dynamical Systems

Abstract

We obtain large deviations for a class of non-square-integrable dependent random variables in the domain of attraction of an α-stable distribution, $\alpha \in (0,1) \cup (1,2]$. This class includes ergodic sums of observables in the domain of attraction of an α-stable distribution driven by Gibbs-Markov maps.

4.1 Introduction

The purpose of this chapter is to obtain large deviations (LD) for a class of dependent random variables satisfying α-stable limit laws, $\alpha \in (0,1) \cup (1,2]$. In particular, we obtain α-stable large deviations, $\alpha \in (0,1) \cup (1,2]$, of a similar form as Theorem 1.8 for a class of deterministic dynamical systems (including Gibbs-Markov maps – recall Definition 1.1 and see §4.4.1 for details). As far as we are aware, this is the first stable LD result for dynamical systems. The case of $\alpha = 1$ has significant difficulties that we do not treat here – see Chapter 5 for related i.i.d. results when $\alpha = 1$.

4

A version of Theorem 1.8 for a class of dependent random variables satisfying certain clustering and limit conditions has been obtained in [MW13], which we discuss briefly in §4.2.1. To our knowledge, this is the only previous work in which a version 'close' to the general form of Theorem 1.8 for dependent random variables has been considered; by close, we mean that (1.12) is shown to hold in a certain range of $N/a_n \to +\infty$ as opposed to the whole region of n and N so that $N/a_n \to +\infty$. It is not clear to us how to verify the limit conditions in [MW13, Theorem 3.1] for the class of dynamical systems considered here, and, in this chapter, we use a different method of proof.

The somewhat related problem of stable local large deviations (LLD)[1] for deterministic dynamical systems (including Gibbs-Markov maps) has been treated in [MT22]. An optimal LLD has been obtained very recently in [MPT24] for (the much more challenging) dynamical systems (with very heavy dependencies) of physical interest known as Lorentz gases. We note that stable LLD does not imply stable LD; even for i.i.d. random variables, stable LLD does not

[1]A stable LLD result is concerned with bounds on $\mathbb{P}(S_n - b_n \in (N-h, N+h))$, for $h > 0$, as $N/a_n \to +\infty$.

imply[2] Theorem 1.8 in the generality of (1.11) though, under ideal smoothness of the tail probabilities (1.11), this is in fact possible; see [Ber19, Theorem 2.4]. For deterministic dynamical systems, the transition from the stable LLD obtained in [MT22; MPT24] to a version of Theorem 1.8 seems impossible. As we shall soon explain, obtaining a version of Theorem 1.8 for dynamical systems is in many respects much more delicate than stable LLD. In this chapter, we restrict ourselves to the class of dynamical systems treated in [MT22].

Throughout the rest of the chapter, we let $N : \mathbb{N} \to (0, +\infty)$ be so that $N(n) \to +\infty$ as $n \to +\infty$ and $N(n)/a_n \to +\infty$ as $n \to +\infty$. We state the main result and refer to Theorem 4.11 in §4.3.2 for a more general version.

Theorem 4.1. *Let $(\Omega, \mathcal{A}, \mu, T)$ be a mixing Gibbs-Markov dynamical system. Suppose that $v : \Omega \to \mathbb{R}$ is a measurable observable with $\int_\Omega v^2 \, \mathrm{d}\mu = +\infty$ that satisfies (1.11) (with μ instead of $\mathbb{P}$). Suppose further that v is locally Lipschitz and satisfies a 'nice' technical condition (more precisely, (4.46) stated in §4.4.1). Write $v_n = \sum_{j=0}^{n-1} v \circ T^j$ for the ergodic sum of v, and assume that v_n satisfies (1.10) (with v_n instead of S_n) with a_n and b_n as in (1.9). Let $\delta > 0$ be arbitrarily small and set*

$$D(x) = \begin{cases} x^{-\alpha}(\log x)\ell(x), & \text{if } \alpha \in (0,1), \\ x^{-(\alpha-\delta)}, & \text{if } \alpha \in (1,2]. \end{cases}$$

Take any $N = N(n)$ such that $N/a_n \to +\infty$. If $\alpha \in (0,1) \cup (1,2)$, then

$$|\mu(v_n > N) - n\mu(v > N)| = \mathcal{O}\big(D(N)\big) + o\Big(\frac{n\ell(N)}{N^\alpha}\Big)$$

as $n \to +\infty$, and a similar statement holds for $\mu(v_n < -N)$. If $\alpha = 2$, then

$$|\mu(v_n > N) - n\mu(v > N)| = \mathcal{O}\big(D(N)\big) + \mathcal{O}\Big(\frac{n \log N}{N^2}\Big) + \mathcal{O}\Big(\frac{n\widehat{\ell}(N)}{N^2}\Big)$$

as $n \to +\infty$, where $\widehat{\ell}$ is as in (1.9), and a similar statement holds for $\mu(v_n < -N)$.

The corollary below gives the range of $N = N(n)$ for which $\mu(v_n > N) = n\mu(v > N)(1 + o(1))$ holds when $\alpha \in (1,2)$. In the case $\alpha = 2$ however, notice when $\ell \equiv 1$ that the asymptotic equality fails; see Remark 1.9. For $\alpha \in (1,2)$, the probability that v_n is larger than N is equivalent to the probability that there is exactly one jump larger than N, i.e. that there exists exactly one $j = 0, \ldots, n-1$ so that $v \circ T^j > N$.

Corollary 4.2. *Assume the setup of Theorem 4.11. Then $\mu(v_n > N) \sim n\mu(v > N)$ holds*

- *for $\alpha \in (0,1)$ and $N \in (a_n, e^{an})$, where a is small and independent of n*
- *for $\alpha \in (1,2)$ and $N \in (a_n, n^{1/\delta})$ for any $\delta > 0$*

Whilst we focus on observables taking values in $\mathbb{R}$, we believe, at the expense of cumbersome notation throughout, that Theorem 4.1 can be extended to $\mathbb{R}^d$-valued observables.

[2] As shown in [CD19; Ber19], under (1.11) with $\alpha \in (0,2)$ (see also [MT22] for a different proof and for the case $\alpha = 2$), for any $h > 0$, there exists $C > 0$ so that $\mathbb{P}(S_n - b_n \in [x-h, x+h]) \le C \frac{n}{a_n} \frac{\ell(|x|)}{1+|x|^\alpha}$, for all $n \in \mathbb{N}$ and for all $x \in \mathbb{R}$.

Although the proof below of Theorem 4.1 is quite technical, the main steps can be summarised as follows:

(i) Rephrase Theorem 1.8 for $\alpha \in (0,1) \cup (1,2)$ (and (1.13) for $\alpha = 2$) for i.i.d. sequences in terms of characteristic functions – see §4.3.1.

(ii) Decompose the Fourier transform (analogue of characteristic function) for the ergodic sum $v_n = \sum_{j=0}^{n-1} v \circ T^j$ into the characteristic function of an i.i.d. sequence and 'some good' quantities – see §4.3.3 and §4.3.4.

(iii) Estimate the various integrals appearing (from the 'good' quantities) in the analytic expression of $\mu(v_n - b_n > x)$ via 'modulus-of-continuity' type arguments (in the sense of [Kat76, Chapter 1]) – see §4.3.5.

(iv) Put (i) and (iii) together to conclude.

As an explicit application of Theorem 4.1 (and Corollary 4.2), we obtain large deviations for the ρ-Hölder mean of continued fraction expansions for all $\rho > 1$.

Recall that the continued fraction map $G : [0,1] \to [0,1]$ defined on the probability space $([0,1], \mathcal{B}([0,1]), \lambda)$ is given by $G(0) = 0$ and $G(x) = x^{-1} \mod 1$ otherwise; see Figure 4.1. In addition, recall that the continued fraction measure μ_G given by

$$\mu_G(A) = \frac{1}{\log 2} \int_A \frac{1}{1+x} \, \mathrm{d}\lambda(x), \qquad A \in \mathcal{B}([0,1])$$

is absolutely continuous with respect to the Lebesgue measure and is G-ergodic. It is known that each $y \in [0,1] \setminus \mathbb{Q}$ has a unique continued fraction expansion. For each $j \in \mathbb{N}$, let $d_j : [0,1] \setminus \mathbb{Q} \to \mathbb{N}$ be the function such that $d_j(y)$ is the j^{th} digit appearing in the continued fraction expansion of $y \in [0,1] \setminus \mathbb{Q}$.

Corollary 4.3. *Fix $\rho > 1$. Then for x in the range $x \in (n, e^{an})$ as $n \to +\infty$, where $a > 0$ is any positive constant,*

$$\mu_G\left(\left\{y \in [0,1] \setminus \mathbb{Q} : \left(\frac{d_1(y)^\rho + \cdots + d_n(y)^\rho}{n}\right)^{1/\rho} \geq x\right\}\right) = \frac{n^{1-1/\rho}}{x \log 2}\big(1 + o(1)\big).$$

4

Note that in order to obtain the result in Corollary 4.3 for the arithmetic (1-Hölder) mean, we would need Theorem 4.1 to hold for $\alpha = 1$; see Corollary 5.3 and Corollary 5.19 in the next chapter for related i.i.d. results in the case of $\alpha = 1$.

The proof of Corollary 4.3 is given in §4.4.2.

4.2 Discussion of Related Results

Before getting to the proof of Theorem 4.1, we will first review some stable large deviations results in the literature for sequences of non-i.i.d. random variables with no second moment and therefore no standard CLT.

4.2.1 [MW13]: Precise Large Deviations for Dependent Regularly Varying Sequences

The closest result to Theorem 1.8 in the literature for non-i.i.d. sequences is that of [MW13]. For sums $S_n := X_1 + \cdots + X_n$ in the domain of an α-stable distribution, where $(X_j)_{j\in\mathbb{N}}$ is a sequence of random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ and $\alpha \in (0,2]$, the authors give the precise asymptotics of $\mathbb{P}(S_n > x)$ as $x \to +\infty$ under certain conditions. As mentioned in Chapter 1, the asymptotics are well understood in the i.i.d. case and are given by Theorem 1.8 – the conclusion of which can be rewritten as

$$\lim_{n\to+\infty} \sup_{x\geq a_n} \left| \frac{\mathbb{P}(S_n - b_n > x)}{n\mathbb{P}(|X_1| > x)} - 1 \right| = 0 \quad \text{and} \quad \lim_{n\to+\infty} \sup_{x\geq a_n} \left| \frac{\mathbb{P}(S_n - b_n < -x)}{n\mathbb{P}(|X_1| > x)} - 1 \right| = 0.$$

The authors of [MW13] extend the precise LD given by Theorem 1.8 to dependent stationary sequences $(X_j)_{j\in\mathbb{N}}$ that satisfy the following conditions:

(MW1) (Regularly-varying-tail condition) X_j satisfies (1.11) for all $j \in \mathbb{N}$;

(MW2) (Anti-clustering condition) For each $n, k \in \mathbb{N}$, there exist constants $\delta_k = o(k^{-2})$ and sets $\Lambda_n \subset (0, +\infty)$ with $b_n := \inf \Lambda_n \to +\infty$ such that $n\mathbb{P}(|X| > b_n) \to 0$ and

$$\lim_{k\to+\infty} \limsup_{n\to+\infty} \sup_{x\in\Lambda_n} \delta_k^{-\alpha} \sum_{j=k}^{n} \mathbb{P}\big(|X_j| > x\delta_k \big| |X_0| > x\delta_k\big) = 0.$$

(MW3) The limit $b_+ := \lim_{k\to+\infty} \big(b_+(k+1) - b_+(k)\big)$ exists, where

$$b_+(k) := \lim_{x\to+\infty} \frac{\mathbb{P}(S_k > x)}{\mathbb{P}(|X_1| > x)} = \lim_{n\to+\infty} n\mathbb{P}(S_k > a_n), \qquad k \in \mathbb{N}.$$

(MW4) For any sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ satisfying $\varepsilon_k = o(k^{-1})$ and $(k+1)\delta_k \leq \varepsilon_k$,

$$\lim_{k\to+\infty} \limsup_{n\to+\infty} \sup_{x\in\Lambda_n} \frac{\mathbb{P}\big(\sum_{i=1}^n X_i \mathbb{1}_{\{|X_i|\leq\delta_k x\}} > \varepsilon_k x\big)}{n\mathbb{P}(|X| > x)} = 0.$$

Note that the quantities $b_+(k)$ and $b_-(k)$, $k \in \mathbb{N}$, in (MW3) exist by (MW1). In addition, the anti-clustering condition (MW2) ensures that long-range dependencies of extremes are avoided by preventing possible clusters of exceedances of high thresholds by (X_j) from being too large. Under these conditions, the authors obtain the following precise large deviation principle.

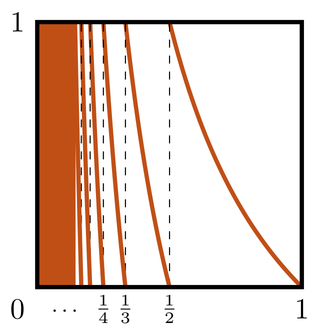


Figure 4.1: The continued fraction map

Theorem 4.4 ([MW13])**.** *Let* $(X_j)_{j\in\mathbb{N}}$ *be a dependent stationary sequence of random variables that satisfy* (MW1)–(MW4)*. Then*

$$\lim_{n\to+\infty}\sup_{x\in\Lambda_n}\left|\frac{\mathbb{P}(S_n>x)}{n\mathbb{P}(|X_1|>x)}-b_+\right|=0 \quad \text{and} \quad \lim_{n\to+\infty}\sup_{x\in\Lambda_n}\left|\frac{\mathbb{P}(S_n<-x)}{n\mathbb{P}(|X_1|>x)}-b_-\right|=0.$$

In particular, Theorem 4.4 gives

$$\mathbb{P}(S_n>x)=b_+n\mathbb{P}(|X_1|>x)\big(1+o(1)\big), \qquad x>b_n\to+\infty,$$

and a similar statement holds for $\mathbb{P}(S_n<-x)$. So, the range of x for which Theorem 4.4 holds is quite comparable to the range for which Theorem 1.8 holds.

We briefly discuss the conditions (MW1)–(MW4) and their applicability to dynamical systems. The condition (MW1), i.e. that the sequence $(X_j)_{j\in\mathbb{N}}$ of random variables satisfies (1.11), is a very natural condition to impose on dynamically defined random variables and is straightforward to verify. The anti-clustering condition (MW2) is commonly assumed when studying limit theory for extremes of dependent sequences and can be verified for dynamical systems that are, for example, strongly mixing. The condition (MW3) is greatly problematic as it is extremely difficult to verify, even for i.i.d. systems, without very strong additional assumptions on the system. The condition (MW4) is also known to hold for various dynamical systems; the authors give examples in [MW13, Remarks 3.2 and 3.3].

In summary, proving that the limit b_+ exists is a major step in proving precise LD, which motivates the question as to whether a precise LD result exists with conditions better catered to dynamically defined random variables.

4.2.2 [Gan96]: Large Deviations for a Heavy-tailed Mixing Sequence

In [Gan96], the author gives a form of large deviations for sequences of dynamically defined random variables $(Y_j)_{j\in\mathbb{N}}$ under the assumption that the dynamics satisfy a certain mixing condition. The type of result obtained is known as a *large deviation principle* (LDP), which, for suitably scaled partial sums Z_n of random variables,

- establishes the speed of convergence to zero as $n\to+\infty$ of the probabilities $\mathbb{P}(Z_n\in\cdot\,)$
- provides an explicit function describing the rate of decay

4

More precisely, an *LDP with rate* s_n *and rate function* I is the property of the probabilities $\mathbb{P}(Z_n\in\cdot\,)$, $n\in\mathbb{N}$, that

$$\limsup_{n\to+\infty}\frac{1}{s_n}\log\mathbb{P}(Z_n\in C)\le-\inf_{x\in C}I(x) \quad \text{for all closed sets } C\subset\mathbb{R},$$
$$\liminf_{n\to+\infty}\frac{1}{s_n}\log\mathbb{P}(Z_n\in O)\ge-\inf_{x\in O}I(x) \quad \text{for all open sets } O\subset\mathbb{R},$$

for some function $I:\mathbb{R}\to[0,+\infty]$ that satisfies the definition of a *rate function*: that I has compact level sets.

The specific mixing condition used involves the β-mixing coefficients $\beta(n)$, $n\in\mathbb{N}$ (see [Dou94]). These coefficients can be thought of as a measure of pairwise independence of the random variables in $(Y_j)_{j\in\mathbb{N}}$ with a value close to zero signifying near-independent behaviour. Assuming that the β-mixing coefficients decay to 0 as $n\to+\infty$ at a sufficiently fast rate allows one to ensure that $(Y_j)_{j\in\mathbb{N}}$ behaves on average like an i.i.d. sequence. Under this condition, the author obtains an LDP with speed $\log n$ of the empirical averages $(S_n/n)_{n\in\mathbb{N}}$.

Theorem 4.5 ([Gan96, Theorem 2.2]). *Fix a sequence $(Y_j)_{j\in\mathbb{N}}$ of random variables satisfying* (1.11) *with $\alpha \in (1,2)$. Assume that $\beta(n) \le n^{-\varphi(n)}$ for some function $\varphi : \mathbb{N} \to (0,+\infty)$ with $\varphi(n) \to +\infty$ as $n \to +\infty$. Then the probabilities $\mathbb{P}(S_n/n \in \cdot)$, $n \in \mathbb{N}$, satisfy an LDP with speed* $\log n$ *and rate function I given by*

$$I(y) := \begin{cases} \alpha - 1, & y > \mathbb{E}X_1, \\ 0, & y = \mathbb{E}X_1, \\ +\infty, & y < \mathbb{E}X_1. \end{cases}$$

Super-exponential decay of β-mixing coefficients is known to occur for a large class of dynamical systems, including Gibbs-Markov maps (see §4.4.1) such as the continued fraction map.

The loss of generality when comparing Theorem 4.5 with Theorem 1.8 is twofold. First, Theorem 4.5 only applies to sequences of α-stable random variables with $\alpha \in (1,2)$, i.e. with finite mean. Second, Theorem 4.5 can give the rate of decay of $\mathbb{P}(S_n \ge ny)$ only when y is independent of n, which is a consequence of having scaled S_n by n rather than a_n as directed by (1.10).

4.2.3 [Tak19]: Large Deviation Principle for Arithmetic Functions in Continued Fraction Expansions

In [Tak19], the author considers a dynamically defined sequence $(Y_j)_{j\in\mathbb{N}}$ of random variables on the probability space $([0,1], \mathcal{B}([0,1]), \lambda)$ given by $Y_j := (\psi \circ d_1) \circ G^{j-1} = \psi \circ d_j$, where $G : [0,1] \to [0,1]$ is the continued fraction map given in (1.4), ψ is any unbounded arithmetic function and $d_j : [0,1] \setminus \mathbb{Q} \to \mathbb{N}$ is the j^{th} continued fraction digit function given in §4.1. Then $\mathbb{E}Y_j = +\infty$ for each $j \in \mathbb{N}$ so we are in the α-stable setting with $\alpha \in (0,1)$. The author obtains the following result.

Theorem 4.6 ([Tak19, Theorem 1(b)]). *For every $a \in \mathbb{R}$,*

$$\lim_{n\to+\infty} \frac{1}{n} \log \lambda\left(\left\{x \in [0,1] \setminus \mathbb{Q} : \frac{1}{n}\sum_{j=1}^{n} \psi \circ d_j(x) \ge a\right\}\right) = 0.$$

Due to the speed n and scaling n having been chosen, Theorem 4.6 tells us that the rate of decay of the arithmetic mean of continued fraction digits is at most sub-exponential but does not give any further information. It is an intriguing question as to whether one can deduce a more quantitative result with the same setup.

4

4.3 Proving Theorem 4.1

4.3.1 Rephrasing Theorem 1.8 in Terms of Characteristic Functions

As mentioned in the summary of the proof of Theorem 4.1 at the end of §4.1, the starting point of the proof is to rephrase the i.i.d. case Theorem 1.8 in terms of characteristic functions. We start with a general lemma that is valid whenever $N(n) \to +\infty$ as $n \to +\infty$.

Lemma 4.7. *Let $(Z_n)_{n\in\mathbb{N}}$ be a sequence of random variables, and let Y be another random variable defined on the same probability space but independent of all the Z_n. Assume that there exist two sequences $N = N(n)$ and $h_N = h_{N(n)}$ such that $\mathbb{P}(|Y| > h_N) = o(\mathbb{P}(Z_n > N))$ and $\mathbb{P}(Z_n > N \pm h_N) = \mathbb{P}(Z_n > N)(1 + o(1))$ as $n \to +\infty$. Then for $\widetilde{Z}_n := Z_n + Y$, we have*

$$\mathbb{P}(\widetilde{Z}_n > N) = \mathbb{P}(Z_n > N)(1 + o(1)) \quad \text{as } n \to +\infty.$$

Proof. Since Z_n and Y are independent, we have (for $h = h_{N(n)}$) that

$$\begin{aligned}\mathbb{P}(Z_n > N + h) &= \mathbb{P}(Z_n > N + h,\, |Y| > h) + \mathbb{P}(Z_n > N + h,\, |Y| \leq h)\\ &\leq \mathbb{P}(|Y| > h) + \mathbb{P}(\widetilde{Z}_n > N).\end{aligned}$$

Therefore

$$\mathbb{P}(\widetilde{Z}_n > N) \geq \mathbb{P}(Z_n > N + h) - \mathbb{P}(|Y| > h) = \mathbb{P}(Z_n > N)\big(1 + o(1)\big).$$

For the other inequality, we have

$$\begin{aligned}\mathbb{P}(Z_n \leq N - h) &= \mathbb{P}(Z_n \leq N - h,\, |Y| > h) + \mathbb{P}(Z_n \leq N - h,\, |Y| \leq h)\\ &\leq \mathbb{P}(|Y| > h) + \mathbb{P}(\widetilde{Z}_n \leq N).\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{P}(\widetilde{Z}_n > N) &= 1 - \mathbb{P}(\widetilde{Z}_n \leq N) \leq 1 - \mathbb{P}(Z_n \leq N - h) + \mathbb{P}(|Y| > h)\\ &= \mathbb{P}(Z_n > N - h) + \mathbb{P}(|Y| > h) = \mathbb{P}(Z_n > N)\big(1 + o(1)\big).\end{aligned}$$

This finishes the proof. □

For $t \in \mathbb{R}$, let $\Psi(t) := \mathbb{E}(e^{itX_1})$ be the characteristic function of X_1. In this section, we translate Theorem 1.8 for $\alpha \in (0,1) \cup (1,2]$ into a statement on the characteristic function $\Psi(t)$. Such a translation is captured in equation (4.1) below, which is the starting point for the dynamic setting. Since $\alpha \neq 1$, we can assume with no loss of generality that $b_n = 0$; for $\alpha \in (0,1)$, this is given, while, for $\alpha \in (1,2]$, we simply replace X_1 by $X_1 - \mathbb{E}(X_1)$.

Proposition 4.8. *Assume the setup of Theorem 1.8 with $\alpha \neq 1$. Let Y be an $\mathcal{L}^2$ random variable, independent of S_n (for each n), with real-valued, even and $\mathcal{C}^2$ characteristic function Ψ_Y, supported in $[-\varepsilon, \varepsilon]$ for some*[3] *small $\varepsilon > 0$. Take any $N = N(n)$ such that $a_n = o(N(n))$ and let $g(n)$ be so that $nN = o(g(n))$. Then*

(i) *If $\alpha \in (0,1) \cup (1,2)$, then the first asymptotic equivalence in (1.12) holds if and only if*

$$\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{\Psi(t)^n - n\Psi(t)}{it}\,\mathrm{d}t\right| = o(n\mathbb{P}(X_1 > N)) \tag{4.1}$$

as $n \to +\infty$.

(ii) *If $\alpha = 2$, then (1.13) holds if and only if*

$$\begin{aligned}&\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{\Psi(t)^n - n\Psi(t)}{it}\,\mathrm{d}t - \Phi\left(\frac{N}{a_n}\right)\right|\\ &\qquad = o\left(n\mathbb{P}(X_1 > N) + \Phi\left(\frac{N}{a_n}\right)\right)\end{aligned}$$

as $n \to +\infty$.

[3]The existence of such a random variable Y on the same probability space and with characteristic function Ψ_Y follows from, for instance, [Gou10a, Proposition 3.8], [KT22, Proof of Theorem 3] and [MT22, Footnote 1].

4

Remark 4.9. (a) The second asymptotic equivalence in (1.12) can be dealt with in the same way by using the fact that $\mathbb{P}(S_n < -N) = \mathbb{P}(-S_n > N)$.

(b) If the random variables X_i are $\mathbb{Z}$-valued (so S_n is $\mathbb{Z}$-valued), then one could exploit a simpler formula, bypassing the presence of Ψ_Y in the statement of Proposition 4.8. Indeed,

$$\mathbb{P}(S_n > N) = \sum_{j>N} \mathbb{P}(S_n = j) = \sum_{j>N} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itj} \Psi(t)^n \, \mathrm{d}t.$$

Proof. To start with, recall the inversion formulae

$$\mathbb{P}(X_1 \in (N, N+g(n)]) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi(t) \, \mathrm{d}t$$

and

$$\mathbb{P}(S_n \in (N, N+g(n)]) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi(t)^n \, \mathrm{d}t,$$

for any choice of $g(n)$. We want to allow that $N(n), g(n) \to +\infty$ and thus obtain the desired formula for $\mathbb{P}(S_n > N)$. Recall that $nN = o(g(n))$, and compute that

$$\mathbb{P}(X_1 > N + g(n)) = p\ell(N+g(n)) \frac{1}{g(n)^\alpha} \frac{1+o(1)}{(1+N/g(n))^\alpha} = o\left(\frac{n\ell(N+g(n))}{g(n)^\alpha}\right) = o(\mathbb{P}(X_1 > N)).$$

Since $N = o\big((N+g(n))/n\big)$, it follows that

$$\mathbb{P}(S_n > N + g(n)) \leq \sum_{j=0}^{n-1} \mathbb{P}\left(|X_j| > \frac{N+g(n)}{n}\right) = o(n\mathbb{P}(X_1 > N)).$$

Thus, given the choice of $g(n)$, we have $\mathbb{P}(X_1 \in (N, N+g(n)]) = \mathbb{P}(X_1 > N)(1+o(1))$ and

$$\begin{aligned} \mathbb{P}(S_n \in (N, N+g(n)]) &= \mathbb{P}(S_n > N) - \mathbb{P}(S_n > N + g(n)) \\ &= \mathbb{P}(S_n > N) + o(n\mathbb{P}(X_1 > N)). \end{aligned}$$

Therefore, as $n \to +\infty$,

$$\mathbb{P}(X_1 > N) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi(t) \, \mathrm{d}t + o(\mathbb{P}(X_1 > N)) \tag{4.2}$$

and

$$\mathbb{P}(S_n > N) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi(t)^n \, \mathrm{d}t + o(n\mathbb{P}(X_1 > N)). \tag{4.3}$$

Next, we argue that one can adjust the domain of integration in (4.3). Put $\widetilde{S}_n := S_n + Y$, where Y is as in the statement, that is, an $\mathcal{L}^2$ random variable, independent of S_n, with real-valued, even and $\mathcal{C}^2$ Fourier transform Ψ_Y, supported in $[-\varepsilon, \varepsilon]$ for some small $\varepsilon > 0$. The analogues of (4.2) and (4.3) are:

$$\mathbb{P}(X_1 + Y > N) = \frac{1+o(1)}{2\pi} \int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi_Y(t)\Psi(t) \, \mathrm{d}t \tag{4.4}$$

and

$$\mathbb{P}(S_n + Y > N) = \frac{1}{2\pi} \int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi_Y(t)\Psi(t)^n \, \mathrm{d}t + o\left(n\mathbb{P}\left(X_1 + \frac{Y}{n} > N\right)\right). \tag{4.5}$$

4

To compare the tails of X_1 and $X_1 + Y$, and of S_n and $\widetilde{S}_n = S_n + Y$, we will use Lemma 4.7 several times, making different choices for Z_n appearing in that lemma.

<u>The case $\alpha \in (0,1) \cup (1,2)$</u>

Take $\delta \in (0, 2-\alpha)$ and $h_N = N^{\frac{\alpha+\delta}{2}} = o(N)$ so that $\mathbb{P}(X_1 > N \pm h_N) = \mathbb{P}(X_1 > N)(1 + o(1))$. Since Y is $\mathcal{L}^2$, we have $\mathbb{P}(|Y| > h_N) = o(h_N^{-2}) = o(N^{-\alpha-\delta}) = o(N^{-\alpha}\ell(N)) = o(\mathbb{P}(X_1 > N))$. Therefore Lemma 4.7 for $Z_n \equiv X_1$ gives

$$\mathbb{P}(X_1 > N) = \mathbb{P}(X_1+Y > N)(1+o(1)) = \frac{1+o(1)}{2\pi}\int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it}\Psi_Y(t)\Psi(t)\,\mathrm{d}t. \quad (4.6)$$

Next we want to obtain a version with S_n in (4.5).

From (1.12) **to** (4.1). If S_n satisfies (1.12), i.e. $\mathbb{P}(S_n > N) = n\mathbb{P}(X_1 > N)(1 + o(1))$, then the previous computation and choice of h_N give also $\mathbb{P}(S_n > N + h_N) = \mathbb{P}(S_n > N)(1 + o(1))$ and $\mathbb{P}(|Y| > h_N) = o(\mathbb{P}(X_1 > N)) = o(\mathbb{P}(S_n > N))$. Therefore, Lemma 4.7 for $Z_n = S_n$ and $\widetilde{S}_n = S_n + Y$ give

$$\mathbb{P}(S_n > N) = \mathbb{P}(\widetilde{S}_n > N)(1 + o(1)).$$

From (4.1) **to** (1.12). Note that $\mathbb{P}(|Y|/n > h_N) < \mathbb{P}(|Y| > h_N) = o(\mathbb{P}(X_1 > N))$. If (4.1) holds, then using (4.5) and (4.6) together with (4.4) yields $\mathbb{P}(\widetilde{S}_n > N) = n\mathbb{P}(X_1 > N)(1+o(1))$. Again, by using Lemma 4.7 with $Z_n = \widetilde{S}_n$, $S_n = \widetilde{S}_n - Y$ and $h_N = N^{\frac{\alpha+\delta}{2}} = o(N)$, we may conclude that

$$\mathbb{P}(\widetilde{S}_n > N) = \mathbb{P}(S_n > N)(1 + o(1)). \quad (4.7)$$

Hence, after subtracting (4.4) from (4.5) and apply (4.6) and (4.7), we obtain

$$\begin{aligned}
&|\mathbb{P}(S_n > N) - n\mathbb{P}(X_1 > N)| \\
&\quad = \frac{1+o(1)}{2\pi}\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{\Psi(t)^n - n\Psi(t)}{it}\,\mathrm{d}t\right| + o(n\mathbb{P}(X_1 > N)) \\
&\quad = o(n\mathbb{P}(X_1 > N)).
\end{aligned}$$

as $n, N \to +\infty$ so that $a_n = o(N)$.

<u>The case $\alpha = 2$</u>

Let $m(N)$ be so that $\mathbb{P}(|Y| > N) = N^{-2}m(N)$ (so $m(N) = o(\ell(N))$). The choice of $m(N)$ ensures that there is a function $\widetilde{m}(N) \to +\infty$ such that also $\widetilde{m}(N)^2 m(N/\widetilde{m}(N)) = o(\ell(N))$. Now take $h_N = N/\widetilde{m}(N) = o(N)$. Then $\mathbb{P}(X_1 > N + h_N) = \mathbb{P}(X_1 > N)(1 + o(1))$. Moreover

$$\mathbb{P}\left(\frac{|Y|}{n} > h_N\right) < \mathbb{P}(|Y| > h_N) \ll \frac{\widetilde{m}(N)^2 m(N/\widetilde{m}(N))}{N^2} = o\left(\frac{\ell(N)}{N^2}\right) = o(\mathbb{P}(X_1 > N)).$$

By the same argument used in obtaining (4.6), $\mathbb{P}(X_1 + Y > N) = \mathbb{P}(X_1 > N)(1 + o(1))$ and

$$\mathbb{P}(X_1 > N) = \frac{1+o(1)}{2\pi}\int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it}\,\Psi_Y(t)\Psi(t)\,\mathrm{d}t. \quad (4.8)$$

The same choice of h_N gives $\mathbb{P}(S_n > N + h_N) = \mathbb{P}(S_n > N)(1 + o(1))$ and $\mathbb{P}(|Y| > h_N) = o(\mathbb{P}(X_1 > N) = o(\mathbb{P}(S_n > N)$. Therefore Lemma 4.7 gives $\mathbb{P}(\widetilde{S}_n > N) = \mathbb{P}(S_n > N)(1 + o(1))$. Combining these estimates with (4.5) gives

$$\mathbb{P}(S_n > N) = \frac{1+o(1)}{2\pi}\int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it}\Psi_Y(t)\Psi(t)^n\,\mathrm{d}t + o(n\mathbb{P}(X_1 > N)). \quad (4.9)$$

Now, we insert the tail (1.13), subtract $\Phi(N/a_n)$ and (4.8) n times. This gives

$$\left|\int_{-\varepsilon}^{+\varepsilon}(e^{-itN}-e^{-it(N+g(n))})\Psi_Y(t)\frac{\Psi(t)^n-n\Psi(t)}{it}\,\mathrm{d}t-\Phi\left(\frac{N}{a_n}\right)\right|=o\left(n\mathbb{P}(X_1>N)+\Phi\left(\frac{N}{a_n}\right)\right),$$

as required. For the converse, (1.13) follows directly from the above formula together with (4.8) and (4.9). □

For use below in the setup of dependent random variables (arising in the context of dynamical systems) we record the following consequence of the proof of Proposition 4.8.

Proposition 4.10. *Let $(v^j)_{j\geq 0}$ be a sequence of random variables that are not necessarily independent but are identically distributed on some probability space (Ω,μ). Suppose that v^j satisfies* (1.11) *(with v^j instead of X_j and μ instead of $\mathbb{P}$) and write $v_n=\sum_{j=0}^{n-1}v^j$. Assume that v_n satisfies* (1.10) *with v_n instead of S_n and a_n and b_n as in* (1.9). *Let Y be an $\mathcal{L}^2$ random variable, independent of v_n (for each n), with real-valued, even and $\mathcal{C}^2$ Fourier transform Ψ_Y, supported in $[-\varepsilon,\varepsilon]$ for some[4] small $\varepsilon>0$. Take any $N=N(n)$ such that $a_n=o(N(n))$ and let $g(n)=N(n)^{1+\varepsilon_0}$ with $\varepsilon_0>0$. Moreover, suppose that there exists $W(N,n)$ so that $n\mathbb{P}(X_1>N)=\mathcal{O}(W(N,n))$ and so that*

(i) *If $\alpha\in(0,1)\cup(1,2)$, as $n\to+\infty$,*

$$\left|\int_{-\varepsilon}^{+\varepsilon}(e^{-itN}-e^{-it(N+g(n))})\Psi_Y(t)\frac{\mathbb{E}_\mu(e^{itv_n})-n\mathbb{E}_\mu(e^{itv^0})}{it}\,\mathrm{d}t\right|=\mathcal{O}(W(N,n)).$$

(ii) *If $\alpha=2$, as $n\to+\infty$,*

$$\left|\int_{-\varepsilon}^{+\varepsilon}(e^{-itN}-e^{-it(N+g(n))})\Psi_Y(t)\frac{\mathbb{E}_\mu(e^{itv_n})-n\mathbb{E}_\mu(e^{itv^0})}{it}\,\mathrm{d}t-\Phi\left(\frac{N}{a_n}\right)\right|$$
$$=\mathcal{O}(W(N,n))+o\left(\Phi\left(\frac{N}{a_n}\right)\right).$$

Then

(i) *If $\alpha\in(0,1)\cup(1,2)$, as $n\to+\infty$,*

$$\left|\mu(v_n>N)-n\mu(v^0>N)\right|=\mathcal{O}(W(N,n)).$$

(ii) *If $\alpha=2$, as $n\to+\infty$,*

$$\left|\mu(v_n>N)-n\mu(v^0>N)\right|=\mathcal{O}(W(N,n))+o\left(\Phi\left(\frac{N}{a_n}\right)\right).$$

Proof. First, note that the only place where we required independence inside the proof of Proposition 4.8 was where we translated the content of Theorem 1.8 with $\alpha\neq 1$ in terms of $\left|\int_{-\varepsilon}^{+\varepsilon}(e^{-itN}-e^{-it(N+g(n))})\frac{\Psi(t)^n-n\Psi(t)}{it}\Psi_Y(t)\,\mathrm{d}t\right|$. As explained below, the same proof (for the converse part) with $\Psi(t)^n$ replaced by $\mathbb{E}_\mu(e^{itv_n})$ yields the conclusion of the current proposition.

[4]In the dynamical setting, ε will be fixed in Fact 4.14.

First, by the same argument as the one used in obtaining (4.2) and (4.3), we have

$$\mu(v_n > N) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \mathbb{E}_\mu(e^{itv_n})\,\mathrm{d}t + o\big(n\mu(v^0 > N)\big)$$

and

$$\mu(v^0 > N) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \mathbb{E}_\mu(e^{itv^0})\,\mathrm{d}t.$$

Also, analogously to (4.4) and (4.5),

$$\mu(v^0 + Y > N) = \frac{1}{2\pi} \int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi_Y(t)\mathbb{E}_\mu(e^{itv^0})\,\mathrm{d}t,$$

$$\mu(v_n + Y > N) = \frac{1}{2\pi} \int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi_Y(t)\mathbb{E}_\mu(e^{itv_n})\,\mathrm{d}t + o\big(n\mu(v^0 > N)\big)$$

and

$$\begin{aligned} &\big|\mu(v_n + Y > N) - n\mu(v^0 > N)\big| - o\big(n\mu(v^0 > N)\big) \\ &\quad= \begin{cases} O(W(N,n)), & \alpha \in (0,1)\cup(1,2), \\ O(W(N,n)) + o\Big(\Phi\Big(\frac{N}{a_n}\Big)\Big) & \alpha = 2. \end{cases} \end{aligned}$$

It remains to argue that $\mu(v_n + Y > N) = \mu(v_n > N)(1 + o(1))$. If $\alpha \in (0,1)\cup(1,2)$, we proceed as in obtaining (4.7), by applying Lemma 4.7 with $Z_n = \widetilde{S}_n$, $S_n = \widetilde{S}_n - Y$ and $h_N = N^{\frac{\alpha+\delta}{2}} = o(N)$. If $\alpha = 2$, we proceed as in the argument used in obtaining (4.9) with the same choice of h_N as there (and S_n replaced by v_n). □

4.3.2 Abstract Setup for Dynamical Systems

Let $T : \Omega \to \Omega$ be a measure-preserving map on a probability space (Ω,μ). Let $v : \Omega \to \mathbb{R}$ be a measurable observable with $\int_\Omega v^2\,\mathrm{d}\mu = +\infty$. We fix $\alpha \in (0,1)\cup(1,2]$ throughout and assume

(H1) $\mu(v > x) = px^{-\alpha}\ell(x)(1+o(1)), \qquad \mu(v \le -x) = qx^{-\alpha}\ell(x)(1+o(1))$

as $x \to +\infty$, where ℓ is slowly varying and $p + q = 1$. Recall ℓ and $\widehat{\ell}$ from (1.9). Throughout, we shall work with

$$\ell_0 := \begin{cases} \ell, & \alpha \in (0,1)\cup(1,2), \\ \widehat{\ell}, & \alpha = 2. \end{cases}$$

Let $R : \mathcal{L}^1 \to \mathcal{L}^1$ be the transfer operator for T defined via $\int_\Omega Rf\,g\,\mathrm{d}\mu = \int_\Omega f\,g\circ T\,\mathrm{d}\mu$. Given $t \in \mathbb{R}$, define the perturbed operator $R(t) : \mathcal{L}^1 \to \mathcal{L}^1$ by $R(t)f = R(e^{itv}f)$. We assume that there is a Banach space $\mathcal{B} \subset \mathcal{L}^\infty$ containing the constant functions, and with norm $\|\cdot\|$ satisfying $|\varphi|_\infty \le \|\varphi\|$ for $\varphi \in \mathcal{B}$, such that

(H2) There exist $\varepsilon > 0$, $C > 0$ such that for all $|t|, |h| \in B_\varepsilon(0)$,

- (i) $\|R(t)\| \le C$ for $\alpha \in (0,1)\cup(1,2]$ and $\|R'(t)\| \le C$ for $\alpha \in (1,2]$.
- (ii) If $\alpha \in (0,1)$, then $\|R(t+h) - R(t)\| \le C|h|^\alpha \ell(1/|h|)$.
- (iii) If $\alpha \in (1,2]$, then $\|R(t+h) - R(t) - ihR'(0)\| \le C|h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$ for any $\alpha' \in (1,\alpha)$.
- (iv) If $\alpha \in (1,2]$, then $\|R'(t+h) - R'(t)\| \le C|h|^{\alpha-1}\ell_0(1/|h|)$.

Item (iii) is maybe not as natural as an estimate on $\|R(t+h) - R(t) - ihR'(t)\|$, but it is what is required in the proofs (it is crucial for the proofs of items (ii) and (iii) of Proposition 4.20), and its validity is checked in §4.4.

Since $R(0) = R$ and $\mathcal{B}$ contains constant functions, $\mathbb{1}$ is an eigenvalue of $R(0)$. We assume the following:

(H3) The eigenvalue $\mathbb{1}$ is simple, and the remainder of the spectrum of $R(0) : \mathcal{B} \to \mathcal{B}$ is contained in a disk of radius less than 1.

Throughout, write $v_n = \sum_{j=0}^{n-1} v \circ T^j$. It is known (see [AD01]) that, under hypotheses (H1)–(H3), distributional convergence $(v_n - b_n)/a_n \to_d Y_\alpha$ to an α-stable random variable Y_α holds, where a_n and b_n are defined as in (1.9). With no loss of generality, when $\alpha \in (1, 2]$, we assume that $\int_\Omega v \, \mathrm{d}\mu = 0$. Hence, $b_n = 0$ for $\alpha \in (0, 1) \cup (1, 2]$. With this specified, we state the main result in the abstract setup.

Theorem 4.11. *Suppose that* (H1)–(H3) *hold. Let $\delta > 0$ be arbitrarily small and set*

$$D(x) = \begin{cases} x^{-\alpha}\ell(x)\log x, & \text{if } \alpha \in (0,1), \\ x^{-(\alpha-\delta)}, & \text{if } \alpha \in (1,2]. \end{cases}$$

Take $N = N(n)$ such that $N/a_n \to +\infty$. Then, as $n \to +\infty$,

(i) *If $\alpha \in (0, 1) \cup (1, 2)$, then*

$$|\mu(v_n > N) - n\mu(v > N)| = \mathcal{O}\big(D(N)\big) + o\left(\frac{n\ell(N)}{N^\alpha}\right),$$

and a similar statement holds for $\mu(v_n \leq -N)$.

(ii) *If $\alpha = 2$, then*

$$|\mu(v_n > N) - n\mu(v > N)| = \mathcal{O}\big(D(N)\big) + \mathcal{O}\left(\frac{n\log N}{N^2}\right) + \mathcal{O}\left(\frac{n\widehat{\ell}(N)}{N^2}\right),$$

and a similar statement holds for $\mu(v_n \leq -N)$.

4

4.3.3 Strategy of the Proof

In the proof below of Theorem 4.11, we shall proceed by reasoning as in §4.3.1 and prove the statement on the right tail. More precisely, as in §4.3.1, we shall work with the inversion formula for distribution functions. As recalled in §4.3.4 below, the relevant characteristic function $\mathbb{E}_\mu(e^{itv_n})$ of the partial sum of dependent quantities $\{v \circ T^j\}_{j\geq 0}$ decomposes into the characteristic function of the partial sum of i.i.d. random variables $\{\hat{v}_j\}_{j\geq 0}$, which we refer to as $\Psi(t)^n$ (see §4.3.4), and 'some other' quantities. As recalled in §4.3.4, $\mathbb{E}_\mu(e^{itv}) = \mathbb{E}_\mu(e^{it\hat{v}_j}) = \Psi(t)$ for all t. This together with a fact recorded inside the proof of Proposition 4.8, namely (4.6) (with $X_j = \hat{v}_j$), ensures that, given $\varepsilon > 0$ small enough,

$$\mu(v > N)(1 + o(1)) = \frac{1}{2\pi}\lim_{n\to+\infty}\int_{-\varepsilon}^{+\varepsilon}\frac{e^{-itN} - e^{-it(N+g(n))}}{it}\Psi_Y(t)\Psi(t)\,\mathrm{d}t, \tag{4.10}$$

where Ψ_Y is as in the statement of Proposition 4.8. By Proposition 4.10 (with $v^j = v \circ T^j$) and (4.10), the conclusion of Theorem 4.11 follows once we show for some $\varepsilon > 0$ that

$$\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{\mathbb{E}_\mu(e^{itv_n}) - n\Psi(t)}{it}\,\mathrm{d}t\right| = \mathcal{O}\big(D(N)\big) + o\Big(\frac{n\ell_0(N)}{N^\alpha}\Big) \tag{4.11}$$

when $\alpha \in (0,1) \cup (1,2)$ and

$$\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{\mathbb{E}_\mu(e^{itv_n}) - n\Psi(t)}{it}\,\mathrm{d}t\right| = \mathcal{O}\left(D(N) + \frac{n\log N + n\widehat{\ell}(N)}{N^2}\right) \tag{4.12}$$

when $\alpha = 2$, using that $\Phi(N/a_n) \ll (N/a_n)^{-2} \ll nN^{-2}\widehat{\ell}(N)$.

If we write $\mathbb{E}_\mu(e^{itv_n}) = \Psi(t)^n + (\mathbb{E}_\mu(e^{itv_n}) - \Psi(t)^n)$ and use the translation recorded in Proposition 4.8, equations (4.11) and (4.12) follow as soon as we show that

$$\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{\mathbb{E}_\mu(e^{itv_n}) - \Psi(t)^n}{it}\,\mathrm{d}t\right| = \mathcal{O}\big(D(N)\big) + o\Big(\frac{n\ell_0(N)}{N^\alpha}\Big) \tag{4.13}$$

when $\alpha \in (0,1) \cup (1,2)$ and

$$\begin{aligned}\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{\mathbb{E}_\mu(e^{itv_n}) - \Psi(t)^n}{it}\,\mathrm{d}t - \Phi\Big(\frac{N}{a_n}\Big)\right|\\ = \mathcal{O}\left(D(N) + \frac{n\log N + n\widehat{\ell}(N)}{N^2}\right)\end{aligned} \tag{4.14}$$

when $\alpha = 2$.

Since $\mathbb{E}_\mu(e^{itv_n}) = \int_\Omega R(t)^n\, \mathbb{1}\, \mathrm{d}\mu$, we take the next section to decompose $\int_\Omega R(t)^n \mathbb{1}\, \mathrm{d}\mu$ into $\Psi(t)^n$ and 'good' quantities. The idea is to isolate $\Psi(t)^n$ in the expression of $\int_\Omega R(t)^n \mathbb{1}\, \mathrm{d}\mu$ so as to be able to use the equivalent of (4.1). The meaning of 'good' quantities will become clear in §4.3.5 below, where we complete the proof of Theorem 4.11. At this stage, we can point out that the proof uses a 'modulus-of-continuity' argument (and also a derivative in the case $\alpha \in (1,2]$). Although this type of argument has some similarity with the ones used in [MT22], the current arguments are much more delicate. To carry out a modulus-of-continuity argument, it is crucial that we obtain enough decay in t in the expression of $\int_\Omega R(t)^n \mathbb{1}\, \mathrm{d}\mu - \Psi(t)^n$; this is necessary to counteract the effect of the division by t in the integrand in (4.13). The details are postponed to §4.3.5.

4.3.4 Decomposing $\int_\Omega R(t)^n \mathbb{1}\, \mathrm{d}\mu$ into $\Psi(t)^n$ and 'Good' Quantities.

The main result of this section is Proposition 4.20 and is stated at the end of the present section. We first recall some general facts under (H1)–(H3), some of which were recently clarified in [MT22, §3.2].

By (H2) and (H3), there exists $\varepsilon > 0$ and a continuous family $\lambda(t)$ of simple eigenvalues of $R(t)$ for $|t| \le 3\varepsilon$ with $\lambda(0) = \mathbb{1}$. The associated spectral projections $P(t)$, $|t| \le 3\varepsilon$, form a continuous family of bounded linear operators on $\mathcal{B}$. Moreover, there is a continuous family of linear operators $Q(t)$ on $\mathcal{B}$ and constants $C > 0$, $\delta_0 \in (0,1)$ such that for $|t| \le 3\varepsilon$

$$R(t) = \lambda(t)P(t) + Q(t) \qquad \text{and} \qquad \|Q(t)^n\| \le C\delta_0^n. \tag{4.15}$$

Let $\zeta(t) := \frac{P(t)\mathbb{1}}{\int P(t)\mathbb{1}\,\mathrm{d}\mu}$ be the normalised eigenvector corresponding to $\lambda(t)$.

Lemma 4.12. *Assume* (H2). *Then there exists $\varepsilon > 0$ such that the properties of $R(t)$ listed in* (H2) *are inherited by $P(t)$, $Q(t)$, $\lambda(t)$ and $\zeta(t)$ for all $|t|$, $|t+h| \le 3\varepsilon$.*

Lemma 4.12 holds with a simplified version of (H2)(iii) obtained by moving $ihR'(0)$ to the other side of the $\ll$-sign); namely, for $\alpha \in (1,2]$,

$$\|R(t+h) - R(t)\| \ll |h|. \tag{4.16}$$

The full strength of (H2)(iii) is exploited in Lemma 4.13 below.

A consequence of (H2)(iii) for the operators P and Q

Lemma 4.13. *Let $\alpha \in (1,2]$. Then there exists $C > 0$ so that the following hold for all $|t|, |t+h| \in B_{3\varepsilon}(0)$ and for any $\alpha' \in (1,\alpha)$.*

(i) $\|P(t+h) - P(t) - ihP'(0)\| \le C|h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$, *and a similar statement holds for* $\zeta(t) := P(t)\mathbb{1}/\int P(t)\mathbb{1}\,\mathrm{d}\mu$.

(ii) $\|Q(t+h) - Q(t) - ihQ'(0)\| \le C|h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$.

Proof. (i) Recall that $\lambda(t)$ is an isolated eigenvalue in the spectrum of $R(t)$ for every $|t| \le 3\varepsilon$. Hence, for any $\delta \in (0, 3\varepsilon)$, we have $P(t) = (2\pi i)^{-1}\int_{|\xi-1|=\delta}(\xi - R(t))^{-1}\,\mathrm{d}\xi$ and so it follows that $P'(t) = (2\pi i)^{-1}\int_{|\xi-1|=\delta}(\xi - R(t))^{-1}R'(t)(\xi-R(t))^{-1}\,\mathrm{d}\xi$. With these specified we see that

$$\begin{aligned}
2\pi i\big(P(t+h) - P(t)\big) &= \int_{|\xi-1|=\delta} \big(\xi - R(t+h)\big)^{-1}\big(R(t+h) - R(t)\big)\big(\xi - R(t)\big)^{-1}\,\mathrm{d}\xi \\
&= \int_{|\xi-1|=\delta} \big(\xi - R(t+h)\big)^{-1}\big(R(t+h) - R(t) - ihR'(0)\big)\big(\xi - R(t)\big)^{-1}\,\mathrm{d}\xi \\
&\quad + ih\int_{|\xi-1|=\delta} \big(\xi - R(t+h)\big)^{-1}R'(0)\big(\xi - R(t)\big)^{-1}\,\mathrm{d}\xi \\
&=: I(t,h) + J(t,h).
\end{aligned}$$

By (H2)(iii), $\|I(t,h)\| \ll |h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$, for any $\alpha' \in (1,\alpha)$. Also,

$$\begin{aligned}
J(t,h) &= ih\int_{|\xi-1|=\delta} \big(\xi - R(0)\big)^{-1}R'(0)\big(\xi - R(t)\big)^{-1}\,\mathrm{d}\xi \\
&\quad + ih\int_{|\xi-1|=\delta} \Big(\big(\xi - R(t+h)\big)^{-1} - \big(\xi - R(0)\big)^{-1}\Big)R'(0)\big(\xi - R(t)\big)^{-1}d\xi \\
&= ih\int_{|\xi-1|=\delta} \big(\xi - R(0)\big)^{-1}R'(0)\big(\xi - R(0)\big)^{-1}\,\mathrm{d}\xi \\
&\quad + ih\int_{|\xi-1|=\delta} \Big(\big(\xi - R(t+h)\big)^{-1} - \big(\xi - R(0)\big)^{-1}\Big)R'(0)\big(\xi - R(t)\big)^{-1}\,\mathrm{d}\xi \\
&\quad + ih\int_{|\xi-1|=\delta} \big(\xi - R(0)\big)^{-1}R'(0)\Big(\big(\xi - R(t)\big)^{-1} - \big(\xi - R(0)\big)^{-1}\Big)d\xi \\
&= ih(2\pi i)P'(0) + ih(J_1(t,h) + J_2(t,h)).
\end{aligned}$$

By (4.16), i.e. the simplified version of (H2)(iii), we have

$$\|R(t+h) - R(0)\| \le \|R(t) - R(0)\| + \|R(t+h) - R(t)\| \ll |t| + |h|$$

4

and thus $\|J_1(t,h)\|$, $\|J_2(t,h)\| \ll |h|(|t|+|h|)$. Altogether, we have found that

$$P(t+h) - P(t) - ihP'(0) = (2\pi i)^{-1} I(t,h) + (2\pi i)^{-1} h(J_1(t,h) + J_2(t,h)),$$

and that each term satisfies the claimed estimate, which concludes the proof of the statement on $P(t)$. The statement on $\zeta(t)$ follows directly from the definition and the statement on $P(t)$.

(ii) We shall use that $Q(t) = R(t)(I - P(t))$, where I denotes the identity operator. So, $Q'(t) = (R(t)(I-P(t)))' = R'(t)(I - P(t)) - R(t)P'(t)$ and thus

$$Q'(0) = R'(0) - R(0)P'(0) - R'(0)P(0). \tag{4.17}$$

With this specified, we compute that

$$\begin{aligned} Q(t+h) - Q(t) &= R(t+h) - R(t) + R(t)P(t) - R(t+h)P(t+h) \\ &= R(t+h) - R(t) + R(t)\big(P(t) - P(t+h)\big) - \big(R(t+h) - R(t)\big)P(t+h) \\ &= \big(R(t+h) - R(t) - ihR'(0)\big) - R(t)\big(P(t+h) - P(t) - ihP'(0)\big) \\ &\quad - \big(R(t+h) - R(t) - ihR'(0)\big)P(t+h) \\ &\quad + ihR'(0) - ihR(t)P'(0) - ihR'(0)P(t+h). \end{aligned}$$

Using (H2)(iii) and item (i) above, we have

$$\big\|Q(t+h) - Q(t) - ih\big(R'(0) - R(t)P'(0) + R'(0)P(t+h)\big)\big\| \ll |h|(|t^{\alpha'-1} + |h|^{\alpha'-1})$$

for any $\alpha' \in (1,\alpha)$. By (H2)(iii) and (4.17),

$$\begin{aligned} ih\big(R'(0) - R(t)P'(0) - R'(0)P(t+h)\big) &= ih\big(R'(0) - R(0)P'(0) - R'(0)P(0)\big) + D(t,h) \\ &= ihQ'(0) + D(t,h), \end{aligned}$$

where $\|D(t,h)\| \ll |h|(|t|+|h|)$. □

Some properties of λ and details on Ψ

Let $\Psi(t)$ be the characteristic function of the i.i.d. random variables $\hat{v}_j : \Omega \to \mathbb{R}$ so that $\hat{v}_j$ and $v \circ T^j$ have the same distribution. Similarly to v, we take $\int_\Omega \hat{v}_j \, d\mu = 0$ (when $\alpha > 1$). Since $\hat{v}_j$ satisfies (H1),

$$1 - \Psi(t) = c_\alpha |t|^\alpha \ell_0\left(\frac{1}{|t|}\right)\big(1 + o(1)\big), \qquad t \to 0. \tag{4.18}$$

Equation (4.18) has been established in [GK49; IL65] (see also [AD01, Theorem 5.1]). Also, as established in these works (in particular, [AD01]) and clarified in [MT22, Lemma 2.1],

Fact 4.14. *There exist constants ε, $c > 0$, such that*

$$|\Psi(t)| \le \exp\left(-c|t|^\alpha \ell_0\left(\frac{1}{|t|}\right)\right), \qquad \forall t \in B_{3\varepsilon}(0).$$

Throughout the rest of the chapter, we fix ε so that both Fact 4.14 and equation (4.15) hold. It is known that there exists $C > 0$ so that for all $|t|, |t+h| \in B_{3\varepsilon}(0)$,

$$|\Psi(t+h) - \Psi(t)| \le \begin{cases} C|h|^\alpha \ell(1/|h|), & \alpha \in (0,1), \\ C|h|, & \alpha \in (1,2]. \end{cases} \tag{4.19}$$

4

Furthermore, if $\alpha \in (1,2]$, then Ψ is differentiable and, writing Ψ' for its derivative,

$$|\Psi'(t)| \le C\ell_0\left(\frac{1}{|t|}\right)|t|^{\alpha-1} \tag{4.20}$$

for some $C > 0$. For a precise reference for the validity of (4.19) and (4.20), see for instance [MT22, Lemma 2.2].

Lemma 4.15. *Let $V(t) = \lambda(t) - \Psi(t)$. Then there exist C, C' so that the following hold for all $|t|, |t+h| \in B_{3\varepsilon}(0)$:*

(i) *If $\alpha \in (0,1)$, then $|V(t)| \le C|t|^{2\alpha}\ell(1/|t|)^2$, $|V(t+h) - V(t)| \le C'|h|^\alpha \ell(1/|h|)|t|^\alpha \ell(1/|t|)$. If $\alpha \in (1,2]$, then $|V(t)| \le C|t|^2$ and $|V(t+h) - V(t)| \le C'|h|(|t| + |h|)$.*

(ii) *If $\alpha \in (1,2]$, then V is differentiable and, writing V' for its derivative, $|V'(t)| \le C|t|$ and $|V'(t+h) - V'(t)| \le C|h|^{\alpha'-1}\ell_0(\frac{1}{|h|})(|t| + |h|) + C'|h|$ for any $\alpha' \in (1,\alpha)$.*

Before proceeding with the proof, we note that Lemma 4.15 together with (4.18) and (4.20) implies that

$$1 - \lambda(t) = c_\alpha \ell_0\left(\frac{1}{|t|}\right)|t|^\alpha(1 + o(1)) \quad \text{and, for } \alpha \in (1,2], \quad |\lambda'(t)| \le C\ell_0\left(\frac{1}{|t|}\right)|t|^{\alpha-1}, \tag{4.21}$$

for some $C > 0$.

Proof of Lemma 4.15. Write

$$\begin{aligned}\lambda(t) &= \int R(t)\zeta(t)\,\mathrm{d}\mu = \int_\Omega R(t)\mathbb{1}\,\mathrm{d}\mu + \int \big(R(t) - R(0)\big)\big(\zeta(t) - \zeta(0)\big)\,\mathrm{d}\mu \\ &= \int_\Omega e^{itv}\,\mathrm{d}\mu + \int_\Omega \big(R(t) - R(0)\big)\big(\zeta(t) - \zeta(0)\big)\,\mathrm{d}\mu = \Psi(t) + V(t).\end{aligned}$$

This gives the (well-known) decomposition of λ into Ψ and V in the statement of the lemma.

4

We still need to clarify the properties of V. Since $\mathcal{B} \subset \mathcal{L}^\infty$, item (i) for the case $\alpha \in (0,1)$ follows immediately from (H2)(ii) and Lemma 4.12. The second part of item (i) for the case $\alpha \in (1,2]$ follows similarly using that, by (4.16), $\|R(t+h) - R(0)\| \ll |t| + |h|$.

For item (ii), we first note that

$$V'(t) = \int_\Omega \big(R(t) - R(0)\big)\zeta'(t)\,\mathrm{d}\mu + \int_\Omega R'(t)\,\big(\zeta(t) - \zeta(0)\big)\,\mathrm{d}\mu.$$

Lemma 4.12 and Fact 4.14 together with $\|R(t) - R(0)\| \ll |t|$ imply that $|V'(t)| \ll |t|$. Next, compute that

$$\begin{aligned}V'(t+h) - V'(t) &= \int_\Omega \big(R(t+h) - R(t)\big)\zeta'(t+h)\,\mathrm{d}\mu + \int_\Omega \big(R(t) - R(0)\big)(\zeta'(t+h) - \zeta'(t))\,\mathrm{d}\mu \\ &\quad + \int_\Omega (R'(t+h) - R'(t))\,\big(\zeta(t) - \zeta(0)\big)\,\mathrm{d}\mu + \int_\Omega R'(t+h)\,(\zeta(t+h) - \zeta(t))\,\mathrm{d}\mu \\ &=: I_1(t,h) + I_2(t,h) + I_3(t,h) + I_4(t,h).\end{aligned}$$

Recall $\|R'(t)\|, |\zeta'(t)| \ll 1$ and $\|R(t+h) - R(0)\| \ll |t| + |h|$. These together with (H2)(iv) and Lemma 4.12 give that $|I_2(t,h)|, |I_4(t,h)| \ll |h|^{\alpha-1}\ell_0(1/|h|)(|t|+|h|)$. Regarding $I_1(t,h)$, compute that

$$\begin{aligned} I_1(t,h) &= ih\int_\Omega R'(0)\zeta'(t+h)\,\mathrm{d}\mu + \int_\Omega (R(t+h) - R(t) - ihR'(0))\zeta'(t+h)\,\mathrm{d}\mu \\ &= ih\int_\Omega R'(0)\zeta'(0)\,\mathrm{d}\mu + ih\int_\Omega R'(0)(\zeta'(t+h) - \zeta'(0))\,\mathrm{d}\mu \\ &\quad + \int_\Omega (R(t+h) - R(t) - ihR'(0))\zeta'(t+h)\,\mathrm{d}\mu. \end{aligned}$$

Similarly,

$$\begin{aligned} I_4(t,h) &= ih\int_\Omega R'(0)\zeta'(0)\,\mathrm{d}\mu + ih\int_\Omega (R'(t+h) - R'(0))\zeta'(0)\,\mathrm{d}\mu \\ &\quad + \int_\Omega R'(t+h)\,(\zeta(t+h) - \zeta(t) - ih\zeta'(0))\,\mathrm{d}\mu. \end{aligned}$$

By (H2)(iii), (H2)(iv) and Lemma 4.12,

$$|I_1(t,h) - ih\int_\Omega R'(0)\zeta'(0)\,\mathrm{d}\mu| \ll |h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$$

for any $\alpha' \in (1,\alpha)$. By (H2)(iv) (together with Lemma 4.12) and Lemma 4.13(i),

$$|I_4(t,h) - ih\int_\Omega R'(0)\zeta'(0)\,\mathrm{d}\mu| \ll |h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$$

for any $\alpha' \in (1,\alpha)$. Hence,

$$|I_1(t,h) + I_4(t,h)| \ll |h| + |h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$$

for any $\alpha' \in (1,\alpha)$, ending the proof. □

For use in the proofs to follow, we recall one more property of λ when $\alpha \in (1,2]$; this is a refined version of the continuity properties recorded in Lemma 4.12 (this type of estimate appears inside [MT22, Proof of Theorem 3.2]).

Lemma 4.16. *If $\alpha \in (1,2]$, then there exists $C > 0$ so that, for all $t, t+h \in B_{3\varepsilon}(0)$ with $|h| \le \frac{|t|}{2}$,*

$$|\lambda(t+h) - \lambda(t)| \le C|h|(|t|^{\alpha-1} + |h|^{\alpha-1})\ell_0\left(\frac{1}{|t|+|h|}\right),$$

and a similar statement holds for Ψ.

Proof. By the mean value theorem, $\lambda(t+h) - \lambda(t) = h\lambda'(t^*)$ for some $t^* \in [t, t+h]$ and the conclusion follows. □

4

A first decomposition of $\int_\Omega R(t)^n \mathbb{1}\,\mathrm{d}\mu$ into $\Psi(t)^n$ and 'good' quantities

Lemma 4.17. *Put*

$$\widehat{V}(t,n) := \int_\Omega R(t)^n \mathbb{1}\,\mathrm{d}\mu - \Psi(t)^n - \int_\Omega Q(t)^n \mathbb{1}\,\mathrm{d}\mu.$$

Then the following hold for all $n \in \mathbb{N}$ and all $|t|, |h| \in B_\varepsilon(0)$ with $|h| \le |t|/2$.

(i) *If* $\alpha \in (0,1)$, *then*

$$\left|\widehat{V}(t,n)\right| \ll \left(|t|^{\alpha}\ell\left(\frac{1}{|t|}\right) + n|t|^{2\alpha}\ell\left(\frac{1}{|t|}\right)^{2}\right)|\Psi(t)|^{n}$$

$$\left|\widehat{V}(t+h,n) - \widehat{V}(t,n)\right| \ll \left(|h|^{\alpha}\ell\left(\frac{1}{|h|}\right) + n|h|^{\alpha}\ell\left(\frac{1}{|h|}\right)|t|^{\alpha}\ell\left(\frac{1}{|t|}\right)\right)|\Psi(t)|^{n}.$$

(ii) *If* $\alpha \in (1,2]$, *then for any* $\alpha' \in (1,\alpha)$,

$$\left|\widehat{V}(t,n) - it\lambda(t)^{n}\int_{\Omega} P'(0)\mathbb{1}\,\mathrm{d}\mu\right| \ll \left(|t|^{\alpha'} + n|t|^{2}\right)|\Psi(t)|^{n}.$$

(iii) *If* $\alpha \in (1,2]$, *then for any* $\alpha' \in (1,\alpha)$,

$$\left|\widehat{V}(t+h,n) - \widehat{V}(t,n) - ih\lambda(t)^{n}\int_{\Omega} P'(0)\mathbb{1}\,\mathrm{d}\mu\right| \ll \left(|h||t|^{\alpha'-1} + n|h||t|^{\alpha-1}\ell_{0}\left(\frac{1}{|t|}\right)\right)|\Psi(t)|^{n}.$$

(iv) *If* $\alpha \in (1,2]$, *then* $\widehat{V}$ *is differentiable in* t *and, writing* $\widehat{V}'(t,n)$ *for its derivative,*

$$\left|\widehat{V}'(t,n) - \lambda(t)^{n}\int_{\Omega} P'(0)\mathbb{1}\,\mathrm{d}\mu\right| \le \left(|t|^{\alpha-1}\ell_{0}\left(\frac{1}{|t|}\right) + n|t| + n^{2}|t|^{\alpha+1}\ell_{0}\left(\frac{1}{|t|}\right)\right)|\Psi(t)|^{n}$$

and

$$\begin{aligned}
&\left|\widehat{V}'(t+h,n) - \widehat{V}'(t,n) - ih \in \mathbb{N}\lambda(t)^{n-1}\lambda'(t)\int_{\Omega} P'(0)\mathbb{1}\,\mathrm{d}\mu\right| \\
&\quad \ll |h|^{\alpha-1}\ell_{0}\left(\frac{1}{|h|}\right)|\Psi(t)|^{n} + n\left(|t||h|^{\alpha'-1}\ell_{0}\left(\frac{1}{|h|}\right) + |h|\right)|\Psi(t)|^{n} \\
&\qquad + n^{2}\left(|t|^{2\alpha-1}|h|\ell_{0}\left(\frac{1}{|t|}\right)^{2} + |t|^{2}|h|^{\alpha-1}\ell_{0}\left(\frac{1}{|h|}\right) + |h||t|^{\alpha}\ell_{0}\left(\frac{1}{|t|}\right)\right)|\Psi(t)|^{n} \\
&\qquad + n^{3}|t|^{2\alpha}|h|\ell_{0}\left(\frac{1}{|t|}\right)^{2}|\Psi(t)|^{n}.
\end{aligned}$$

4

Proof. By (4.15),

$$\begin{aligned}
\int_{\Omega} R(t)^{n}\mathbb{1}\,\mathrm{d}\mu - \int_{\Omega} Q(t)^{n}\mathbb{1}\,\mathrm{d}\mu &= \lambda(t)^{n}\int_{\Omega} P(t)\mathbb{1}\,\mathrm{d}\mu \\
&= \lambda(t)^{n}\int_{\Omega} P(0)\mathbb{1}\,\mathrm{d}\mu + \lambda(t)^{n}\int_{\Omega}\big(P(t)-P(0)\big)\mathbb{1}\,\mathrm{d}\mu \\
&= \Psi(t)^{n} + (\lambda(t)^{n} - \Psi(t)^{n}) + \lambda(t)^{n}\int_{\Omega}\big(P(t)-P(0)\big)\mathbb{1}\,\mathrm{d}\mu.
\end{aligned}$$

Hence,

$$\widehat{V}(t,n) = \big(\lambda(t)^{n} - \Psi(t)^{n}\big) + \lambda(t)^{n}\int_{\Omega}\big(P(t)-P(0)\big)\mathbb{1}\,\mathrm{d}\mu =: A(t,n) + B(t,n). \tag{4.22}$$

Throughout the rest of the proof, we will use that, by Lemma 4.15, $|\lambda(t)^n| \ll |\Psi(t)|^n$.

(i) Note by Lemma 4.12 and (H2)(ii) that $|B(t,n)| \ll |t|^\alpha \ell(1/|t|)|\Psi(t)|^n$. Next, using (H2)(ii) and Lemma 4.12,

$$|B(t+h,n) - B(t,n)| \leq |h|^\alpha \ell\Big(\frac{1}{|h|}\Big)|\Psi(t)|^n + |\lambda(t+h)^n - \lambda(t)^n||t|^\alpha \ell\Big(\frac{1}{|t|}\Big).$$

Note that $\lambda(t+h)^n - \lambda(t)^n = (\lambda(t+h) - \lambda(t))\sum_{j=0}^{n-1}\lambda(t+h)^j\lambda(t)^{n-j}$. This together with Lemma 4.12 implies that $|\lambda(t+h)^n - \lambda(t)^n| \ll n|h|^\alpha \ell(1/|h|)|\Psi(t)|^n$. Hence,

$$|B(t+h,n) - B(t,n)| \leq |h|^\alpha \ell\Big(\frac{1}{|h|}\Big)\Big(n|t|^\alpha \ell\Big(\frac{1}{|t|}\Big) + 1\Big)|\Psi(t)|^n.$$

It remains to deal with $A(t,n)$ in (4.22). Given $V(t) = \lambda(t) - \Psi(t)$ as in Lemma 4.15, we have $A(t,n) = V(t)\sum_{j=0}^{n-1}\lambda(t)^j\Psi(t)^{n-j-1}$. This together with the properties of V in Lemma 4.15(i) implies that $|A(t,n)| \ll n|t|^{2\alpha}\ell(1/|t|)^2|\Psi(t)|^n$. To deal with the continuity properties of $A(t,n)$, we first compute that

$$\begin{aligned}
&\big(\lambda(t+h)^n - \Psi(t+h)^n\big) - \big(\lambda(t)^n - \Psi(t)^n\big)\\
&\quad= V(t+h)\sum_{j=0}^{n-1}\lambda(t+h)^j\Psi(t+h)^{n-j-1} - V(t)\sum_{j=0}^{n-1}\lambda(t)^j\Psi(t)^{n-j-1}\\
&\quad= \big(V(t+h) - V(t)\big)\sum_{j=0}^{n-1}\lambda(t+h)^j\Psi(t+h)^{n-j-1}\\
&\qquad+ V(t)\sum_{j=0}^{n-1}\big(\lambda(t+h)^j - \lambda(t)^j\big)\Psi(t+h)^{n-j-1}\\
&\qquad+ V(t)\sum_{j=0}^{n-1}\lambda(t)^j\big(\Psi(t+h)^{n-j-1} - \Psi(t)^{n-j-1}\big)
\end{aligned} \tag{4.23}$$

Using (4.23), we compute that $|A(t+h,n) - A(t,n)| \ll I_1 + I_2 + I_3$, where

$$I_1 = |V(t+h) - V(t)|\sum_{j=0}^{n-1}|\lambda(t+h)^j|\,|\Psi(t+h)^{n-j-1}| \ll n|h|^\alpha \ell\Big(\frac{1}{|h|}\Big)|t|^\alpha \ell\Big(\frac{1}{|t|}\Big)|\Psi(t+h)^n|,$$

$$I_2 = |V(t)|\sum_{j=0}^{n-1}|\lambda(t+h)^j - \lambda(t)^j|\,|\Psi(t+h)^{n-j-1}| \ll n|h|^\alpha \ell\Big(\frac{1}{|h|}\Big)|t|^{2\alpha}\ell\Big(\frac{1}{|t|}\Big)^2|\Psi(t+h)^n|,$$

using Lemma 4.12 and Lemma 4.15 again, and

$$I_3 = |V(t)|\sum_{j=0}^{n-1}|\lambda(t)^j|\,|\Psi(t+h)^{n-j-1} - \Psi(t)^{n-j-1}| \ll n|h|^\alpha \ell\Big(\frac{1}{|h|}\Big)|t|^{2\alpha}\ell\Big(\frac{1}{|t|}\Big)^2|\Psi(t)|^n,$$

where we have used (4.19) as well.

(ii) We continue from (4.22). First, $|A(t,n)| \ll n|V(t)||\Psi(t)|^n$, and, using Lemma 4.15(i) (the statement on V for the case $\alpha \in (1,2]$), $|A(t,n)| \ll n|t|^2|\Psi(t)|^n$. Next, we rewrite $B(t,n)$ in a convenient way:

$$B(t,n) = it\lambda(t)^n\int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu + \lambda(t)^n\int_\Omega\big(P(t) - P(0) - itP'(0)\big)\mathbb{1}\,\mathrm{d}\mu.$$

4

By Lemma 4.13(i), $\left|\lambda(t)^n \int_\Omega \big(P(t) - P(0) - itP'(0)\big)\mathbb{1}\,\mathrm{d}\mu\right| \ll |t|^{\alpha'}|\Psi(t)|^n$ for any $\alpha' \in (1,\alpha)$, and the conclusion follows.

(iii) We start from (4.22). First,

$$\begin{aligned}|A(t+h,n) - A(t,n)| &\ll |\lambda(t+h)^n - \lambda(t)^n| + |\Psi(t+h)^n - \Psi(t)^n| \\ &\ll n|\lambda(t+h) - \lambda(t)|\,|\lambda(t)^n| + n|\Psi(t+h) - \Psi(t)|\,|\Psi(t)|^n \\ &\ll n|h||t|^{\alpha-1}\,\ell_0\Big(\frac{1}{|t|}\Big)|\Psi(t)|^n,\end{aligned}$$

where we have used Lemma 4.16. It remains to deal with B defined in (4.22). Observe that

$$\begin{aligned}B(t+h,n) - B(t,n) &= \lambda(t)^n \int_\Omega \big(P(t+h) - P(t)\big)\mathbb{1}\,\mathrm{d}\mu \\ &\quad + (\lambda(t+h)^n - \lambda(t)^n) \int_\Omega \big(P(t+h) - P(0)\big)\mathbb{1}\,\mathrm{d}\mu \\ &=: J_1(t,h,n) + J_2(t,h,n).\end{aligned}$$

First, by Lemma 4.12 and Lemma 4.16,

$$|J_2(t,h,n)| \ll n|\lambda(t+h) - \lambda(t)|\,|\lambda(t)^n||t+h| \ll n\,|h||t|^\alpha\,\ell_0\Big(\frac{1}{|t|}\Big)|\Psi(t)|^n.$$

Now

$$J_1(t,h,n) = ih\lambda(t)^n \int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu + \lambda(t)^n \int_\Omega \big(P(t+h) - P(t) - ihP'(0)\big)\mathbb{1}\,\mathrm{d}\mu.$$

Using Lemma 4.13(i), we obtain that for any $\alpha' \in (1,\alpha)$,

$$\left|J_1(t,h,n) - ih\lambda(t)^n \int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu\right| \ll |h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})\,|\Psi(t)|^n,$$

ending the proof of (iii).

(iv) Differentiating in (4.22),

$$\begin{aligned}\widehat{V}'(t,n) &= n\big(\lambda(t)^{n-1}\lambda'(t) - \Psi(t)^{n-1}\Psi'(t)\big) + n\lambda(t)^{n-1}\lambda'(t)\int_\Omega \big(P(t) - P(0)\big)\mathbb{1}\,\mathrm{d}\mu \\ &\quad + \lambda(t)^n \int_\Omega P'(t)\mathbb{1}\,\mathrm{d}\mu \\ &=: A_0(t,n) + A_1(t,n) + A_2(t,n). \qquad (4.24)\end{aligned}$$

Recalling from Lemma 4.15 that $V(t) = \lambda(t) - \Psi(t)$, we compute that

$$\begin{aligned}A_0(t,n) &= n\lambda'(t)\big(\lambda(t)^{n-1} - \Psi(t)^{n-1}\big) + n\Psi(t)^{n-1}(\lambda'(t) - \Psi'(t)) \\ &= n\lambda'(t)\big(\lambda(t)^{n-1} - \Psi(t)^{n-1}\big) + n\Psi(t)^{n-1}V'(t). \qquad (4.25)\end{aligned}$$

It follows from equation (4.21) and Lemma 4.15(i) and (ii) that

$$\begin{aligned}|A_0(t,n) &\ll n|t|^{\alpha-1}\ell_0\Big(\frac{1}{|t|}\Big)n|V(t)|\,|\Psi(t)|^n + n|V'(t)|\,|\Psi(t)|^n \\ &\ll n^2|t|^{\alpha+1}\ell_0\Big(\frac{1}{|t|}\Big)|\Psi(t)|^n + n|t|\,|\Psi(t)|^n.\end{aligned}$$

Regarding A_1, using equation (4.21) (the statement on the derivative) and Lemma 4.12, we obtain $|A_1(t,n)| \ll n|t|^\alpha \ell_0(1/|t|)|\Psi(t)|^n$. For A_2, write

$$\begin{aligned} A_2(t,n) &= \lambda(t)^n \int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu + \lambda(t)^n \int_\Omega \big(P'(t)-P'(0)\big)\mathbb{1}\,\mathrm{d}\mu \\ &=: \lambda(t)^n \int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu + A_2^1(t,n). \end{aligned} \tag{4.26}$$

By Lemma 4.12 and (H2)(iv), $|A_2^1(t,n)| \ll |t|^{\alpha-1}\ell_0(1/|t|)|\Psi(t)|^n$. Thus,

$$\left|A_2(t,n) - \lambda(t)^n \int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu\right| \ll |t|^{\alpha-1}\ell_0\left(\frac{1}{|t|}\right)|\Psi(t)|^n.$$

Putting together the estimates for A_0, A_1, A_2,

$$\begin{aligned} \left|\widehat{V}'(t,n) - \lambda(t)^n \int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu\right| &\ll \left(n^2|t|^{\alpha+1}\ell_0\left(\frac{1}{|t|}\right) + n|t|\right)|\Psi(t)|^n + n|t|^\alpha\ell_0\left(\frac{1}{|t|}\right)|\Psi(t)|^n \\ &\quad + |t|^{\alpha-1}\ell_0\left(\frac{1}{|t|}\right)|\Psi(t)|^n \\ &\ll \left(|t|^{\alpha-1}\ell_0\left(\frac{1}{|t|}\right) + n|t| + n^2|t|^{\alpha+1}\ell_0\left(\frac{1}{|t|}\right)\right)|\Psi(t)|^n, \end{aligned}$$

as claimed.

We continue with the continuity properties of $\widehat{V}'$ studying each term in (4.22).

Term A_0 in (4.24)**.** Using the expression (4.25), we compute that

$$\begin{aligned} |A_0(t+h,n) - A_0(t,n)| &\ll n^2|\lambda'(t+h)-\lambda'(t)|\,|\Psi(t)^{n-1}|\,|V(t)| \\ &\quad + n|\lambda'(t)|\big|\big(\lambda(t+h)^{n-1} - \Psi(t+h)^{n-1}\big) - \big(\lambda(t)^{n-1} - \Psi(t)^{n-1}\big)\big| \\ &\quad + n\big|\Psi(t+h)^{n-1} - \Psi(t)^{n-1}\big||V'(t)| + n|\Psi(t)^{n-1}||V'(t+h)-V'(t)| \\ &=: I_1(t,h,n) + I_2(t,h,n) + I_3(t,h,n) + I_4(t,h,n). \end{aligned}$$

By Lemma 4.15(ii),

$$I_4(t,h,n) \ll n|h|^{\alpha'-1}\ell_0\left(\frac{1}{|h|}\right)(|t|+|h|)|\Psi(t)^{n-1}| + n|h||\Psi(t)^{n-1}|.$$

By Lemma 4.16 (the statement on Ψ), Lemma 4.15(ii) and (4.20),

$$\begin{aligned} I_3(t,h,n) &\ll n^2|h||t|\ell_0\left(\frac{1}{|t|+|h|}\right)(|t|^{\alpha-1} + |h|^{\alpha-1})\,|\Psi(t)|^n \\ &\ll n^2|h||t|\ell_0\left(\frac{1}{|t|+|h|}\right)|t|^{\alpha-1}\,|\Psi(t)|^n. \end{aligned}$$

We continue with I_2. First, recalling equation (4.20) and using Lemma 4.15 and Lemma 4.16,

$$\begin{aligned} &\big|\big(\lambda(t+h)^n - \Psi(t+h)^n\big) - \big(\lambda(t)^n - \Psi(t)^n\big)\big| \\ &\quad \ll n\big|V(t+h)-V(t)\big|\,|\Psi(t)^{n-1}| + n^2|V(t)|\,\big|\lambda(t+h)-\lambda(t)\big|\,|\Psi(t)^{n-1}| \\ &\qquad + n^2|V(t)|\,\big|\Psi(t+h)-\Psi(t)\big|\,|\Psi(t)^{n-1}| \\ &\quad \ll n|h|(|t|+|h|)\,|\Psi(t)^{n-1}| + n^2|h||t|^2\ell_0\left(\frac{1}{|t|+|h|}\right)\big(|t|^{\alpha-1}+|h|^{\alpha-1}\big)|\Psi(t)^{n-1}|. \end{aligned} \tag{4.27}$$

4

By equations (4.27) and (4.21),

$$I_2(t,h,n) \ll n^2|h||t|^\alpha\, \ell_0\Big(\frac{1}{|t|+|h|}\Big)|\Psi(t)^{n-1}| + n^3|h||t|^{2\alpha}\ell_0\Big(\frac{1}{|t|+|h|}\Big)^2 |\Psi(t)^{n-1}|.$$

Finally, by Lemma 4.12 and (H2)(iv), $|\lambda'(t+h)-\lambda'(t)| \ll h^{\alpha-1}\ell_0(1/|h|)$. This together with Lemma 4.15(i) gives

$$I_1(t,h,n) \ll n^2|h|^{\alpha-1}|t|^2\ell_0\Big(\frac{1}{|t|+|h|}\Big)|\Psi(t)|^n.$$

Putting together the estimates for I_4, I_3, I_2, I_1, we obtain

$$\begin{aligned}|A_0(t+h,n)-A_0(t,n)| \ll \Big(n|h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big)(|t|+|h|)+n|h|\Big)|\Psi(t)^{n-1}| \\ + \Big(n^2|h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big)|t|^2 + n^2|h||t|^\alpha\ell_0\Big(\frac{1}{|t|+|h|}\Big)\Big)|\Psi(t)^{n-1}| \\ + n^3|h||t|^{2\alpha}\,\ell_0\Big(\frac{1}{|t|+|h|}\Big)^2 |\Psi(t)^{n-1}|.\end{aligned} \tag{4.28}$$

Term A_1 in (4.24). We compute that

$$\begin{aligned}A_1(t+h,n)-A_1(t,n) &= n\lambda(t)^{n-1}\lambda'(t)\int_\Omega \big(P(t+h)-P(t)\big)\mathbb{1}\,\mathrm{d}\mu \\ &\quad + n(\lambda(t+h)^{n-1}-\lambda(t)^{n-1})\lambda'(t)\int_\Omega \big(P(t+h)-P(0)\big)\mathbb{1}\,\mathrm{d}\mu \\ &\quad + n\lambda(t+h)^{n-1}(\lambda'(t+h)-\lambda'(t))\int_\Omega \big(P(t+h)-P(0)\big)\mathbb{1}\,\mathrm{d}\mu \\ &=: J_1(t,h,n)+J_2(t,h,n)+J_3(t,h,n).\end{aligned}$$

By Lemma 4.12 and arguments similar to the ones used in estimating A_0 above,

$$|J_2(t,h,n)| \ll n^2|h||t|^{2\alpha-1}\ell_0\Big(\frac{1}{|t|+|h|}\Big)^2|\Psi(t)|^n \text{ and } |J_3(t,h,n)| \ll n|h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big)|t|\,|\Psi(t)|^n.$$

4

Finally, for any $\alpha' \in (1,\alpha)$,

$$\begin{aligned}J_1(t,h,n) &= ihn\,\lambda(t)^{n-1}\lambda'(t)\int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu + n\lambda(t)^{n-1}\lambda'(t)\int_\Omega\big(P(t+h)-P(t)-ihP'(0)\big)\mathbb{1}\,\mathrm{d}\mu \\ &= ihn\,\lambda(t)^{n-1}\lambda'(t)\int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu + \mathcal{O}\Big(n|h||t|^{\alpha'+\alpha-2}\,|\Psi(t)|^n\Big),\end{aligned}$$

where in the last line we have used equation (4.21) and Lemma 4.13(i).

Altogether,

$$\begin{aligned}&\Big|A_1(t+h,n)-A_1(t,n)-ihn\,\lambda(t)^{n-1}\lambda'(t)\int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu\Big| \\ &\ll \Big(n|h||t|^{\alpha'+\alpha-2}+n^2|h||t|^{2\alpha-1}\ell_0\Big(\frac{1}{|t|+|h|}\Big)^2 + n|h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big)|t|\Big)|\Psi(t)|^n.\end{aligned} \tag{4.29}$$

Term A_2 in (4.24)**.** Here we use the definition of A_2^1 in (4.26). By Lemma 4.12 and Lemma 4.16,

$$|A_2^1(t+h,n) - A_2^1(t,n)| \ll |\lambda(t+h)^n - \lambda(t)^n| \, \|P'(t)\| + |\lambda(t)^n| \, \|P'(t+h) - P'(t)\| \\ \ll n|h||t|^{\alpha-1}\ell_0\Big(\frac{1}{|t|+|h|}\Big) |\Psi(t)|^n + |h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big)|\Psi(t)|^n.$$

The second statement of item (iv) follows by putting together the estimate on A_2^1 with the equations (4.28) and (4.29). More precisely, putting all these together (recorded in the order A_0, A_1 and A_2) and recalling (4.29),

$$\Big|\widehat{V}'(t+h,n) - \widehat{V}'(t,n) - ihn\,\lambda(t)^{n-1}\lambda'(t)\int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu\Big| \\ \ll |h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big)|\Psi(t)|^n + n\bigg(|t||h|^{\alpha'-1}\ell_0\Big(\frac{1}{|h|}\Big) + |h| + |t|^{\alpha'+\alpha-2}|h|\ell_0\Big(\frac{1}{|t|}\Big)\bigg)|\Psi(t)|^n \\ + n^2\bigg(|t|^{2\alpha-1}|h|\ell_0\Big(\frac{1}{|t|+|h|}\Big)^2 + |t|^2|h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big) + |h||t|^\alpha\ell_0\Big(\frac{1}{|t|+|h|}\Big)\bigg)|\Psi(t)|^n \\ + n^3|t|^{2\alpha}|h|\ell_0\Big(\frac{1}{|t|+|h|}\Big)^2|\Psi(t)|^n.$$

Using that $|h| \le |t|/2$, we can ignore the term $|t|^{\alpha'+\alpha-2}|h|\ell_0(1/|t|)$. More precisely, we see that for any $\varepsilon > 0$,

$$|t|^{\alpha'+\alpha-2}|h|\ell_0\Big(\frac{1}{|t|}\Big) = |t|\,|t|^{\alpha'+\alpha-3}\,|h|^{2-\alpha-\varepsilon}\ell_0\Big(\frac{1}{|t|}\Big)|h|^{\alpha-1+\varepsilon} \ll |t|\,|t|^{\alpha'-1-\varepsilon}\ell_0\Big(\frac{1}{|t|}\Big)|h|^{\alpha-1+\varepsilon}$$

Since $\alpha' > 1$ and $\varepsilon > 0$ is arbitrarily small, $|t|^{\alpha'-1-\varepsilon}\ell_0(1/|t|) = o(1)$ as $|t| \to 0$. Also, we have $|h|^{\alpha-1+\varepsilon} \ll |h|^{\alpha-1}\ell_0(1/|h|)$. Thus,

$$|t|^{\alpha'+\alpha-2}|h|\ell_0\Big(\frac{1}{|t|}\Big) \ll |t||h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big) \ll |t||h|^{\alpha'-1}\ell_0\Big(\frac{1}{|h|}\Big).$$

This completes the proof of (iv). □

Properties of $Q(t)^n$

4

Lemma 4.18. *For $n \ge 1$, consider the operator $W(t,n) = Q(t)^n - Q(0)^n$. Let δ_0 be as in* (4.15)*. Let $|t|, |h| \in B_\varepsilon(0)$. The following hold for some $\delta_0 < \delta_1 < \delta_2 < 1$ and for some $C', C'' > 0$.*

(i) *If $\alpha \in (0,1)$, then*

$$\|W(t,n)\| \le C'\delta_1^n|t|^\alpha\ell\Big(\frac{1}{|t|}\Big) \quad \text{and} \quad \|W(t+h,n) - W(t,n)\| \le C''\delta_1^n|h|^\alpha\ell\Big(\frac{1}{|h|}\Big).$$

(ii) *If $\alpha \in (1,2]$, then for any $\alpha' \in (1,\alpha)$.*

$$\big\|\big(W(t+h,n) - W(t,n) - ihQ(0)^{n-1}Q'(0)\big)\mathbb{1}\big\| \\ \le C'\delta_2^n|h|(|t|^{\alpha'-1} + |h|^{\alpha'-1}) + C''n|h|(|t|+|h|).$$

(iii) *If $\alpha \in (1,2]$, then W is differentiable and noticing $W'(t,n) = (Q(t)^n)'$,*

$$\|W'(t+h,n) - W'(t,n)\| = \|(Q(t+h)^n)' - (Q(t)^n)'\| \le C''\delta_1^n|h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big).$$

Proof. Compute that

$$\|W(t,n)\| = \sum_{j=0}^{n-1} \|Q(t)^j\|\,\|Q(t)-Q(0)\|\,\|Q(0)^{n-j-1}\|,$$

$$\|W(t+h,n)-W(t,n)\| \ll \sum_{j=0}^{n-1} \|Q(t+h)^j\|\,\|Q(t+h)-Q(t)\|\,\|Q(t)^{n-j-1}\|$$

and

$$(Q(t)^n)' = \sum_{j=0}^{n-1} Q(t)^j Q'(t) Q(t)^{n-j-1}.$$

Using these equations, as well as $\|Q(t+h)-Q(t)\| \ll |t|+|h|$ (which follows from (4.16) and $\|Q^n(t)\| \le C\delta_0^n$ from (4.15), we have

$$\|Q(t+h)-Q(0)\| \ll \delta_0^n(|t|+|h|). \tag{4.30}$$

(i) and (iii) The proof of these are included in [MT22, Proof of Lemma 3.7]; the claimed estimates follow directly from the previous three displayed formulae, equation (4.15), assumptions (H2)(i) and (H2)(iv), respectively and Lemma 4.12.

(ii) Note that $W(t+h,n)-W(t,n)-ih(Q(0)^n)' = Q(t+h)^n - Q(t)^n - ih(Q(0)^n)'$. We take $M \in \mathbb{N}$ such that $\|Q^M(t)\| \le C\delta_0^M \le \frac{1}{2}$ in (4.15).

We will need a two-step induction argument, first for $n \le M$, and then for $n > M$.

Step 1: when $n \le M$

Take the induction hypothesis: for any $\alpha' \in (1,\alpha)$, there is $C_{\mathrm{ind}} > 0$ such that

$$\begin{aligned}\Big\|Q(t+h)^n - Q(t)^n - ihQ(0)^{n-1}Q'(0) - ih\Big(Q(0)Q'(0)+Q'(0)Q(0)\Big)\sum_{j=0}^{n-1}Q(0)^j\Big\| \\ \ll C_{\mathrm{ind}}\, n|h|\Big(C^n\delta_1^{n-1}(|t|^{\alpha'-1}+|h|^{\alpha'-1}) + (|t|+|h|)\Big),\end{aligned} \tag{4.31}$$

for C from (4.15). We will first prove this statement by induction but only for $1 \le n \le M$. Afterwards we explain how to adjust the induction for $n > M$ and with a new uniform constant $\widehat{C}_{\mathrm{ind}}$ instead of $C_{\mathrm{ind}}C^n$. It is very likely that the second term $n|h|(|t|+|h|)$ can be improved, but this estimate suffices for the proof of the main result.

4

Since $Q(0)\mathbb{1} = R(0)(I-P(0))\mathbb{1} = 0$, we have $\sum_{j=1}^{n-1} Q(0)^j\mathbb{1} = 0$ and $n\delta_0^n \le \delta_1^n$ for some $\delta_1 \in (\delta_0, 1)$, (ii) follows.

We show that (4.31) holds using an induction argument. A straightforward (but lengthy) calculation verifies the statement for $n = 2$. To see this, compute that

$$\begin{aligned}Q(t+h)^2 - Q(t)^2 &= Q(t+h)\big(Q(t+h)-Q(t)\big) + \big(Q(t+h)-Q(t)\big)Q(t) \\ &= Q(t+h)\big(Q(t+h)-Q(t)-ihQ'(0)\big) \\ &\quad + \big(Q(t+h)-Q(t)-ihQ'(0)\big)Q(t) + ihQ(t+h)Q'(0) + ihQ'(0)Q(t).\end{aligned}$$

So,

$$\begin{aligned}&\big(Q(t+h)^2 - Q(t)^2\big) - ihQ(0)Q'(0) - ihQ'(0)Q(0) \\ &\quad = Q(t+h)\big(Q(t+h)-Q(t)-ihQ'(0)\big) + \big(Q(t+h)-Q(t)-ihQ'(0)\big)Q(t) \\ &\qquad + ih\big(Q(t+h)-Q(0)\big)Q'(0) + ihQ'(0)\big(Q(t)-Q(0)\big).\end{aligned}$$

Using Lemma 4.13(ii), (H2)(iii) and Lemma 4.12,

$$\big\|Q(t+h)^2 - Q(t)^2 - ihQ(0)Q'(0) - ihQ'(0)Q(0)\big\| \ll 2\delta_0|h|(|t|^{\alpha'-1} + |h|^{\alpha'-1}) + |h|(|t| + |h|)$$

for any $\alpha' \in (1, \alpha)$. This verifies the statement for $n = 2$. To make the assumption (4.31) more clear, we also record $n = 3$.

$$\begin{aligned} Q(t+h)^3 - Q(t)^3 &= Q(t+h)^2\big(Q(t+h) - Q(t)\big) + \big(Q(t+h)^2 - Q(t)^2\big)Q(t) \\ &= Q(t+h)^2\big(Q(t+h) - Q(t) - ihQ'(0)\big) \\ &\quad + \big(Q(t+h)^2 - Q(t)^2 - ihQ(0)Q'(0) - ihQ'(0)Q(0)\big)Q(t) \\ &\quad + ihQ(t+h)^2Q'(0) + ih\big(Q(0)Q'(0) + Q'(0)Q(0)\big)Q(t). \end{aligned}$$

So,

$$\begin{aligned} &Q(t+h)^3 - Q(t)^3 - ihQ(0)^2Q'(0) + ih\big(Q(0)Q'(0) + Q'(0)Q(0)\big)Q(0) \\ &\quad = Q(t+h)^2\big(Q(t+h) - Q(t) - ihQ'(0)\big) \\ &\qquad + \big(Q(t+h)^2 - Q(t)^2 - ihQ(0)Q'(0) - ihQ'(0)Q(0)\big)Q(t) \\ &\qquad + ih\big(Q(t+h)^2 - Q(0)^2\big)Q'(0) + ih\big(Q(0)Q'(0) + Q'(0)Q(0)\big)\big(Q(t) - Q(0)\big) \\ &\quad =: I_1 + I_2 + I_3 + I_4. \end{aligned}$$

Using the same estimates as for $n = 2$, we have $|I_1| \ll \delta_0^2|h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$. Using the statement on $n = 2$ and that $\|Q(t)\| \ll \delta_0$, $|I_2| \ll 2\delta_0^2\,|h|(|t|^{\alpha'-1} + \delta_0\,|h|^{\alpha'-1}) + |h|(|t| + |h|)$, $|I_3| \ll \delta_0^2|h|(|t| + |h|)$ (using (4.30) as well) and $|I_4| \ll |h|(|t| + |h|)$. Thus,

$$|I_1 + I_2 + I_3 + I_4| \ll 3\delta_0^2|h|(|t|^{\alpha'-1} + |h|^{\alpha'-1}) + (1 + \delta_0)|h|(|t| + |h|).$$

Since $1 + \delta_0 < 3$, the statement for $n = 3$ is verified.

The general case $n \in \mathbb{N}$ follows by induction. So, assume that (4.31) holds for n. We want to prove the hypothesis for $n + 1$. Note that

$$\begin{aligned} &Q(t+h)^{n+1} - Q(t)^{n+1} \\ &\quad = Q(t+h)^n\big(Q(t+h) - Q(t)\big) + \big(Q(t+h)^n - Q(t)^n\big)Q(t) \\ &\quad = Q(t+h)^n(Q(t+h) - Q(t) - ihQ'(0)) \\ &\qquad + \left(Q(t+h)^n - Q(t)^n - ihQ(0)^{n-1}Q'(0) - ih(Q(0)Q'(0) + Q'(0)Q(0))\sum_{j=0}^{n-1} Q(0)^j\right)Q(t) \\ &\qquad + ihQ(t+h)^nQ'(0) + ihQ(0)^{n-1}Q'(0)Q(t) + ih\Big(Q(0)Q'(0) + Q'(0)Q(0)\Big)\sum_{j=0}^{n-1} Q(0)^jQ(t). \end{aligned}$$

Thus,

$$\begin{aligned} &Q(t+h)^{n+1} - Q(t)^{n+1} - ihQ(0)^nQ'(0) - ih\big(Q(0)Q'(0) + Q'(0)Q(0)\big)\sum_{j=0}^{n-1} Q(0)^j\,Q(0) \\ &\quad = Q(t+h)^n\big(Q(t+h) - Q(t) - ihQ'(0)\big) \\ &\qquad + \left(Q(t+h)^n - Q(t)^n - ihQ(0)^{n-1}Q'(0) - ih\big(Q(0)Q'(0) + Q'(0)Q(0)\big)\sum_{j=0}^{n-1} Q(0)^j\right)Q(t) \\ &\qquad + ih(Q(t+h)^n - Q(0)^n)Q'(0) + ih\big(Q(0)Q'(0) + Q'(0)Q(0)\big)\sum_{j=0}^{n-1} Q(0)^j\big(Q(t) - Q(0)\big) \\ &\quad =: J_1 + J_2 + J_3 + J_4. \end{aligned}$$

4

Now, $|J_1| \ll \delta_0^n |h|(|t|^{\alpha'-1} + |h|^{\alpha'-1})$. Using (4.31) and (4.15),

$$|J_2| \le C C_{\mathrm{ind}} n|h| \Big(C^n \delta_1^{n-1} \left(|t|^{\alpha'-1} + |h|^{\alpha'-1}\right) + (|t| + |h|) \Big).$$

Using (4.30), $|J_3| \ll \delta_0^1 |h| \left(|t| + |h|\right)$. Since $\|Q(t) - Q(0)\| \ll |t|$ and $\|\sum_{j=0}^{n-1} Q(0)^j\| = \mathcal{O}(1)$, we get $|J_4| \ll |h|(|t| + |h|)$. The statement on $n+1$ follows by adding these estimates.

Step 2: when $n > M$

Next, we will do the induction in steps of M iterates. That is, we consider

$$\widehat{R}(t) := R^M(t) = \lambda^M(t) P(t) + Q^M(t) =: \widehat{\lambda}(t) P(t) + \widehat{Q}(t),$$

and recall that $\|\widehat{Q}\| \le \frac{1}{2}$ by the choice of M. The reason why the previous induction is of no use for all n is due to the factor C^n, for which we have no control; however, using steps of M iterates, C can be replaced by $\frac{1}{2}$, so if we use the induction hypothesis: for any $\alpha' \in (1, \alpha)$, there is $\widehat{C}_{\mathrm{ind}} > 0$ such that

$$\begin{aligned} &\left\| \widehat{Q}(t+h)^n - \widehat{Q}(t)^n - ih\widehat{Q}(0)^{n-1}\widehat{Q}'(0) - ih\Big(\widehat{Q}(0)\widehat{Q}'(0) + \widehat{Q}'(0)\widehat{Q}(0)\Big) \sum_{j=0}^{n-1} \widehat{Q}(0)^j \right\| \\ &\qquad \ll \widehat{C}_{\mathrm{ind}}\, n|h| \big(\delta_2^{n-1} \left(|t|^{\alpha'-1} + |h|^{\alpha'-1}\right) + (|t| + |h|) \big), \end{aligned}$$

for $\delta_2 := \delta_1^{1/M}$, the above induction proof works. Apart from M-dependent constants, the estimates of J_1, J_3 and J_4 still hold. This gives the required result for all $n = n_0 + jM$ and $j \ge 0$ (where the initial step $j = 0$ follows from the first induction for $1 \le n_0 \le M$). □

A final form of the decomposition of $\int_\Omega R(t)^n \mathbb{1}\, \mathrm{d}\mu$ into $\Psi(t)^n$ and 'good' quantities

To simplify the statement of Lemmas 4.17 and 4.18 further, we record the following facts. Recall that we have assumed that $\int_\Omega v \, \mathrm{d}\mu = 0$ (when $\alpha > 1$).

Lemma 4.19. *Let $(Q(0)^n)' = (Q(t)^n)'|_{t=0}$, $n \in \mathbb{N}$. Then*

$$\int_\Omega P'(0)\mathbb{1}\, \mathrm{d}\mu = 0 \quad \textit{and} \quad \int_\Omega (Q(0)^n)' \mathbb{1}\, \mathrm{d}\mu = \int_\Omega Q(0)^{n-1} Q'(0) \mathbb{1}\, \mathrm{d}\mu = 0.$$

Proof. As in the proof of Lemma 4.13,

$$P'(0)\mathbb{1} = \frac{1}{2\pi i} \int_{|\xi - 1| = \delta} \big(\xi - R(0)\big)^{-1} R'(0) \big(\xi - R(0)\big)^{-1} \mathbb{1}\, \mathrm{d}\xi.$$

However, $(\xi - R(0))^{-1}\mathbb{1} = (\xi - 1)^{-1}$ for all $\xi \ne 1$. Since we also know that $R'(0)\mathbb{1} = iR(0)v$, using Fubini's theorem we obtain

$$\begin{aligned} 2\pi i \int_\Omega P'(0)\mathbb{1}\, \mathrm{d}\mu &= i \int_\Omega \int_{|\xi-1|=\delta} (\xi - 1)^{-1} \big(\xi - R(0)\big)^{-1} R(0) v \, \mathrm{d}\xi \, \mathrm{d}\mu \\ &= i \int_{|\xi-1|=\delta} (\xi - 1)^{-1} \sum_{j \in \mathbb{N}_0} \xi^{-j-1} \int_\Omega R(0)^j R(0) v \, \mathrm{d}\mu \, \mathrm{d}\xi \\ &= i \int_{|\xi-1|=\delta} (\xi - 1)^{-1} \sum_{j \in \mathbb{N}_0} \xi^{-j-1} \int_\Omega R(0)^{j+1} v \, d\mu \, \mathrm{d}\xi \\ &= i \int_{|\xi-1|=\delta} (\xi - 1)^{-1} \sum_{j \in \mathbb{N}_0} \xi^{-j-1} \int_\Omega v \, \mathrm{d}\mu \, \mathrm{d}\xi = 0, \end{aligned}$$

concluding the argument for the first part of the statement.

4

Next, recall that $(Q(t)^n)' = \sum_{j=0}^{n-1} Q(t)^j Q'(t) Q(t)^{n-j}$. Hence,

$$\int_\Omega (Q(0)^n)' \mathbb{1} \, \mathrm{d}\mu = \int_\Omega \sum_{j=0}^{n-1} Q(0)^j Q'(0) Q(0)^{n-j-1} \mathbb{1} \, \mathrm{d}\mu.$$

Since $Q(0)\mathbb{1} = 0$ and $Q(0)^{n-j-1}\mathbb{1} = 0$ for all $j < n-1$, we thus have

$$\int_\Omega (Q(0)^n)' \mathbb{1} \, \mathrm{d}\mu = \int_\Omega Q(0)^{n-1} Q'(0) \mathbb{1} \, \mathrm{d}\mu.$$

Recall from (4.17) that $Q'(0)\mathbb{1} = -R(0)P'(0)\mathbb{1}$. As a consequence,

$$\begin{aligned}
\int_\Omega (Q(0)^n)' \mathbb{1} \, \mathrm{d}\mu &= -\int_\Omega Q(0)^{n-1} R(0) P'(0) \mathbb{1} \, \mathrm{d}\mu \\
&= -\int_\Omega R(0)^{n-1} (I - P(0)) R(0) P'(0) \mathbb{1} \, \mathrm{d}\mu \\
&= -\int_\Omega R(0)^n P'(0) \mathbb{1} \, \mathrm{d}\mu + \int_\Omega R(0)^{n-1} P(0) R(0) P'(0) \mathbb{1} \, \mathrm{d}\mu \\
&= -\int_\Omega P'(0) \mathbb{1} \, \mathrm{d}\mu + \int_\Omega R(0) P'(0) \mathbb{1} \, \mathrm{d}\mu = 0,
\end{aligned}$$

which ends the proof. □

Putting the last three lemmas together, we obtain the following.

Proposition 4.20. *Suppose that* (H1)–(H3) *hold. Set* $U(t,n) := \int_\Omega R(t)^n \mathbb{1} \, \mathrm{d}\mu - \Psi(t)^n$. *Then the following hold for all* $n \in \mathbb{N}$ *and all* $|t|, |h| \in B_\varepsilon(0)$ *with* $|h| \leq |t|/2$.

(i) *For* $\alpha \in (0,1)$,

$$|U(t,n)| \leq \left(C_0 |t|^\alpha \ell\left(\frac{1}{|t|}\right) + n|t|^{2\alpha} \ell\left(\frac{1}{|t|}\right)^2 \right) |\Psi(t)|^n, \quad \textit{and}$$

$$|U(t+h,n) - U(t,n)| \leq \left(|h|^\alpha \ell\left(\frac{1}{|h|}\right) + n|h|^\alpha \ell\left(\frac{1}{|h|}\right) |t|^\alpha \ell\left(\frac{1}{|t|}\right) \right) |\Psi(t)|^n.$$

(ii) *If* $\alpha \in (1,2]$, *we have for any* $\alpha' \in (1,\alpha)$ *that*

$$|U(t,n)| \leq \left(|t|^{\alpha'} + n|t|^2\right) |\Psi(t)|^n, \quad \textit{and}$$

$$|U(t+h,n) - U(t,n)| \leq |h||t|^{\alpha'-1} |\Psi(t)|^n + n|h||t|.$$

(iii) *If* $\alpha \in (1,2]$, *then* U *is differentiable in* t *and, writing* $U'(t,n)$ *for its derivative,*

$$U'(t,n) \ll \left(|t|^{\alpha-1} \ell_0\left(\frac{1}{|t|}\right) + n|t| + n^2 |t|^{\alpha+1} \ell_0\left(\frac{1}{|t|}\right) \right) |\Psi(t)|^n$$

and

$$\begin{aligned}
&\left| U'(t+h,n) - U'(t,n) \right| \\
&\quad \ll |h|^{\alpha-1} \ell_0\left(\frac{1}{|h|}\right) |\Psi(t)|^n + n\left(|t||h|^{\alpha'-1} \ell_0\left(\frac{1}{|h|}\right) + |h| \right) |\Psi(t)|^n \\
&\quad + n^2 \left(|t|^{2\alpha-1} |h| \ell_0\left(\frac{1}{|t|}\right)^2 + |t|^2 |h|^{\alpha-1} \ell_0\left(\frac{1}{|h|}\right) + |h||t|^\alpha \ell_0\left(\frac{1}{|t|}\right) \right) |\Psi(t)|^n \\
&\quad + n^3 |t|^{2\alpha} |h| \ell_0\left(\frac{1}{|t|}\right)^2 |\Psi(t)|^n.
\end{aligned}$$

Proof. In the notation of Lemma 4.17 and Lemma 4.18,

$$U(t,n) = \int_\Omega R(t)^n \mathbb{1}\,\mathrm{d}\mu - \Psi(t)^n = \widehat{V}(t,n) + \int_\Omega W(t,n)\mathbb{1}\,\mathrm{d}\mu$$

since $W(t,n)\mathbb{1} = Q(t)^n\mathbb{1} - Q(0)^n\mathbb{1} = Q(t)^n\mathbb{1}$ (because $Q(0)^n\mathbb{1} = 0$ for all $n \in \mathbb{N}$).

(i) This follows immediately from Lemma 4.17(i) and Lemma 4.18(i).

(ii) By Lemma 4.19, $\int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu = \int_\Omega (Q(0)^n)'\mathbb{1}\,\mathrm{d}\mu = \int_\Omega Q(0)^{n-1}Q'(0)\mathbb{1}\,\mathrm{d}\mu = 0$. Then Lemma 4.17(ii) gives,

$$|U(t,n)| \ll \big(|t|^{\alpha'} + n|t|^2\big)|\Psi(t)|^n,$$

which proves the first statement of item (ii). We already know that $\int_\Omega Q(0)^{n-1}Q'(0)\mathbb{1}\,\mathrm{d}\mu = 0$. This together with Lemma 4.18(ii) (and recalling that $|h| \le |t|/2$) implies that

$$\left|\int_\Omega (W(t+h,n) - W(t,n))\mathbb{1}\,\mathrm{d}\mu\right| \ll \delta_1^n |h||t|^{\alpha'-1} + n|h||t|.$$

Combining this with Lemma 4.17(iii),

$$\begin{aligned}|U(t+h,n) - U(t,n)| &\le \Big(|h||t|^{\alpha'-1} + n|h||t|^{\alpha-1}\ell_0(1/|t|)\Big)|\Psi(t)|^n + n|h||t| \\ &\ll |h||t|^{\alpha'-1}\,|\Psi(t)|^n + n|h||t|,\end{aligned}$$

which proves the second statement of item (ii).

(iii) This follows immediately from Lemma 4.17(iv), Lemma 4.18(iii) and also the fact that $\int_\Omega P'(0)\mathbb{1}\,\mathrm{d}\mu = 0$. □

4.3.5 Proof of Theorem 4.11

As already mentioned in the paragraph after the statement of Theorem 4.11, it suffices to show that (4.13) and (4.14) hold. In the notation of Proposition 4.20, $\mathbb{E}_\mu(e^{itv_n}) = \int_\Omega R(t)^n\mathbb{1}\,\mathrm{d}\mu = U(t,n) + \Psi(t)^n$ and the LHS of equation (4.13), for $\alpha \in (0,1)\cup(1,2)$, becomes

4

$$\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{U(t,n)}{it}\,\mathrm{d}t\right| = \mathcal{O}\big(D(N)\big) + o\Big(\frac{n\ell_0(N)}{N^\alpha}\Big). \tag{4.32}$$

where (in the notation of Theorem 4.11),

$$D(N) = \begin{cases} N^{-\alpha}\ell(N)\log N, & \text{if } \alpha \in (0,1), \\ N^{-(\alpha-\delta)}, & \text{if } \alpha \in (1,2]. \end{cases}$$

Similarly, for $\alpha = 2$, equation (4.14) becomes

$$\left|\int_{-\varepsilon}^{+\varepsilon} (e^{-itN} - e^{-it(N+g(n))})\Psi_Y(t)\frac{U(t,n)}{it}\,\mathrm{d}t\right| = \mathcal{O}\big(D(N)\big) + \mathcal{O}\Big(\frac{n\log N}{N^2}\Big) + \mathcal{O}\Big(\frac{n\widehat{\ell}(N)}{N^2}\Big). \tag{4.33}$$

In what follows we show that (4.32) and (4.33) hold by showing that for $\alpha \in (0,1)\cup(1,2)$,

$$\left|\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Psi_Y(t)\frac{U(t,n)}{it}\,\mathrm{d}t\right| = \mathcal{O}\big(D(N)\big) + o\big(nN^{-\alpha}\ell_0(N)\big). \tag{4.34}$$

and for $\alpha = 2$,

$$\left|\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Psi_Y(t)\frac{U(t,n)}{it}\,\mathrm{d}t\right| = \mathcal{O}\big(D(N)\big) + \mathcal{O}\left(\frac{n\log N}{N^2}\right) + \mathcal{O}\left(\frac{n\widehat{\ell}(N)}{N^2}\right). \tag{4.35}$$

A similar argument shows for $\alpha \in (0,1) \cup (1,2)$ that

$$\left|\int_{-\varepsilon}^{+\varepsilon} e^{-it(N+g(n))}\Psi_Y(t)\frac{U(t,n)}{it}\,\mathrm{d}t\right| = \mathcal{O}\big(D(N+g(n))\big) + \mathcal{O}\left(\frac{n\log(N+g(n))}{(N+g(n))^2}\right) + o\left(\frac{n\ell_0(N+g(n))}{(N+g(n))^\alpha}\right),$$

and a corresponding statement for $\alpha = 2$, concluding the proof of Theorem 4.11.

The validity of equations (4.34) and (4.35) will be shown via a 'modulus-of-continuity' argument. To carry out such an argument, it is crucial that we have enough decay in t in the expression of $U(t,n)$ as to counteract the effect of the division by t in the integrand of (4.34). Proposition 4.20 tracks this type of decay in t (along with finer continuity properties in h) and it is in this sense that we call $U(t,n)$ a 'good quantity'.

Before embarking on the final calculations, we need to recall one more technical lemma.

Lemma 4.21 ([MT22, Lemma 2.3])**.** *Let $L : (0,+\infty) \to (0,+\infty)$ be a slowly varying function. For all $\beta > 0$, there exists $C > 0$ so that for all $n \in \mathbb{N}$,*

$$\int_0^\varepsilon t^\beta L(t^{-1})|\Psi(t)|^n\,\mathrm{d}t \le C\frac{L(a_n)}{a_n^{1+\beta}}.$$

We may now proceed to the proof of (4.34).

Proof of Equation (4.34).

The case $\alpha \in (0,1)$**.** Set $f(t,n) := \Psi_Y(t)\frac{U(t,n)}{t}$ and $K(n,N) := \int_{-\varepsilon}^{+\varepsilon} e^{-itN}f(t,n)\,\mathrm{d}t$. To be able to apply the modulus-of-continuity argument we need to split the domain of integration (so as to be able to deal with the difference coming from $\frac{1}{t} - \frac{1}{t-\frac{\pi}{N}}$). For this purpose, write

$$\begin{aligned} K(n,N) &= \int_{-\frac{2\pi}{N}}^{\frac{2\pi}{N}} e^{-itN}f(t,n)\,\mathrm{d}t + \int_{-\varepsilon}^{-\frac{2\pi}{N}} e^{-itN}f(t,n)\,\mathrm{d}t + \int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}f(t,n)\,\mathrm{d}t \\ &=: K_0(n,N) + K^-(n,N) + K^+(n,N). \end{aligned}$$

By Proposition 4.20(i), $|f(t,n)| \ll (n|t|^{2\alpha-1}\ell(1/|t|)^2 + |t|^{\alpha-1}\ell(1/|t|))|\Psi(t)|^n$. Thus, we compute using Karamata's Theorem (see e.g. [BGT87, Theorem 1.5.11]) that

$$|K_0(n,N)| \ll \int_0^{\frac{2\pi}{N}} t^{\alpha-1}\ell\left(\frac{1}{t}\right)\mathrm{d}t + n\int_0^{\frac{2\pi}{N}} t^{2\alpha-1}\ell\left(\frac{1}{t}\right)^2\mathrm{d}t \ll \frac{\ell(N)}{N^\alpha} + \frac{n\ell(N)^2}{N^{2\alpha}} = o\left(\frac{n\ell(N)}{N^\alpha}\right).$$

We estimate K^+ via modulus-of-continuity; the estimate for K^- is similar. With the change of variables $t \to t - \pi/N$,

$$K^+(n,N) = -\int_{(2\pi+\pi)/N}^{\varepsilon+\pi/N} e^{-itN}f\left(t - \frac{\pi}{N}, n\right)\mathrm{d}t,$$

4

and summing the two expressions of K^+ gives

$$\begin{aligned}2K^+(n,N) &= -\int_{\varepsilon}^{\varepsilon+\pi/N} e^{-itN} f\Big(t-\frac{\pi}{N},n\Big)\,\mathrm{d}t + \int_{\frac{2\pi}{N}}^{\frac{2\pi+\pi}{N}} e^{-itN} f\Big(t-\frac{\pi}{N},n\Big)\,\mathrm{d}t \\ &\quad + \int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}\Big(f(t,n) - f\Big(t-\frac{\pi}{N},n\Big)\Big)\,\mathrm{d}t \\ &=: K_1 + K_2 + K_3.\end{aligned}$$

Since the integrands of K_1 and K_2 are bounded, $|K_1|, |K_2| = o(nN^{-\alpha}\ell(N))$. We continue with K_3, computing that

$$\begin{aligned}|K_3| &\le \left|\int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}\Big(\Psi_Y(t) - \Psi_Y\Big(t-\frac{\pi}{N}\Big)\Big)\frac{U(t,n)}{t}\,\mathrm{d}t\right| \\ &\quad + \left|\int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}\Psi_Y\Big(t-\frac{\pi}{N}\Big)\frac{U(t,n)-U(t-\frac{\pi}{N},n)}{t}\,\mathrm{d}t\right| \\ &\quad + \left|\int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}\Psi_Y\Big(t-\frac{\pi}{N}\Big)U\Big(t-\frac{\pi}{n},n\Big)\Big(\frac{1}{t}-\frac{1}{t-\frac{\pi}{N}}\Big)\,\mathrm{d}t\right| \\ &=: K_3^1 + K_3^2 + K_3^3.\end{aligned}$$

Since Ψ_Y is Lipschitz and $|f(t,n)| \ll (n|t|^{2\alpha-1}\ell(1/|t|)^2 + |t|^{\alpha-1}\ell(1/|t|))|\Psi(t)|^n$,

$$\begin{aligned}K_3^1 &\ll \frac{1}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-1}\ell\Big(\frac{1}{t}\Big)\,\mathrm{d}t + \frac{n}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{2\alpha-1}\ell\Big(\frac{1}{t}\Big)^2\,\mathrm{d}t \\ &\ll \frac{\ell(N)}{N^\alpha} + \frac{n\ell(N)}{N^{2\alpha}} = o(nN^{-\alpha}\ell(N)).\end{aligned}$$

So far, all the terms are as good as desired. The loss in the statement comes from treating the remaining integrals K_3^2 and K_3^3. By Proposition 4.20(i),

$$|U(t+h,n) - U(t,n)| \ll \Big(|h|^\alpha\ell\Big(\frac{1}{|h|}\Big) + n|h|^\alpha\ell\Big(\frac{1}{|h|}\Big)|t|^\alpha\ell\Big(\frac{1}{|t|}\Big)\Big)|\Psi(t)|^n.$$

Taking $h = \pi/N$,

$$\begin{aligned}|K_3^2| &\ll \frac{\ell(N)}{N^\alpha}\int_{\frac{2\pi}{N}}^{\varepsilon}\frac{1}{t}\,\mathrm{d}t + \frac{n\ell(N)}{N^\alpha}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-1}\ell\Big(\frac{1}{t}\Big)|\Psi(t)|^n\,\mathrm{d}t \\ &\ll \frac{\ell(N)\log N}{N^\alpha} + \frac{\ell(N)}{N^\alpha}\, n\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-1}\ell\Big(\frac{1}{t}\Big)|\Psi(t)|^n\,\mathrm{d}t. \qquad (4.36)\end{aligned}$$

The second term is unproblematic. By Lemma 4.21, with $\beta = \alpha - 1$ and $L = \ell$,

$$n\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-1}\ell\Big(\frac{1}{t}\Big)\,|\Psi(t)|^n\,\mathrm{d}t \ll \frac{n}{a_n^\alpha}\ell(a_n) \ll 1.$$

Hence, $K_3^2 \ll N^{-\alpha}\ell(N)\log N$.

For K_3^3, we use again that $|U(t,n)/t| \ll n|t|^{2\alpha-1}\ell(1/|t|)^2 + |t|^{\alpha-1}\ell(1/|t|)$ and that $\left|\frac{1}{t} - \frac{1}{t-\frac{\pi}{N}}\right| \ll N^{-1}t^{-2}$. So, by Karamata's Theorem,

$$|K_3^3| \ll \frac{1}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-2}\ell\Big(\frac{1}{t}\Big)\,\mathrm{d}t + \frac{n}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{2\alpha-2}\ell\Big(\frac{1}{t}\Big)^2\,\mathrm{d}t \ll \frac{\ell(N)}{N^\alpha} + \frac{n\ell(N)^2}{N^{2\alpha}} \ll \frac{\ell(N)}{N^\alpha} + o\Big(\frac{n\ell(N)}{N^\alpha}\Big),$$

ending the argument for case $\alpha \in (0,1)$.

The case $\alpha \in (1,2]$**.** Set $M(n,N) := \int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Psi_Y(t)U(t,n)/it\,\mathrm{d}t$. Since the boundary terms cancel, integration by parts gives

$$\begin{aligned} M(n,N) &= \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Psi_Y'(t)\frac{U(t,n)}{it}\,\mathrm{d}t + \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Psi_Y(t)\frac{U'(t,n)}{it}\,\mathrm{d}t \\ &\quad - \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Psi_Y(t)\frac{U(t,n)}{t^2}\,\mathrm{d}t \\ &=: \frac{1}{iN}M_1(n,N) + \frac{1}{iN}M_2(n,N) + \frac{1}{iN}M_3(n,N). \end{aligned} \tag{4.37}$$

Throughout the rest of the proof, we take α' close to α (that is, $\alpha - \alpha'$ is positive but as small as we want). By Proposition 4.20(ii), $|U(t,n)| \ll n|t|^2 + |t|^{\alpha'}$. By Proposition 4.20(iii), $|U'(t,n)| \ll (n|t| + |t|^{\alpha-1}\ell_0(1/|t|) + n^2|t|^{\alpha+1}\ell_0(1/|t|))|\Psi(t)|^n$. Hence, integrating by parts once more,

$$\begin{aligned} \left|\frac{1}{iN}M_1(n,N)\right| &\ll \frac{1}{N^2}\int_0^{\varepsilon} |\Psi_Y''(t)|\frac{|U(t,n)|}{t}\,\mathrm{d}t + \frac{1}{N^2}\int_0^{\varepsilon} |\Psi_Y'(t)|\frac{|U'(t,n)|}{t}\,\mathrm{d}t \\ &\quad + \frac{1}{N^2}\int_0^{\varepsilon} |\Psi_Y'(t)|\frac{|U(t,n)|}{t^2}\,\mathrm{d}t \\ &\ll \frac{n}{N^2} = o\left(\frac{n\ell_0(N)}{N^\alpha}\right), \end{aligned} \tag{4.38}$$

using that the variance is infinite by assumption when $\alpha = 2$ so $\ell_0(N) \to +\infty$. Here, we have also used that Ψ_Y is $\mathcal{C}^2$ (with bounded first and second derivative).

Next we estimate M_2 and M_3 via the modulus-of-continuity argument similarly to the proof in the case $\alpha \in (0,1)$ above.

Regarding M_2,

$$\begin{aligned} M_2(n,N) &= \int_{-\frac{2\pi}{N}}^{\frac{2\pi}{N}} e^{-itN}\Psi_Y(t)\frac{U'(t,n)}{t}\,\mathrm{d}t + \int_{-\varepsilon}^{-\frac{2\pi}{N}} e^{-itN}\Psi_Y(t)\frac{U'(t,n)}{t}\,\mathrm{d}t \\ &\quad + \int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}\Psi_Y(t)\frac{U'(t,n)}{t}\,\mathrm{d}t \\ &=: M_2^0(n,N) + M_2^-(n,N) + M_2^+(n,N). \end{aligned} \tag{4.39}$$

Next,

$$\begin{aligned} M_2^0(n,N) &\ll n\int_0^{\frac{2\pi}{N}} 1\,\mathrm{d}t + \int_0^{\frac{2\pi}{N}} t^{\alpha-2}\ell_0\left(\frac{1}{t}\right)\mathrm{d}t + n^2\int_0^{\frac{2\pi}{N}} t^{\alpha}\ell_0\left(\frac{1}{t}\right)\mathrm{d}t \\ &\ll \frac{n}{N} + \frac{\ell_0(N)}{N^{\alpha-1}} + \frac{n^2\ell_0(N)}{N^{\alpha+1}} = \mathcal{O}\left(\frac{n}{N}\right) + o\left(\frac{n\ell_0(N)}{N^{\alpha-1}}\right). \end{aligned}$$

Similarly to the argument used above for the case $\alpha \in (0,1)$ (when estimating K^+, K^-), it suffices to estimate $M_2^+(n,N)$. Set $g(t,n) := \Psi_Y(t)\frac{U'(t,n)}{t}$ and compute that

$$\begin{aligned} 2M_2^+(n,N) &= -\int_{\varepsilon}^{\varepsilon+\frac{\pi}{N}} e^{-itN}g\left(t-\frac{\pi}{N},n\right)\mathrm{d}t + \int_{\frac{2\pi}{N}}^{\frac{2\pi+\pi}{N}} e^{-itN}g\left(t-\frac{\pi}{N},n\right)\mathrm{d}t \\ &\quad + \int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}\left(g(t,n) - g\left(t-\frac{\pi}{N},n\right)\right)\mathrm{d}t \\ &=: M_2^1 + M_2^2 + M_2^3. \end{aligned} \tag{4.40}$$

The estimates for M_2^1 and M_2^2 go similarly to the estimate for $M_2^0(n,N)$ above and give $|M_2^1| = \mathcal{O}(1/N) = o(nN^{-(\alpha-1)}\ell_0(N))$ and $|M_2^2| = o(nN^{-(\alpha-1)}\ell_0(N))$.

Regarding M_2^3, recalling that Ψ_Y is bounded, we compute that

$$\begin{aligned}|M_2^3| &\ll \int_{\frac{2\pi}{N}}^{\varepsilon} \left|\Psi_Y(t) - \Psi_Y\Big(t - \frac{\pi}{N}\Big)\right| \frac{|U'(t,n)|}{t}\,\mathrm{d}t + \int_{\frac{2\pi}{N}}^{\varepsilon} \Psi_Y\Big(t-\frac{\pi}{N}\Big)\frac{\big|U'(t,n) - U'\big(t-\frac{\pi}{N},n\big)\big|}{t}\,\mathrm{d}t \\ &\quad + \int_{\frac{2\pi}{N}}^{\varepsilon} \Psi_Y\Big(t-\frac{\pi}{N}\Big)\Big|U'\Big(t-\frac{\pi}{N},n\Big)\Big|\left|\frac{1}{t} - \frac{1}{t-\frac{\pi}{N}}\right|\mathrm{d}t \\ &=: L_1 + L_2 + L_3.\end{aligned}$$

Recall that $|U'(t,n)/t| \ll (n + |t|^{\alpha-2}\ell_0(1/|t|) + n^2|t|^{\alpha}\ell_0(1/|t|))|\Psi(t)|^n$. Since Ψ_Y is Lipschitz, by Lemma 4.21 with $\beta \in \{\alpha-2, \alpha\}$ and $L = \ell$, we get

$$L_1 \ll \frac{n}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} |\Psi(t)|^n\,\mathrm{d}t \ll \frac{n}{Na_n}, \tag{4.41}$$

using that $a_n^{\alpha} \sim n\ell_0(a_n)$. By Proposition 4.20(iii),

$$\begin{aligned}&\big|U'(t+h,n) - U'(t,n)\big| \\ &\quad\ll |h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big)|\Psi(t)|^n + n\Big(|t||h|^{\alpha'-1}\ell_0\Big(\frac{1}{|h|}\Big) + |h|\Big)|\Psi(t)|^n \\ &\qquad + n^2\left(|t|^{2\alpha-1}|h|\ell_0\Big(\frac{1}{|t|}\Big)^2 + |t|^2|h|^{\alpha-1}\ell_0\Big(\frac{1}{|h|}\Big) + |h||t|^{\alpha}\ell_0\Big(\frac{1}{|t|}\Big)\right)|\Psi(t)|^n \\ &\qquad + n^3|t|^{2\alpha}|h|\ell_0\Big(\frac{1}{|t|}\Big)^2|\Psi(t)|^n.\end{aligned}$$

Using this estimate with $h = \pi/N$ and recalling that $|\Psi(t)| \le e^{-ct^{\alpha}\ell_0(1/t)}$,

$$\begin{aligned}L_2 &\ll \frac{\ell_0(N)}{N^{\alpha-1}}\int_{\frac{2\pi}{N}}^{\varepsilon}\frac{1}{t}\,\mathrm{d}t + \frac{n\ell_0(N)}{N^{\alpha'-1}}\int_{\frac{2\pi}{N}}^{\varepsilon}|\Psi(t)|^n\,\mathrm{d}t + \frac{n}{N}\int_{\frac{2\pi}{N}}^{\varepsilon}\frac{1}{t}\,\mathrm{d}t \\ &\quad + \frac{n^2\ell_0(N)}{N^{\alpha-1}}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{2\alpha-2}\ell_0\Big(\frac{1}{t}\Big)|\Psi(t)|^n\,\mathrm{d}t + \frac{n^2\ell_0(N)}{N^{\alpha-1}}\int_{\frac{2\pi}{N}}^{\varepsilon} t\,|\Psi(t)|^n\,\mathrm{d}t \\ &\quad + \frac{n^2}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-1}\ell_0\Big(\frac{1}{t}\Big)|\Psi(t)|^n\,\mathrm{d}t + \frac{n^3}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{2\alpha-1}\ell_0\Big(\frac{1}{t}\Big)^2|\Psi(t)|^n\,\mathrm{d}t.\end{aligned}$$

To treat the latter four integrals, we apply Lemma 4.21 with various choice of β and L: taking $\beta = \alpha - 1$ and $L = \ell_0$, we get

$$\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-1}\ell_0\Big(\frac{1}{t}\Big)|\Psi(t)|^n\,\mathrm{d}t \ll \frac{\ell_0(a_n)}{a_n^{\alpha}} \ll \frac{1}{n}. \tag{4.42}$$

Taking $\beta = 2\alpha - 1$ and $L = \ell_0$, and using that $a_n^{\alpha} \sim n\ell_0(a_n)$, we get

$$\begin{aligned}&\int_{\frac{2\pi}{N}}^{\varepsilon} t^{2\alpha-2}\ell_0\Big(\frac{1}{t}\Big)|\Psi(t)|^n\,\mathrm{d}t \ll \frac{\ell_0(a_n)}{a_n^{2\alpha-1}} \ll \frac{1}{n^{2-1/\alpha-\delta^*}}, \quad \text{for any } \delta^* > 0, \quad \text{and} \\ &\int_{\frac{2\pi}{N}}^{\varepsilon} t^{2\alpha-1}\ell_0\Big(\frac{1}{t}\Big)^2|\Psi(t)|^n\,\mathrm{d}t \ll \frac{\ell_0(a_n)^2}{a_n^{2\alpha}} \ll \frac{1}{n^2}.\end{aligned} \tag{4.43}$$

Taking $\beta = 1$ and $L = 1$,

$$\int_{\frac{2\pi}{N}}^{\varepsilon} t\,|\Psi(t)|^n\,\mathrm{d}t \ll \frac{1}{a_n^2} \ll \frac{1}{(n\ell_0(a_n))^{2/\alpha}} \ll \frac{1}{n^{2/\alpha-\delta^*}}, \quad \text{for any } \delta^* > 0. \tag{4.44}$$

Using the same computation as in (4.41), $\int_{\frac{2\pi}{N}}^{\varepsilon} |\Psi(t)|^n\,\mathrm{d}t \ll a_n^{-1}$. These estimates together with (4.42), (4.43) and (4.44) (after a change of variables $t \to \sigma/a_n$) imply that

$$\begin{aligned} L_2 &\ll \frac{n\ell_0(N)}{N^{\alpha'-1}}\frac{1}{a_n} + \frac{n\log N}{N} + \frac{\log N\,\ell_0(N)}{N^{\alpha-1}} + \frac{n^2\ell_0(N)}{N^{\alpha-1}}\frac{1}{n^{2-1/\alpha-\delta^*}} \\ &\quad + \frac{n^2}{N}\frac{1}{n} + \frac{n^2\ell_0(N)}{N^{\alpha-1}}\frac{1}{n^{2/\alpha-\delta^*}} + \frac{n^3}{N}\frac{1}{n^2} \\ &= \mathcal{O}\Big(\frac{\ell_0(N)\log N}{N^{\alpha-1}}\Big) + \mathcal{O}\Big(\frac{n\log N}{N}\Big) + o\Big(\frac{n\ell_0(N)}{N^{\alpha'-1}}\Big); \end{aligned}$$

here, we have used that $\frac{n^2}{n^{2-1/\alpha-\delta^*}} < n$, $\frac{n^2}{n^{2/\alpha-\delta^*}} < n$ and that $a_n^\alpha \sim n\ell_0(a_n)$. The argument for estimating L_3 goes similarly to the argument used for K_3^2 in (4.36) (inside the proof for the case $\alpha \in (0,1)$ above). Recall that $|U'(t,n)| \ll (|t|^{\alpha-1}\ell_0(1/|t|) + n|t| + n^2|t|^{\alpha+1}\ell_0(1/|t|))|\Psi(t)|^n$ and $\left|\frac{1}{t} - \frac{1}{t-\frac{\pi}{N}}\right| \ll N^{-1}t^{-2}$. Hence, using Karamata's Theorem,

$$\begin{aligned} L_3 &\ll \frac{1}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-3}\ell_0\Big(\frac{1}{t}\Big)\,\mathrm{d}t + \frac{n}{N}\int_{\frac{2\pi}{N}}^{\varepsilon}\frac{1}{t}\,\mathrm{d}t + \frac{n^2}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha-1}\ell_0\Big(\frac{1}{t}\Big)|\Psi(t)|^n\,\mathrm{d}t \\ &\ll \frac{\ell_0(N)}{N^{\alpha-1}} + \frac{n\log N}{N} + \frac{n}{N}. \end{aligned}$$

where we have used (4.42). Putting all the above together and recalling (4.40),

$$\left|\frac{1}{iN}M_2(n,N)\right| = \mathcal{O}\Big(\frac{\ell_0(N)\log N}{N^{\alpha'}}\Big) + \mathcal{O}\Big(\frac{n\log N}{N^2}\Big) + o\Big(\frac{n\ell_0(N)}{N^{\alpha}}\Big). \tag{4.45}$$

It remains to estimate $M_3(n,N)$ in (4.37). We use the modulus-of-continuity argument again. Similarly to (4.39),

$$\begin{aligned} M_3(n,N) &= \int_{-\frac{2\pi}{N}}^{\frac{2\pi}{N}} e^{-itN}\Psi_Y(t)\frac{U(t,n)}{t^2}\,\mathrm{d}t + \int_{-\varepsilon}^{-\frac{2\pi}{N}} e^{-itN}\Psi_Y(t)\frac{U(t,n)}{t^2}\,\mathrm{d}t \\ &\quad + \int_{\frac{2\pi}{N}}^{\varepsilon} e^{-itN}\Psi_Y(t)\frac{U(t,n)}{t^2}\,\mathrm{d}t \\ &=: M_3^0(n,N) + M_3^-(n,N) + M_3^+(n,N). \end{aligned}$$

Recall that, by Proposition 4.20(ii), $|U(t,n)| \ll (|t|^{\alpha'} + n|t|^2)|\Psi(t)|^n$ for any $\alpha' \in (1,\alpha)$. An easy computation shows that $|M_3^0(n,N)| = \mathcal{O}(N^{1-\alpha'}) + \mathcal{O}(nN^{-1}) = \mathcal{O}(N^{1-\alpha'}) + o(n\ell_0(N)N^{\alpha-1})$, noting that we have infinite variance by assumption when $\alpha = 2$ so $\ell_0(N) \to +\infty$.

Similarly to the argument used for (4.39), it suffices to estimate $M_3^+(n,N)$. We compute that

$$\begin{aligned} |M_3^+| &\ll \int_{\frac{2\pi}{N}}^{\varepsilon}\Big|\Psi_Y(t) - \Psi_Y\Big(t-\frac{\pi}{N}\Big)\Big|\frac{|U(t,n)|}{t^2}\,\mathrm{d}t + \int_{\frac{2\pi}{N}}^{\varepsilon}\Psi_Y\Big(t-\frac{\pi}{N}\Big)\frac{\big|U(t,n) - U\big(t-\frac{\pi}{N},n\big)\big|}{t^2}\,\mathrm{d}t \\ &\quad + \int_{\frac{2\pi}{N}}^{\varepsilon}\Psi_Y\Big(t-\frac{\pi}{N}\Big)\Big|U\Big(t-\frac{\pi}{N},n\Big)\Big|\left|\frac{1}{t^2} - \frac{1}{\big(t-\frac{\pi}{N}\big)^2}\right|\mathrm{d}t \\ &=: S_1 + S_2 + S_3. \end{aligned}$$

Clearly, $S_1 \ll nN^{-1} = o(nN^{-(\alpha-1)}\ell_0(N))$. By Proposition 4.20(ii), for any $\alpha' \in (1,\alpha)$,

$$|U(t+h,n) - U(t,n)| \ll |h||t|^{\alpha'-1}|\Psi(t)|^n + n|h||t|.$$

Taking $h = \pi/N$,

$$S_2 \ll \frac{1}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha'-3}\,\mathrm{d}t + \frac{n}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{-1}\,\mathrm{d}t \ll \frac{1}{N^{\alpha'-1}} + \frac{n\log N}{N}.$$

Finally, recall again that, by Proposition 4.20(ii), $|U(t,n)| \ll |t|^{\alpha'} + n|t|^2$ for any $\alpha' \in (1,\alpha)$. Hence,

$$S_3 \ll \frac{1}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{\alpha'-3}\,\mathrm{d}t + \frac{n}{N}\int_{\frac{2\pi}{N}}^{\varepsilon} t^{-1}\,\mathrm{d}t \ll \frac{1}{N^{\alpha'-1}} + \frac{n\log N}{N}.$$

Thus,

$$|M_3^+| = \begin{cases} \mathcal{O}(N^{-(\alpha'-1)}) + o\big(nN^{-(\alpha-1)}\ell_0(N)\big), & \text{if } \alpha \in (1,2),\\ \mathcal{O}(N^{-(\alpha'-1)}) + \mathcal{O}\big(nN^{-1}(\ell_0(N)+\log N)\big), & \text{if } \alpha = 2. \end{cases}$$

The integral $|M_3^-|$ can be estimated similarly and gives similar estimates. Altogether,

$$\left|\frac{1}{iN}M_3(n,N)\right| = \begin{cases} \mathcal{O}(N^{-\alpha'}) + o\big(nN^{-\alpha}\ell_0(N)\big), & \text{if } \alpha \in (1,2),\\ \mathcal{O}(N^{-\alpha'}) + \mathcal{O}\big(nN^{-2}(\ell_0(N)+\log N)\big) & \text{if } \alpha = 2. \end{cases}$$

The conclusion follows from the previous displayed equation together with (4.37), (4.38), (4.45) and also by recalling that $\ell_0(N)\log N \ll N^{\delta_0}$ for any δ_0 arbitrarily small. □

4.4 Applications

4.4.1 Gibbs-Markov Maps

4 Let (X,μ) be a probability space with an at most countable measurable partition $\{I_j\}_{j\in\mathbb{N}}$, and let $T : X \to X$ be an ergodic measure-preserving transformation. Recall from Definition 1.1 that T is a Gibbs-Markov map if it is a Markov map with the big images property and bounded distortion. Fix $\theta \in (0,1)$, and recall from §1.2 that this gives a metric d_θ on X defined by $d_\theta(x,x') = \theta^{s(x,x')}$, where $s(x,x')$ denotes the least integer $n \in \mathbb{N}_0$ such that $T^n(x)$ and $T^n(x')$ lie in distinct partition elements. Given $f : X \to \mathbb{R}$, set

$$D_j f := \sup\left\{\frac{|f(x)-f(x')|}{d_\theta(x,x')} : x,x' \in I_j,\, x \neq x'\right\}, \qquad |f|_\theta := \sup\{D_j f : j \in \mathbb{N}\}.$$

We let the Banach space $\mathcal{B}_\theta \subset \mathcal{L}^\infty$ consist of functions $f : X \to \mathbb{R}$ such that $|f|_\theta < +\infty$ with norm $\|f\|_\theta = |f|_\infty + |f|_\theta < +\infty$.

Proposition 4.22. *Suppose that T is a mixing Gibbs-Markov map, and take $v : X \to \mathbb{R}$ with $\int_X v^2\,\mathrm{d}\mu = +\infty$ and $|v|_\theta < +\infty$. Assume moreover that there exists $C > 0$ so that*

$$\big|v|_{I_j}\big|_\theta \le C \inf_{j\in\mathbb{N}}\big|v|_{I_j}\big|, \qquad \forall j \in \mathbb{N}. \tag{4.46}$$

Fix $\alpha \in (0,1) \cup (1,2]$, and assume that the tails of v satisfy (H1). *Then conditions* (H1)–(H3) *are satisfied with Banach space $\mathcal{B} = \mathcal{B}_\theta$.*

Proof. Condition (H1) is an assumption on v and condition (H3) is well known; see [Aar97, Chapter 4] and [AD01]. Conditions (H2)(i), (ii) are again well known, and condition (H2)(iv) has been verified in [MT22, Proof of Proposition 3.10].

It remains to verify (H2)(ii). First note that

$$\|R(t+h) - R(t) - ihR'(0)\| \leq \big\|R\big(e^{itv}\big(e^{ihv} - 1 - ihv\big)\big)\big\| + \big\|R\big(ihv(e^{itv} - 1)\big)\big\|.$$

Let $\varphi \in \{v(e^{itv} - 1), e^{itv}\big(e^{ihv} - 1 - ihv\big)\}$, and recall that φ satisfies assumption (4.46). As shown in [MT17, Proof of Proposition 12.1], under (4.46), there exists $C > 0$ so that, for all $f \in \mathcal{B}_\theta$,

$$\|R(\varphi f)\| \leq C\|f\|\,\|\varphi\|_{\mathcal{L}^1(\mu)}.$$

Since we know v satisfies (H1), we have for any $\alpha' \in (1, \alpha)$,

$$\|v(e^{itv} - 1)\|_{\mathcal{L}^1(\mu)} \ll |t|^{\alpha'-1} \text{ and } \big\|e^{ihv} - 1 - ihv\big\|_{\mathcal{L}^1(\mu)} \ll |h|^{\alpha'}.$$

Thus, $\big\|R(e^{itv}\big(e^{ihv} - 1 - ihv\big)\big\| \ll |h|^{\alpha'}$ and $\|R(ihv(e^{itv} - 1))\| \ll |h||t|^{\alpha'-1}$. □

4.4.2 Hölder Mean of Continued Fraction Expansion Digits

In this section, we prove Corollary 4.3 on the Hölder mean of continued fraction digits as an application of Theorem 4.11. We will need the following lemma, which states that a sequence of random variables defined via continued fraction digits satisfies (H1) with $\alpha \in (0, 1) \cup (1, 2]$.

Lemma 4.23. *Suppose that $\psi : (0, +\infty) \to (0, +\infty)$ is a regularly varying function of index $\rho > 0$ of the form $\psi(x) = x^\rho \ell^a(x^{\frac{\rho}{a}})\big(1 + o(1)\big)$ for some slowly varying function ℓ and constant $a > 0$. Suppose further that ψ is locally bounded on $[X, +\infty)$ for some $X \geq 0$. Then as $x \to +\infty$,*

$$\mu_G(\psi \circ d_1 > x) = \frac{1 + o(1)}{\log 2} x^{-1/\rho}\big(\ell^{\#}(x^{1/a})\big)^{-a/\rho}.$$

Proof of Corollary 4.3. Fix $\rho > 1$, $a > 0$ and take x in the range $x \in (n, e^{an})$ as $n \to +\infty$. To mimic earlier notation, set $\alpha := 1/\rho \in (0, 1)$ and $N = N(n, x) = x^\rho n$. Define $\psi : \mathbb{N} \to (0, +\infty)$ by $\psi(n) = n^\rho$. Since $d_j = d_1 \circ G^{j-1}$ for all $j \in \mathbb{N}$, we have

$$\mu_G\left(\left\{y \in [0, 1] \setminus \mathbb{Q} : \left(\frac{d_1^\rho(y) + \cdots + d_n^\rho(y)}{n}\right)^{\frac{1}{\rho}} > x\right\}\right) = \mu_G\big(S_n(\psi \circ d_1) > N\big).$$

By Lemma 4.23, we have $\mu_G(\psi \circ d_1 \geq x) = x^{-\alpha}\ell(x)\big(1 + o(1)\big)$, where $\ell(x) \equiv (\log 2)^{-1}$. In addition, it was shown in [Aar97] that the continued fraction map is Gibbs-Markov. Consequently, we may apply Proposition 4.22 to see that (H1)–(H3) are satisfied. Since $a_n = n^\rho(1 + o(1))$ and $b_n = 0$ (recall (1.9)), and since $x \in (n, e^{an})$ as $n \to +\infty$, it follows that $N \in (a_n, e^{a'n})$ as $n \to +\infty$ for some constant $a' > 0$ so we may apply Theorem 4.11 (and Corollary 4.2) with $v = \psi \circ d_1$ to see that

$$\mu_G\big(S_n(\psi \circ d_1) > N\big) = n\mu_G(\psi \circ d_1 > N)\big(1 + o(1)\big) = \frac{n}{N^\alpha \log 2}\big(1 + o(1)\big).$$

Putting the above two equations together completes the proof. □

It remains to prove Lemma 4.23.

4

Proof of Lemma 4.23. Note that ψ is locally bounded on $[X, +\infty)$ so the results of [BGT87, §1.5] apply. First, we may apply [BGT87, Theorem 1.5.3] to obtain an increasing function $\psi_1 = \big(1 + o(1)\big)\psi$. Fix $x > 0$, and set

$$A_x := \{n \in \mathbb{N} : \psi_1(n) \geq x\}, \qquad \varphi_1(x) := \min\{n \in \mathbb{N} : \psi_1(n) \geq x\}.$$

Since ψ_1 is increasing, it follows that $A_x = \{n \in \mathbb{N} : n \geq \varphi_1(x)\}$. Now, the set of points $x \in [0,1] \setminus \mathbb{Q}$ whose first digit in the Lüroth series expansion is $d_1(x) = n$ is $[\frac{1}{n+1}, \frac{1}{n}] \setminus \mathbb{Q}$ so we have by the dominated convergence theorem that

$$\begin{aligned}\mu_G(\psi_1 \circ d_1 \geq x) &= \int_{[0,1]} \sum_{n \in A_x} \mathbb{1}_{[\frac{1}{n+1}, \frac{1}{n}] \setminus \mathbb{Q}} \,\mathrm{d}\mu_G = \sum_{n \in A_x} \mu_G\Big(\Big[\frac{1}{n+1}, \frac{1}{n}\Big]\Big) \\ &= \sum_{n \geq \varphi_1(x)} \mu_G\Big(\Big[\frac{1}{n+1}, \frac{1}{n}\Big]\Big) = \mu_G\Big(\Big(0, \frac{1}{\varphi_1(x)}\Big]\Big) \\ &= \frac{1}{\log 2} \log\Big(1 + \frac{1}{\varphi_1(x)}\Big) = \frac{1 + o(1)}{\varphi_1(x) \log 2},\end{aligned}$$

using that $\mu_G(\mathbb{Q}) = 0$ (since $\mu_G \ll \lambda$). It remains to find the asymptotic behaviour of φ_1. By [BGT87, Theorem 1.5.2], φ_1 is an asymptotic inverse of ψ_1 so, by [BGT87, Theorem 1.5.3], it follows that

$$\varphi_1(x) = x^{1/\rho}\big(\ell^{\#}(x^{1/a})\big)^{a/\rho}\big(1 + o(1)\big).$$

Therefore,

$$\mu_G(\psi_1 \circ d_1 \geq x) = \frac{1 + o(1)}{\log 2} x^{-1/\rho}\big(\ell^{\#}(x^{1/a})\big)^{-a/\rho}.$$

Since $\mu_G(\psi \circ d_1 \geq x)\big) = \mu_G\big(\psi_1 \circ d_1 \geq x(1 + o(1))\big)$, the result follows. ☐

4.4.3 Non-Markov Expanding Interval Maps (a.k.a. AFU maps)

4

Let $\{I_j\}_j$ be a measurable partition of $[0,1]$ consisting of open intervals, and recall the definition of an AFU map on the partition $\{I_j\}$ from Definition 1.2. Since AFU maps are not necessarily Markov, Hölder spaces are not preserved by the transfer operator of T so it is therefore standard to consider the space of bounded variation functions instead. Accordingly, we define the Banach space $\mathcal{B} = BV \subset \mathcal{L}^\infty$ to consist of functions $f : \Omega \to \mathbb{R}$ such that $\mathrm{Var}(f) < +\infty$ with norm $\|f\| = |f|_\infty + \mathrm{Var}(f)$, where

$$\mathrm{Var}(f) = \inf_{g \sim f} \sup_{0 = y_0 < \ldots < y_k = 1} \sum_{i=1}^{k} |g(y_i) - g(y_{i-1})|,$$

denotes the variation of the (equivalence class) of f, where $g \sim f$ if f and g differ on a null set. Also, we let $\mathrm{Var}_I(f)$ denote the variation of f on I.

We suppose that $T : \Omega \to \Omega$ is topologically mixing. Then there is a unique absolutely continuous T-invariant probability measure μ, and μ is mixing.

Proposition 4.24. *Assume T is a topologically mixing AFU map, and take $v : \Omega \to \mathbb{R}$ with $\int_\Omega v^2 \,\mathrm{d}\mu = +\infty$ and $\sup_I \mathrm{Var}_I(v) < +\infty$. Fix $\alpha \in (0,1) \cup (1,2]$, and assume that the tails of v satisfy* (H1). *Then conditions* (H1)–(H3) *are satisfied with Banach space $\mathcal{B} = \mathrm{BV}$.*

Proof. Condition (H1) is an assumption and condition (H3) is well-known for mixing AFU maps; see [Zwe98]. Apart from (H2)(iii), the remaining items of (H2) have been clarified in [MT22, Proof of Proposition 3.11].

The verification of (H2)(iii) is similar to the one in Proposition 4.22 with the only change being that the assumption in (4.46) is replaced by $\sup_I \operatorname{Var}_I(v) < +\infty$. The latter ensures that $\operatorname{Var}_I(e^{itv}) \ll |t| \operatorname{Var}_I(v) \ll |t|$ and similarly $\operatorname{Var}_I(e^{ihv}) \ll |h|$. Using this and proceeding as in [MT22, Proof of Proposition 3.11], we have for $\varphi = v(e^{itv}-1)$ that $\|R(\varphi f)\| \le C\|f\| \, \|\varphi\|_{\mathcal{L}^1(\mu)}$.

A similar calculation, which we sketch below, gives the same estimate for $\varphi = e^{itv}(e^{ihv}-1-ihv)$. Compute that $\|R(e^{itv}(e^{ihv}-1-ihv)f)\| = S_1 + S_2 + S_3$, where

$$\begin{aligned}
S_1 &= \sum_I \mu(I) \sup_I |e^{ihv} - 1 - ihv| \sup_I |f| \ll \|e^{ihv} - 1 - ihv\|_{\mathcal{L}^1(\mu)}, \\
S_2 &= \sum_I \mu(I) \sup_I |e^{itv}| \big(\operatorname{Var}_I(e^{ihv}) + |h| \operatorname{Var}_I(v)\big) \sup_I |f| \ll |h|, \\
S_3 &= \sum_I \mu(I) \sup_I |e^{itv}\big(e^{ihv} - 1 - ihv\big)| \operatorname{Var}_I(f) \ll \|e^{ihv} - 1 - ihv\|_{\mathcal{L}^1(\mu)}.
\end{aligned}$$

From here onwards, the proof goes exactly the same as in Proposition 4.22. □

4

Chapter 5

Error Rates for I.I.D. Cauchy Stable Large and Quasi-large Deviations

Abstract

We obtain error rates for large deviations of sums of i.i.d. random variables given by a Cauchy stable distribution. We treat a slightly reduced range of 1-stable large deviations and the full range of modified large deviations, which we call quasi-large deviations. As an application, we give the error rates for the large and quasi-large deviations of the arithmetic mean of digits coming from Lüroth series expansions.

5.1 Introduction

We recall the statement of the Cauchy ($\alpha = 1$) case of Theorem 1.8, which was proved recently in [Ber19]. Consider a sequence $(X_j)_{j\in\mathbb{N}}$ of i.i.d. random variables on the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$. Suppose that $(S_n - b_n)/a_n$ is in the domain of attraction of a Cauchy distribution, where $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are the scaling and centring sequences given in (1.9) by $a_n\big(1+o(1)\big) = n\ell(a_n)$ and $b_n = n\mathbb{E}(X_1 \mathbb{1}_{\{|X_1|\le a_n\}})$, $n \in \mathbb{N}$, for some slowly varying function ℓ, and recall from §1.4 that this is equivalent to having assumed that

$$\mathbb{P}(X_1 > x) = px^{-1}\ell(x)\big(1 + o(1)\big) \quad \text{and} \quad \mathbb{P}(X_1 < -x) = qx^{-1}\ell(x)\big(1 + o(1)\big), \tag{5.1}$$

where $p, q > 0$ with $p + q = 1$. Then for x such that x/a_n as $n \to +\infty$,

$$\mathbb{P}(S_n - b_n > x) = n\mathbb{P}(X_1 > x)\big(1 + o(1)\big), \quad \mathbb{P}(S_n - b_n < -x) = n\mathbb{P}(X_1 > x)\big(1 + o(1)\big). \tag{5.2}$$

In this chapter, we will be interested in obtaining error rates for the large deviations in (5.2). Due to the application to Lüroth series expansions (recall (1.1)) that we have in mind, we shall focus on the case with no first moment. We will also restrict to the easiest possible case, namely that $(X_j)_{j\in\mathbb{N}}$ is a sequence of i.i.d. random variables given by

$$\mathbb{P}(X_1 > x) = \sum_{j\in\mathbb{N}} \frac{c_j}{x^j}, \qquad x > 0, \tag{5.3}$$

where $(c_j)_{j\in\mathbb{N}}$ is a (summable) sequence of real constants with $c_1 > 0$, but all that is really needed is for the characteristic function of X_1 to have two derivatives; see §5.2.1. Given (5.3), it follows that $(S_n - b_n)/a_n$ is in the domain of attraction of a Cauchy distribution, where $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are as above, so (5.1) holds with $p = 1$, $q = 0$ and $\ell \equiv 1$. In fact, (5.3) implies for each $n \in \mathbb{N}$ that $a_n(1 + o(1)) = n$ and $b_n(1 + o(1)) = n \log n$.

For notational convenience throughout, write $N = N(x, n) := x + b_n$, and note that

$$\frac{N}{a_n} = \frac{N}{n}\big(1 + o(1)\big) \geq \big(1 + o(1)\big)\log n \to +\infty, \quad \text{as } n \to +\infty.$$

We obtain the following large deviations result.

Theorem 5.1. *Assume* (5.3)*. Then for any* $\mathfrak{p} \geq 2$ *and all* x *such that* $N = N(x, n) := x + b_n$ *satisfies* $n \log N = o(N)$ *as* $n \to +\infty$*,*

$$\mathbb{P}(S_n - b_n > x) = n\mathbb{P}(X_1 > N) + o\left(\frac{n^2(\log N)^2 + N^{2/\mathfrak{p}}}{N^2}\right).$$

Obtaining error rates for the large deviations in (5.2) ($p = 1, q = 0, \ell \equiv 1$) comes at the cost of reducing the range of x for which the result is optimal. In order to obtain error rates for the full range of x, i.e. such that $x/a_n \to +\infty$, we introduce an upper bound to the deviation of S_n considered. This gives the following result.

Theorem 5.2. *Assume* (5.3)*. Then as* $x/a_n \to +\infty$*,*

$$\mathbb{P}\big(S_n - b_n \in [x, x + g(n)]\big) = n\mathbb{P}\big(X_1 \in [N, N + g(n)]\big) + o\left(\frac{n^2(\log n)^2 + N^{2/\mathfrak{p}}}{N^2}\right)$$

for any $\mathfrak{p} \geq 2$ *and any choice of function* $g : \mathbb{R} \to (0, +\infty)$ *with* $g(n) \to +\infty$ *as* $n \to +\infty$ *such that*

$$\frac{g(n)(\log N)^2}{N} = o\big((\log n)^2\big) \qquad \text{and} \qquad \frac{ng(n)}{N + g(n)} = o\big((\log n)^2\big) \qquad \text{as } n \to +\infty. \tag{5.4}$$

We refer to the result in Theorem 5.2, i.e. a deviation of centred partial sums $S_n - b_n$ of the form $S_n - b_n \in [x, x + g(n)]$, where $g : \mathbb{R} \to (0, +\infty)$ satisfies $g(n) \to +\infty$ and $g(n) = o(N)$ as $n \to +\infty$, as a *quasi-large deviation*. Note that the condition $g(n) = o(N)$ prevents a quasi-large deviation from being a genuine large deviation. Note also that a quasi-large deviation differs from a local large deviation because the interval $[x, x + g(n)]$ considered here is large, given that $g(n) \to +\infty$, which is in contrast with the intervals $[x, x + h]$, $h > 0$ fixed, that are considered for local large deviations.

We then consider the following application of Theorems 5.1 and 5.2. Recall the definition of the Lüroth map $L : [0, 1] \to [0, 1]$ in (1.2) (shown below in Figure 5.1) and that every $x \in [0, 1] \setminus \mathbb{Q}$ has a unique associated Lüroth series expansion of the form

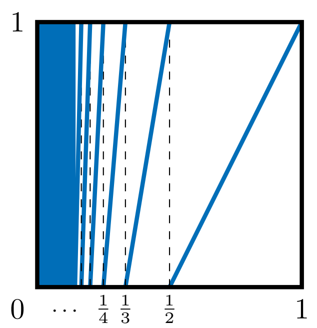


Figure 5.1: The Lüroth map

$$x = \sum_{n\in\mathbb{N}} \frac{d_n(x)}{\prod_{1\le\ell\le n} d_\ell(x)(d_\ell(x)+1)},$$

where $d_n(x) \coloneqq k$ whenever $L^{n-1}(x) \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$ for $n \in \mathbb{N}$. Let λ be the Lebesgue measure restricted to the unit interval. In this setting, Theorems 5.1 and 5.2 give the following corollary.

Corollary 5.3. *Fix $y \in [0,1] \setminus \mathbb{Q}$ and consider the arithmetic average*

$$S_n(y) \coloneqq \frac{d_1(y) + \cdots + d_n(y)}{n} = \frac{1}{n} \sum_{1\le\ell\le n} d_1 \circ L^{\ell-1}(y)$$

of the Lüroth digits of y. Then the following statements hold for any $\mathfrak{p} \ge 2$.

(i) *For all x such that $N \coloneqq x + b_n$ satisfies $n \log N = o(N)$ as $n \to +\infty$,*

$$\lambda\big(S_n(y) - b_n \ge x\big) = \frac{n}{N} + o\left(\frac{n^2(\log N)^2 + N^{2/\mathfrak{p}}}{N^2}\right).$$

(ii) *As $x/a_n \to +\infty$,*

$$\lambda\big(S_n(y) - b_n \in [x, x+g(n)]\big) = \frac{ng(n)}{N(N+g(n))} + o\left(\frac{n^2(\log n)^2 + N^{2/\mathfrak{p}}}{N^2}\right)$$

for any choice of function $g : \mathbb{R} \to (0,+\infty)$ satisfying (5.4).

The outline of this chapter is the following. In §5.2, we give the main strategy of the proofs of Theorems 5.1 and 5.2, which will follow the same method up to the end of §5.3. More precisely, we show in §5.2.1 that Theorems 5.1 and 5.2 follow once we expand a certain integral (I in (5.5) below) involving the characteristic function of S_n and a function g. In §5.2.2, we determine the precise form of the characteristic function given (5.3), and we derive estimates for related quantities in §5.2.3. In §5.3, we integrate the integral obtained in §5.2.1 by parts and split it into more manageable sub-integrals. Each is computed up to certain quantities that depend on the choice of g. It is at this point that the argument takes two paths; in §5.4, we take g in the range of large deviations, namely with $g(n) \gg N$, and complete the proof of Theorem 5.1, while, in §5.5, we take g in the range of quasi-large deviations, namely as in (5.4), and complete the proof of Theorem 5.2. In §5.6, we apply Theorems 5.1 and 5.2 to the digit functions associated with Lüroth series expansions to obtain Corollary 5.3. We also obtain a generalisation of this result to GLS expansions under a certain condition on the associated branch-support partition.

5.2 Technical Ingredients of the Proofs

To prove Theorems 5.1 and 5.2, we will use the same method as in §4.3.1, but we must use slightly different techniques in order to get error rates. Namely, we first use the Fourier inversion formula for probability measures (see e.g. [Bil12, Theorem 26.2]) to translate the statements of Theorems 5.1 and 5.2, which are on the measure $\mathbb{P}$, into statements on the characteristic function $\Psi(t) \coloneqq \int_0^{+\infty} e^{itx} \,\mathrm{d}\mathbb{P}(X_1 \le x)$ of X_1. We show during §5.2.1 that the conclusions of Theorems 5.1 and 5.2 follow once we obtain the power expansion in N up to second order terms of the integral

5

$$I^* \coloneqq \int_{-\varepsilon}^{+\varepsilon} \big(e^{-itN} - e^{-it(N+g(n))}\big) \frac{\Psi(t)^n - n\Psi(t)}{it} \Psi_Y(t)\,\mathrm{d}t,$$

where $\varepsilon > 0$ is a constant and Y is a random variable to be introduced later with a characteristic function $\Psi_Y(t)$ satisfying convenient properties that allow us to restrict the domain of integration to a neighbourhood of zero; §5.2.1. The main difficulty in expanding the integral I^* comes from the fact that the factor of t^{-1} in the integrand blows up. To overcome this, we rewrite $\Psi(t)^n - n\Psi(t)$ as a multiple of $\Psi(t) - 1$ plus a quantity that is independent of $\Psi(t)$, which splits I^* into two parts. Using the proof of [AD98, Theorem 2] and the form of $\mathbb{P}(X_1 > x)$ given in (5.3), it follows that

$$\Psi(t) - 1 = b\,\mathrm{sgn}(t)t\big(1 + o(1)\big) + i\big(-ct\log|t|\big)\big(1 + o(1)\big), \qquad t \in [-\varepsilon, \varepsilon],$$

for constants $b, c > 0$, which has a factor of t, so we can use $\Psi(t) - 1$ in the first part of I^* to counteract the effect of t^{-1}. We can show that the second part of I^* is negligible by using the Fourier inversion formula again.

The next step is to expand the first part of I^* using integration by parts and a modulus-of-continuity argument as in [Kat76, Chapter 1] to produce factors of N^{-1}. We expand up to quantities whose behaviour depends on the range of $g(n)$; these quantities can be found in (5.29) and are treated in §5.4 (resp. §5.5) for g in the range of large (resp. quasi-large) deviations. The expansion can be found in Proposition 5.9.

In order to keep the power of n at most two throughout the calculations to follow, we use Lemma 4.21, e.g. [MT22, Lemma 2.3], which we restate here for $\alpha = 1$ (although not stated in [MT22] for $\alpha = 1$, the proof there generalises immediately to this case).

Lemma 5.4 ([MT22, Lemma 2.3])**.** *Let $\ell : (0, +\infty) \to (0, +\infty)$ be a continuous slowly varying function. Fix $0 \le b < a \le \varepsilon$, $n \in \mathbb{N}$ and $\delta > 0$. Then for any $\rho \ge 0$,*

$$\int_b^a t^\rho \ell(t^{-1})|\Psi(t)|^n \,\mathrm{d}t = \mathcal{O}\big(n^{-(\rho+1)}\ell(n)\big).$$

5.2.1 Rephrasing the Problem in Terms of Characteristic Functions

We translate the statements of Theorems 5.1 and 5.2 into statements on the characteristic function $\Psi(t)$ by using the method in §4.3.1 but recording the error terms along the way. We will start with the following two lemmas, the first of which is an analogue of Lemma 4.7. Recall that $N := x + b_n$ is a function of n, and x is taken in the range $x/a_n \to +\infty$ as $n \to +\infty$.

Lemma 5.5. *Let $(Z_j)_{j\in\mathbb{N}}$ be a sequence of random variables on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$. Fix $\mathfrak{p} \ge 2$, and let Y be a centred random variable independent of the Z_j that is in $\mathcal{L}^{\mathfrak{p}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ and has a real-valued even characteristic function $\Psi_Y \in \mathcal{C}^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ supported in $[-\varepsilon, \varepsilon]$ for some small constant $\varepsilon > 0$. Assume that there is a function $\kappa : \mathbb{R} \to (0, +\infty)$ with $\kappa(n) = o(N^2)$ and $\kappa(n) \to +\infty$ as $n \to +\infty$ such that $\mathbb{P}\big(Z_n > N \pm \kappa(n)\big) = \mathbb{P}(Z_n > N) + o\big(\kappa(n)/N^2\big)$. Then as $n \to +\infty$,*

$$\mathbb{P}(Z_n > N) = \mathbb{P}(Z_n + Y > N) + o\left(\frac{1}{\kappa(n)^{\mathfrak{p}}} + \frac{\kappa(n)}{N^2}\right).$$

Proof. Conditioning on Y gives

$$\begin{aligned}\mathbb{P}\big(Z_n > N + \kappa(n)\big) &= \mathbb{P}\big(Z_n > N + \kappa(n) \,\big|\, |Y| > \kappa(n)\big) + \mathbb{P}\big(Z_n > N + \kappa(n) \,\big|\, |Y| \le \kappa(n)\big) \\ &\le \mathbb{P}\big(|Y| > \kappa(n)\big) + \mathbb{P}(Z_n + Y > N).\end{aligned}$$

Since $Y \in \mathcal{L}^{\mathfrak{p}}$, it follows by the assumption on κ that

$$\mathbb{P}(Z_n + Y > N) \ge \mathbb{P}\big(Z_n > N + \kappa(n)\big) - \mathbb{P}\big(|Y| > \kappa(n)\big) = \mathbb{P}(Z_n > N) + o\left(\frac{1}{\kappa(n)^{\mathfrak{p}}} + \frac{\kappa(n)}{N^2}\right).$$

Similarly,

$$\begin{aligned}\mathbb{P}\big(Z_n \le N - \kappa(n)\big) &= \mathbb{P}\big(Z_n \le N - \kappa(n) \,\big|\, |Y| > \kappa(n)\big) + \mathbb{P}\big(Z_n \le N - \kappa(n) \,\big|\, |Y| \le \kappa(n)\big) \\ &\le \mathbb{P}\big(|Y| > \kappa(n)\big) + \mathbb{P}(Z_n + Y \le N)\end{aligned}$$

and so

$$\begin{aligned}\mathbb{P}(Z_n + Y > N) = 1 - \mathbb{P}(Z_n + Y \le N) &\le 1 - \mathbb{P}\big(Z_n \le N - \kappa(n)\big) + \mathbb{P}\big(|Y| > \kappa(n)\big) \\ &= \mathbb{P}\big(Z_n > N - \kappa(n)\big) + \mathbb{P}\big(|Y| > \kappa(n)\big) \\ &= \mathbb{P}(Z_n > N) + o\left(\frac{1}{\kappa(n)^{\mathfrak{p}}} + \frac{\kappa(n)}{N^2}\right).\end{aligned}$$

□

The existence of a random variable Y as in Lemma 5.5 follows for every $\mathfrak{p} \ge 2$ from [Gou10a, Proposition 3.8]. Throughout the rest of the chapter, we fix $\mathfrak{p} \ge 2$ and such a random variable $Y \in \mathcal{L}^{\mathfrak{p}}$ that is independent of the X_j, $j \in \mathbb{N}$. Write $\Psi_Y(t)$, $t \in [-\varepsilon, \varepsilon]$, for its characteristic function.

Lemma 5.6. *Let $g : \mathbb{R} \to (0, +\infty)$ be such that $g(n) \to +\infty$ as $n \to +\infty$. Then*

$$\mathbb{P}\big(S_n \in [N, N + g(n)]\big) = \mathbb{P}\big(S_n + Y \in [N, N + g(n)]\big) + o\left(\frac{N^{2/\mathfrak{p}}}{N^2}\right).$$

Furthermore, if we additionally assume that $g(n) \ge N^2$, then

$$\mathbb{P}(S_n > N) = \mathbb{P}\big(S_n + Y \in [N, N + g(n)]\big) + o\left(\frac{N^{2/\mathfrak{p}}}{N^2}\right).$$

Proof. First, note for any choice of positive function κ that (5.3) gives

$$\begin{aligned}\mathbb{P}(S_n > N) - \mathbb{P}\big(S_n > N + \kappa(n)\big) &= \sum_{j \in \mathbb{N}} c_j \left(\frac{1}{N^j} - \frac{1}{(N + \kappa(n))^j}\right) = \sum_{j \in \mathbb{N}} c_j \frac{(N + \kappa(n))^j - N^j}{N^j (N + \kappa(n))^j} \\ &= \frac{\kappa(n)}{N(N + \kappa(n))} \sum_{j \in \mathbb{N}} \sum_{0 \le k < j} \frac{c_j}{N^k (N + \kappa(n))^{j-1-k}} \\ &= \frac{\kappa(n)}{N^2}\big(1 + o(1)\big)\end{aligned}$$

and similarly $\mathbb{P}(S_n > N) - \mathbb{P}\big(S_n > N - \kappa(n)\big) = \kappa(n) N^{-2}\big(1 + o(1)\big)$. Additionally, since $Y \in \mathcal{L}^{\mathfrak{p}}$, we know that $\mathbb{P}\big(|Y| > \kappa(n)\big) = o\big(\kappa(n)^{-\mathfrak{p}}\big)$. Choose $\kappa(n) := N^{2/(\mathfrak{p}+1)}$. Then it follows that $\mathbb{P}\big(S_n > N \pm \kappa(n)\big) = \mathbb{P}(S_n > N) + o\big(N^{-(2-2/\mathfrak{p})}\big)$ and $\mathbb{P}\big(|Y| > \kappa(n)\big) = o\big(N^{-(2-2/\mathfrak{p})}\big)$ so we may apply Lemma 5.5 to see that

$$\begin{aligned}\mathbb{P}\big(S_n \in [N, N + g(n)]\big) &= \mathbb{P}(S_n \ge N) - \mathbb{P}\big(S_n > N + g(n)\big) \\ &= \mathbb{P}(S_n + Y \ge N) - \mathbb{P}\big(S_n + Y > N + g(n)\big) + o\big(N^{-(2-2/\mathfrak{p})}\big) \\ &= \mathbb{P}\big(S_n + Y \in [N, N + g(n)]\big) + o\big(N^{-(2-2/\mathfrak{p})}\big),\end{aligned}$$

which completes the proof of the first statement. To prove the second statement, note that $g(n) \ge N^2$ and (5.3) imply that

$$\mathbb{P}\big(S_n > N + g(n)\big) \le n\mathbb{P}\left(X_1 > \frac{N + g(n)}{n}\right) = \frac{c_1 n^2}{N + g(n)}\big(1 + o(1)\big) = \mathcal{O}\left(\frac{n^2}{N^2}\right).$$

5

Consequently,

$$\mathbb{P}\big(S_n \in [N, N+g(n)]\big) = \mathbb{P}(S_n \geq N) - \mathbb{P}\big(S_n > N+g(n)\big) = \mathbb{P}(S_n > N) + \mathcal{O}\bigg(\frac{n^2}{N^2}\bigg)$$

and so the second statement follows from the first. □

Fix $n \in \mathbb{N}$, and let $g : \mathbb{R} \to (0, +\infty)$ be an increasing function with $g(n) \to +\infty$ as $n \to +\infty$ to be specified later (in §5.4, we will take g as in Theorem 5.1 to get large deviations, and we will take g as in Theorem 5.2 to give quasi-large deviations in §5.5). In the next proposition, we reduce proving Theorems 5.1 and 5.2 to analysing a certain integral written in terms of characteristic functions, which is analogous to the result in Proposition 4.8 but catered towards obtaining error rates.

Proposition 5.7. *Let $g : \mathbb{R} \to (0, +\infty)$ be such that $g(n) \to +\infty$ as $n \to +\infty$. Fix $\mathfrak{p} \geq 2$, and let $Y \in \mathcal{L}^{\mathfrak{p}}$ be as in Lemma 5.5. Put*

$$I := \int_{-\varepsilon}^{+\varepsilon} \big(e^{-itN} - e^{-it(N+g(n))}\big) \frac{\Psi(t)-1}{it} \Psi_Y(t) F(t,n)\,\mathrm{d}t. \tag{5.5}$$

Then

$$\mathbb{P}\big(S_n \in [N, N+g(n)]\big) = n\mathbb{P}\big(X_1 \in [N, N+g(n)]\big) + \frac{I}{2\pi} + o\bigg(\frac{N^{2/\mathfrak{p}}}{N^2}\bigg).$$

Furthermore, if we additionally assume that $g(n) \geq N^2$, then

$$\mathbb{P}(S_n > N) = n\mathbb{P}(X_1 > N) + \frac{I}{2\pi} + o\bigg(\frac{N^{2/\mathfrak{p}}}{N^2}\bigg).$$

Proof. Since the characteristic function of $X_1 + Y$ is $\Psi(t)\Psi_Y(t)$ and the characteristic function of $S_n + Y$ is $\Psi(t)^n\Psi_Y(t)$, then, using that the X_j, $j \in \mathbb{N}$, are i.i.d. random variables, the Fourier inversion formula yields

$$\begin{aligned}
\mathbb{P}\big(X_1 + Y \in [N, N+g(n)]\big) &= \lim_{a\to+\infty} \frac{1}{2\pi} \int_{-a}^{a} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi(t)\Psi_Y(t)\,\mathrm{d}t,\\
\mathbb{P}\big(S_n + Y \in [N, N+g(n)]\big) &= \lim_{a\to+\infty} \frac{1}{2\pi} \int_{-a}^{a} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi(t)^n\Psi_Y(t)\,\mathrm{d}t.
\end{aligned} \tag{5.6}$$

The purpose of having introduced the random variable Y is to adjust the domain of integration in (5.6) and remove the limit in a. Indeed, since the characteristic function Ψ_Y of Y has support $[-\varepsilon, \varepsilon]$, we see that (5.6) simplifies to give

$$\begin{aligned}
\mathbb{P}\big(X_1 + Y \in [N, N+g(n)]\big) &= \frac{1}{2\pi} \int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi_Y(t)\Psi(t)\,\mathrm{d}t,\\
\mathbb{P}\big(S_n + Y \in [N, N+g(n)]\big) &= \frac{1}{2\pi} \int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it} \Psi_Y(t)\Psi(t)^n\,\mathrm{d}t.
\end{aligned} \tag{5.7}$$

Next, observe that

$$\begin{aligned}
\Psi(t)^n - n\Psi(t) &= \big(\Psi(t)^{n-1} - 1\big)\Psi(t) - (n-1)\Psi(t) = \big(\Psi(t) - 1\big)\sum_{k=1}^{n-1} \Psi(t)^k - (n-1)\Psi(t)\\
&= \big(\Psi(t) - 1\big)\sum_{k=1}^{n-1} \big(\Psi(t)^k - 1\big) + (n-1)\big(\Psi(t) - 1\big) - (n-1)\Psi(t)\\
&= \big(\Psi(t) - 1\big)\sum_{k=1}^{n-1} \big(\Psi(t)^k - 1\big) - (n-1).
\end{aligned} \tag{5.8}$$

5

so, by setting

$$F(t,n) := \sum_{k=1}^{n-1} \big(\Psi(t)^k - 1\big), \tag{5.9}$$

it follows from (5.7) and (5.8) that

$$\begin{aligned}
&\mathbb{P}\big(S_n + Y \in [N, N+g(n)]\big) - n\mathbb{P}\big(X_1 + Y \in [N, N+g(n)]\big) \\
&= \frac{1}{2\pi}\int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it}\Psi_Y(t)\big(\Psi(t)^n - n\Psi(t)\big)\,\mathrm{d}t \\
&= \frac{I}{2\pi} - \frac{n-1}{2\pi}\int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it}\Psi_Y(t)\,\mathrm{d}t. \qquad (5.10)
\end{aligned}$$

Applying the Fourier inversion formula again and using that $Y \in \mathcal{L}^2$ gives

$$\begin{aligned}
\frac{n-1}{2\pi}\int_{-\varepsilon}^{+\varepsilon} \frac{e^{-itN} - e^{-it(N+g(n))}}{it}\Psi_Y(t)\,\mathrm{d}t &= (n-1)\lim_{a\to+\infty}\frac{1}{2\pi}\int_{-a}^{+a} \frac{e^{-itN} - e^{-it(N+g(n))}}{it}\Psi_Y(t)\,\mathrm{d}t \\
&= (n-1)\mathbb{P}\big(Y \in [N, N+g(n)]\big) \\
&\le n\mathbb{P}(Y \ge N) = \mathcal{O}\Big(\frac{n}{N^2}\Big), \qquad (5.11)
\end{aligned}$$

Substituting (5.11) into (5.10) gives

$$\mathbb{P}\big(S_n + Y \in [N, N+g(n)]\big) - n\mathbb{P}\big(X_1 + Y \in [N, N+g(n)]\big) = \frac{I}{2\pi} + \mathcal{O}\Big(\frac{n}{N^2}\Big),$$

and the proof concludes by applying Lemma 5.6. □

Hence, to prove Theorems 5.1 and 5.2, it suffices to estimate the integral I for g in the range prescribed by large deviations and quasi-large deviations respectively.

5.2.2 Form of the Characteristic Function

Before we can expand the integral I defined in (5.5), we will need to determine the exact form of the characteristic function $\Psi(t)$, $t \in [-\varepsilon, \varepsilon]$.

Lemma 5.8. *Assume* (5.3)*. Then the characteristic function* $\Psi(t)$ *of* X_1 *is given by*

$$\Psi(t) = 1 + Cit + \sum_{r\in\mathbb{N}} t^r \Psi_r(t), \qquad t \in [-\varepsilon, \varepsilon],$$

where $C := \int_0^1 e^{itx}\mathbb{P}(X_1 > x)\,\mathrm{d}x \ge 0$ *is a constant and* $\Psi_r(t) := \frac{1}{(r-1)!}\big(\zeta_r + \xi_r e^{it} - c_r i^r \log|t|\big)$ *for bounded constants* $\zeta_r, \xi_r \in \mathbb{C}$.

5

An immediate consequence of Lemma 5.8 is that $\Psi(t)$ is of the form $\Psi(t) = 1 - c_1 it\log|t| + tf(t)$ with

$$f(t) := Ci + \zeta_1 + \xi_1 e^{it} + \sum_{r\in\mathbb{N}} t^r \Psi_{r+1}(t) = A + \mathcal{O}\big(\big|t\log|t|\big|\big), \tag{5.12}$$

where $A := Ci + \zeta_1 + \xi_1 \in \mathbb{C}$.

Proof of Lemma 5.8. Briefly following the method at the beginning of [AD98, Theorem 2], integration by parts gives

$$\Psi(t) = 1 + \int_0^{+\infty} (e^{itx} - 1)\,\mathrm{d}\mathbb{P}(X_1 \le x) = 1 + it\int_0^{+\infty} e^{itx}\mathbb{P}(X_1 > x)\,\mathrm{d}x$$
$$= 1 + it\int_0^1 e^{itx}\mathbb{P}(X_1 > x)\,\mathrm{d}x + it\int_1^{+\infty} e^{itx}\sum_{j\in\mathbb{N}}\frac{c_j}{x^j}\,\mathrm{d}x = 1 + Cit + it\sum_{j\in\mathbb{N}} c_j\int_1^{+\infty}\frac{e^{itx}}{x^j}\,\mathrm{d}x.$$

For each $j \in \mathbb{N}$, integration by parts $j-1$ times yields

$$\int_1^{+\infty}\frac{e^{itx}}{x^j}\,\mathrm{d}x = \sum_{1\le k<j}\frac{e^{it}(it)^{j-1-k}}{(j-1)\cdots\big(j-(j-k)\big)t^k} + \frac{(it)^{j-1}}{(j-1)!}\int_1^{+\infty}\frac{e^{itx}}{x}\,\mathrm{d}x$$
$$= \frac{(it)^{j-1}}{(j-1)!}\left(e^{it}\sum_{1\le k<j}\frac{(k-1)!}{(it)^k} + \int_1^{+\infty}\frac{e^{itx}}{x}\,\mathrm{d}x\right).$$

It is known (e.g. see [AS64]) that the exponential integral $E_1(z)$ satisfies

$$E_1(z) := \int_1^{+\infty}\frac{e^{-xz}}{x}\,\mathrm{d}x = -\operatorname{Log} z - \gamma - \sum_{k\in\mathbb{N}}\frac{(-z)^k}{k!k}$$

for all $z \in \mathbb{C}$ with $\operatorname{Re} z \ge 0$ and $\arg z \in (-\pi, \pi]$, where $\operatorname{Log} z$ is the principal branch of the complex logarithm and γ is the Euler-Mascheroni constant. Thus,

$$\int_1^{+\infty}\frac{e^{itx}}{x}\,\mathrm{d}x = E_1(-it) = -\log|t| - \gamma - \frac{i\pi}{2} - \sum_{k\in\mathbb{N}}\frac{(it)^k}{k!\,k}.$$

Consequently,

$$\Psi(t) - 1 - Cit = \sum_{j\in\mathbb{N}}\frac{c_j(it)^j}{(j-1)!}\left(e^{it}\sum_{1\le k<j}\frac{(k-1)!}{(it)^k} - \log|t| - \gamma - \frac{i\pi}{2} - \sum_{k\in\mathbb{N}}\frac{(it)^k}{k!\,k}\right). \quad (5.13)$$

Observe that each infinite series in (5.13) is absolutely convergent so we may change the order of summation. Then by collecting like-terms in (5.13), we see for each $r \in \mathbb{N}$ that

$$\zeta_r := -c_r i^r\left(\gamma + \frac{i\pi}{2}\right) - (r-1)!\sum_{\substack{j,k\in\mathbb{N}\\ j+k=r}}\frac{c_j i^r}{(j-1)!k!\,k} \quad\text{and}\quad \xi_r := (r-1)!\sum_{\substack{j,k\in\mathbb{N}\\ j-k=r}}\frac{c_j i^r(k-1)!}{(j-1)!}$$

yield $\Psi_r(t) = \frac{1}{(r-1)!}(\zeta_r + \xi_r e^{it} - c_r i^r\log|t|)$ as the r^{th} coefficient in the power expansion of $\Psi(t) - 1 - Cit$. □

5.2.3 Estimates for Quantities Involving the Characteristic Function

In this section, we shall collect various estimates for quantities involving the characteristic function that will be needed later. Fix $0 < |t| \le \varepsilon$ and $r \in \mathbb{N}$. From the definition of Ψ_r in Lemma 5.8, it is immediate that

$$\Psi_r(t) = -\frac{c_r i^r\log|t|}{(r-1)!} + \mathcal{O}(1). \quad (5.14)$$

The derivative of $\log|t|$ is t^{-1}, which blows up at $t=0$ in a controlled way. Thus, we have that

$$\begin{aligned}\Psi_r'(t) &= \frac{i\xi_r e^{it} - c_r i^r t^{-1}}{(r-1)!} = \mathcal{O}\left(\frac{|t|^{-1}}{(r-1)!}\right),\\ \Psi_r''(t) &= \frac{-\xi_r e^{it} + c_r i^r t^{-2}}{(r-1)!} = \mathcal{O}\left(\frac{t^{-2}}{(r-1)!}\right).\end{aligned} \tag{5.15}$$

Recall the definition of $f(t)$ from (5.12). By (5.14) and (5.15),

$$\begin{aligned}f'(t) &= \xi_1 i e^{it} + \sum_{r\in\mathbb{N}}\big(rt^{r-1}\Psi_{r+1}(t) + t^r\Psi_{r+1}'(t)\big) = \Psi_2(t) + \mathcal{O}(1) = \frac{c_2}{2}\log|t| + \mathcal{O}(1),\\ f''(t) &= -\xi_1 e^{it} + \sum_{r\in\mathbb{N}}\big(r(r-1)t^{r-2}\Psi_{r+1}(t) + 2rt^{r-1}\Psi_{r+1}'(t) + t^r\Psi_{r+1}''(t)\big) = \mathcal{O}(|t|^{-1}).\end{aligned} \tag{5.16}$$

Since $\Psi(t) = 1 - c_1 it\log|t| + tf(t)$, (5.12) and (5.16) yield

$$\begin{aligned}\Psi'(t) &= -c_1 i\big(1+\log|t|\big) + f(t) + tf'(t) = -c_1 i\log|t| + B + \mathcal{O}\big(\big|t\log|t|\big|\big),\\ \Psi''(t) &= -c_1 it^{-1} + 2f'(t) + tf''(t) = -c_1 it^{-1} + \mathcal{O}(\log|t|),\end{aligned} \tag{5.17}$$

where $B := (C-c_1)i + \zeta_1 + \xi_1 \in \mathbb{C}$. Recall the definition of $F(t,n)$ from (5.9). Factorising and using that $\Psi(t) - 1 = \mathcal{O}\big(\big|t\log|t|\big|\big)$, we find that

$$F(t,n) = \big(\Psi(t)-1\big)\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\Psi(t)^k = \mathcal{O}(n^2\big|t\log|t|\big|). \tag{5.18}$$

Differentiating the expression for $F(t,n)$ in (5.9) and applying (5.17), we obtain

$$F'(t,n) = \Psi'(t)\sum_{j=1}^{n-1} j\Psi(t)^{j-1} = \big(-c_1 i\log|t| + B\big)\sum_{j=1}^{n-1} j\Psi(t)^{j-1} + \mathcal{O}\big(n^2\big|t\log|t|\big|\big). \tag{5.19}$$

In §5.3.3, we will need an estimate for $F'(t,n) - F'(t-h,n)$, $0<h<|t|<\varepsilon$, such that any term including a logarithm is explicit; see (5.25) below. This requires several auxiliary estimates, which we shall derive first. It follows from the mean value theorem that there is a $\xi \in [t-h,t]$ such that $f'(t) - f'(t-h) = f''(\xi)h$ and so (5.16) yields

$$f'(t) - f'(t-h) = \mathcal{O}(h|t|^{-1}). \tag{5.20}$$

By the same reasoning but with (5.15) in place of (5.16), we have that

$$t^r\Psi_r(t) - (t-h)^r\Psi_r(t-h) = \mathcal{O}\big(h\big|t^{r-1}\log|t|\big|\big)$$

so it follows from (5.12), (5.14) and the definition of $\Psi_r(t)$ in Lemma 5.8 that

$$\begin{aligned}f(t) - f(t-h) &= \xi_1(e^{it} - e^{i(t-h)}) + t\big(\Psi_2(t) - \Psi_2(t-h)\big) + h\Psi_2(t-h) + \mathcal{O}(h|t\log|t||)\\ &= -\frac{c_2}{2}t\log\left(1-\frac{h}{t}\right) + \frac{c_2}{2}h\log|t-h| + \mathcal{O}(h)\\ &= \frac{c_2}{2}h\log|t-h| + \mathcal{O}(h).\end{aligned} \tag{5.21}$$

5

Thus, it follows from Lemma 5.8, the definition of f in (5.12) and (5.21) that

$$\begin{aligned}\Psi(t) - \Psi(t-h) &= c_1 i t \log\left(1 - \frac{h}{t}\right) + t\big(f(t) - f(t-h)\big) + h f(t-h) \\ &= c_1 i t \log\left(1 - \frac{h}{t}\right) + Ah + \mathcal{O}(h|t\log|t||),\end{aligned} \tag{5.22}$$

where $A := Ci + \zeta_1 + \xi_1 \in \mathbb{C}$. Consequently, we have

$$\begin{aligned}\Psi(t)^r - \Psi(t-h)^r &= \big(\Psi(t) - \Psi(t-h)\big) \sum_{j=0}^{r-1} \Psi(t)^j \Psi(t-h)^{r-1-j} \\ &= \left(c_1 i t \log\left(1 - \frac{h}{t}\right) + Ah\right) \sum_{j=0}^{r-1} \Psi(t)^j \Psi(t-h)^{r-1-j} \\ &\quad + \mathcal{O}\big(rh\big|t\log|t|\Psi(t)^{r-1}\big|\big).\end{aligned} \tag{5.23}$$

By (5.16), (5.20) and (5.21),

$$\begin{aligned}\Psi'(t) - \Psi'(t-h) &= c_1 i \log\left(1 - \frac{h}{t}\right) + f(t) - f(t-h) + t\big(f'(t) - f'(t-h)\big) + h f'(t-h) \\ &= c_1 i \log\left(1 - \frac{h}{t}\right) + c_2 h \log|t-h| + \mathcal{O}(h).\end{aligned} \tag{5.24}$$

Therefore, it follows by adding and subtracting terms, factorising and then applying (5.17), (5.23) and (5.24) that

$$\begin{aligned}&F'(t,n) - F'(t-h,n) \\ &\quad = \big(\Psi'(t) - \Psi'(t-h)\big) \sum_{j=1}^{n-1} j\Psi(t)^{j-1} + \Psi'(t-h) \sum_{j=1}^{n-1} j\big(\Psi(t)^{j-1} - \Psi(t-h)^{j-1}\big) \\ &\quad = \left(c_1 i \log\left(1 - \frac{h}{t}\right) + c_2 h \log|t-h|\right) \sum_{j=1}^{n-1} j\Psi(t)^{j-1} \\ &\qquad - c_1 i\left(c_1 i t \log\left(1 - \frac{h}{t}\right) + Ah\right) \log|t-h| \sum_{j=1}^{n-1}\sum_{k=0}^{j-2} j\Psi(t)^k \Psi(t-h)^{j-2-k}\big) \\ &\qquad + \mathcal{O}\left(h \sum_{j=1}^{n-1} j^2 |\Psi(t)|^{j-1}\right).\end{aligned} \tag{5.25}$$

Throughout §5.3, we will need estimates for the function $\Theta(t) := (\Psi(t) - 1)/it$ and its derivatives. Observe by Lemma 5.8 and (5.12) that

$$\Theta(t) = -c_1 \log|t| - if(t) = -c_1 \log|t| - Ai + \mathcal{O}\big(\big|t\log|t|\big|\big). \tag{5.26}$$

Next, observe by (5.16) that

$$\begin{aligned}\Theta'(t) &= -c_1 t^{-1} - if'(t) = -c_1 t^{-1} + \mathcal{O}(|\log|t||), \\ \Theta''(t) &= c_1 t^{-2} - if''(t) = c_1 t^{-2} + \mathcal{O}(|t|^{-1}).\end{aligned} \tag{5.27}$$

In addition, it follows from (5.22) and (5.26) that

$$\begin{aligned}\Theta(t)-\Theta(t-h) &= \frac{\Psi(t)-1}{it}-\frac{\Psi(t-h)-1}{i(t-h)} = \frac{\Psi(t)-\Psi(t-h)}{it}-ht^{-1}\Theta(t-h)\\ &= c_1\log\left(1-\frac{h}{t}\right)+c_1ht^{-1}\log|t-h|+\mathcal{O}(h|\log|t||). \qquad (5.28)\end{aligned}$$

We will also use the following facts, which will allow us to ignore Ψ_Y and $\Psi_Y'(t)$ when integrating in §5.3. Since $\Psi_Y(0)=1$ and $\Psi_Y\in\mathcal{C}^2(\mathbb{R},\mathcal{B}(\mathbb{R}),\mathbb{P})$, we have for any $-\varepsilon\le a<b\le\varepsilon$ and any function y that is integrable on $[a,b]$,

$$\int_a^b y(t)\Psi_Y(t)\,\mathrm{d}t=\int_a^b y(t)\,\mathrm{d}t+\int_a^b y(t)\big(\Psi_Y(t)-\Psi_Y(0)\big)\,\mathrm{d}t=\int_a^b y(t)\,\mathrm{d}t+\mathcal{O}\bigg(\int_a^b|ty(t)|\,\mathrm{d}t\bigg). \quad (\dagger)$$

In addition, since $\Psi_Y'(0)=\mathbb{E}Y=0$ and $\Psi_Y'\in\mathcal{C}^1(\mathbb{R},\mathcal{B}(\mathbb{R}),\mathbb{P})$, we also have

$$\int_a^b y(t)\Psi_Y'(t)\,\mathrm{d}t=\int_a^b y(t)\big(\Psi_Y'(t)-\Psi_Y'(0)\big)\,\mathrm{d}t=\mathcal{O}\bigg(\int_a^b|ty(t)|\,\mathrm{d}t\bigg). \quad (\ddagger)$$

5.3 Expanding the Integral I

Recall that $\Theta(t):=(\Psi(t)-1)/it$ and $F(t,n):=\sum_{j=1}^{n-1}(\Psi(t)^j-1)$. In this section, we will prove Proposition 5.9 below, which expands the integral I in (5.5) up to quantities that are asymptotically equivalent (or can be treated analogously) to one of the following:

$$\begin{aligned}\mathcal{X}_1 &:= \Big[\big(e^{-itN}-e^{-it(N+g(n))}\big)\Theta'(t)F(t,n)\Big]_{-h}^{h},\\ \mathcal{X}_2 &:= \int_h^\varepsilon\big(e^{-itN}-e^{-it(N+g(n))}\big)t^{-1}(\log t)^2\Psi(t)^n\,\mathrm{d}t,\\ \mathcal{X}_3 &:= \int_0^h\big(e^{-itN}-e^{-it(N+g(n))}\big)(\log t)^2\Psi(t)^n\,\mathrm{d}t,\\ \mathcal{X}_4 &:= n\int_h^\varepsilon\big(e^{-itN}-e^{-it(N+g(n))}\big)(\log t)^2\Psi(t)^n\,\mathrm{d}t.\end{aligned} \qquad (5.29)$$

The estimates of these quantities are affected by the choice of g; it is these quantities and our choice of g that determines whether we get the partial-range large deviations in Theorem 5.1 (see Proposition 5.16) or the full-range quasi-large deviations in Theorem 5.2 (see Proposition 5.17).

Proposition 5.9. *Recall that $n\log n=o(N)$ as $n\to+\infty$, and let $g:\mathbb{N}\to(0,+\infty)$ be an increasing function such that $g(n)\to+\infty$ as $n\to+\infty$. Then as $n\to+\infty$,*

$$\begin{aligned}I &:= \int_{-\varepsilon}^{+\varepsilon}\big(e^{-itN}-e^{-it(N+g(n))}\big)\Theta(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t\\ &= \mathcal{O}\bigg(\frac{1}{N^2}|\mathcal{X}_1|+\frac{n^2}{N^2}|\mathcal{X}_2|+\frac{n^2}{N}|\mathcal{X}_3|+\frac{n^2}{N^2}|\mathcal{X}_4|\bigg)+\mathcal{O}\bigg(\frac{n^2\log n}{N^2}\bigg).\end{aligned}$$

Throughout the current section, we fix $N=N(n)$ such that $N/n\to+\infty$ as $n\to+\infty$ and $g:\mathbb{N}\to(0,+\infty)$ an increasing function such that $g(n)\to+\infty$ as $n\to+\infty$. To prove Proposition 5.9, we split I into two parts: $I=I(N)-I(N+g(n))$, where

$$\begin{aligned}I(N) &:= \int_{-\varepsilon}^{+\varepsilon}e^{-itN}\Theta(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t,\\ I\big(N+g(n)\big) &:= \int_{-\varepsilon}^{+\varepsilon}e^{-it(N+g(n))}\Theta(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t.\end{aligned}$$

Integrating $I(N)$ by parts gives

$$\begin{aligned} I(N) &= \left[\frac{ie^{-itN}}{N}\Theta(t)\Psi_Y(t)F(t,n)\right]_{-\varepsilon}^{+\varepsilon} + \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t \\ &\quad + \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta'(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t + \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t \\ &=: I_0(N) + I_1(N) + I_2(N) + I_3(N). \end{aligned} \tag{5.30}$$

In the same way, we can write

$$I(N+g(n)) = I_0(N+g(n)) + I_1(N+g(n)) + I_2(N+g(n)) + I_3(N+g(n))$$

by replacing N with $N+g(n)$ throughout (5.30). Put $I_j := I(N) - I(N+g(n))$ for each $j = 0,1,2,3$ so that $I = I_0 + I_1 + I_2 + I_3$. To proceed, we give the expansion of each term I_j as a lemma. I_0 and I_1 are simple and are treated in the next section in Lemmas 5.10 and 5.11. The integrals I_2 and I_3 are less straightforward so they will be computed up to the quantities in (5.29) in §5.3.2 and §5.3.3 respectively.

5.3.1 Expanding I_0 and the Sub-integral I_1 of I

Lemma 5.10. *For any $n \in \mathbb{N}$,*

$$I_0 := \left[\frac{ie^{-itN}}{N}\Theta(t)\Psi_Y(t)F(t,n)\right]_{-\varepsilon}^{+\varepsilon} - \left[\frac{ie^{-it(N+g(n))}}{N+g(n)}\Theta(t)\Psi_Y(t)F(t,n)\right]_{-\varepsilon}^{+\varepsilon} = 0.$$

Proof. Since Ψ_Y is $\mathcal{C}^2$, even and supported in $[-\varepsilon, \varepsilon]$, we have that

$$\Psi_Y(\pm\varepsilon) = \Psi_Y(\varepsilon) - \Psi_Y\left(\varepsilon + \frac{1}{N^a}\right) = \mathcal{O}\left(\frac{1}{N^a}\right)$$

for any $a > 0$. Thus, we must have $\Psi_Y(\pm\varepsilon) = 0$, whence the result follows. □

Lemma 5.11. *As $n \to +\infty$,*

$$\begin{aligned} I_1 &:= \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t - \frac{1}{i(N+g(n))}\int_{-\varepsilon}^{+\varepsilon} e^{-it(N+g(n))}\Theta(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t \\ &= \mathcal{O}\left(\frac{n^2}{N^2}\right). \end{aligned}$$

Proof. Integrating $I_1(N) := \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t$ by parts gives

$$\begin{aligned} I_1(N) &= \left[\frac{e^{-itN}}{N^2}\Theta(t)\Psi_Y'(t)F(t,n)\right]_{-\varepsilon}^{+\varepsilon} - \frac{1}{N^2}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta'(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t \\ &\quad - \frac{1}{N^2}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta(t)\Psi_Y''(t)F(t,n)\,\mathrm{d}t - \frac{1}{N^2}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta(t)\Psi_Y'(t)F'(t,n)\,\mathrm{d}t \\ &= \mathcal{O}\left(\frac{n^2}{N^2} + \frac{n^2}{N^2}\int_0^\varepsilon |\log t|\,\mathrm{d}t + \frac{n^2}{N^2}\int_0^\varepsilon t|\log t|\,\mathrm{d}t + \frac{n^2}{N^2}\int_0^\varepsilon |\log t|\,\mathrm{d}t\right) = \mathcal{O}\left(\frac{n^2}{N^2}\right). \end{aligned}$$

In the same way, we get

$$\begin{aligned} I_1\big(N+g(n)\big) &:= \frac{1}{i\big(N+g(n)\big)}\int_{-\varepsilon}^{+\varepsilon} e^{-it(N+g(n))}\Theta(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t \\ &= \mathcal{O}\left(\frac{n^2}{\big(N+g(n)\big)^2}\right) = \mathcal{O}\left(\frac{n^2}{N^2}\right), \end{aligned}$$

which completes the proof. □

5

5.3.2 Expanding the Sub-integral I_2 of I

Recall that $\Theta(t) := (\Psi(t) - 1)/it$ and $F(t,n) := \sum_{j=1}^{n-1}(\Psi(t)^j - 1)$. In this section, we treat the integral I_2 by decomposing it into quantities of type $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{X}_3$ (recall (5.29)).

Lemma 5.12. *As $n \to +\infty$,*

$$\begin{aligned} I_2 &:= \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta'(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t - \frac{1}{i(N+g(n))}\int_{-\varepsilon}^{+\varepsilon} e^{-it(N+g(n))}\Theta'(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t \\ &= \mathcal{O}\left(\frac{1}{N^2}|\mathcal{X}_1| + \frac{n^2}{N^2}|\mathcal{X}_2| + \frac{n^2}{N}|\mathcal{X}_3|\right) + \mathcal{O}\left(\frac{n^2}{N^2}\right). \end{aligned}$$

Proof. We begin by splitting the range of $I_2(N)$:

$$\begin{aligned} I_2(N) &= \frac{1}{iN}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t + \frac{1}{iN}\int_{-h}^{h} e^{-itN}\Theta'(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t \\ &= I_2^+(N) + I_2^{(0)}(N). \end{aligned}$$

We split the range of $I_2(N+g(n)) = I_2^+(N+g(n)) + I_2^{(0)}(N+g(n))$ in the same way so that

$$\begin{aligned} I_2 &:= I_2(N) - I_2(N+g(n)) \\ &= I_2^+(N) - I_2^+\big(N+g(n)\big) + I_2^{(0)}(N) - I_2^{(0)}\big(N+g(n)\big). \end{aligned} \tag{5.31}$$

<u>The range excluding zero</u>

We integrate the first range $I_2^+(N)$ of $I_2(N)$ by parts:

$$\begin{aligned} I_2^+(N) = &\left[\frac{e^{-itN}}{N^2}\Theta'(t)\Psi_Y(t)F(t,n)\right]_{[h,\varepsilon]\cup[-\varepsilon,-h]} \\ &- \frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta''(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t \\ &- \frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t \\ &- \frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t. \end{aligned} \tag{5.32}$$

By (5.18) and (5.27), we can apply (†) to the first term on the right-hand side of (5.32) to obtain

$$\left[\frac{e^{-itN}}{N^2}\Theta'(t)\Psi_Y(t)F(t,n)\right]_{[h,\varepsilon]\cup[-\varepsilon,-h]} = -\left[\frac{e^{-itN}}{N^2}\Theta'(t)F(t,n)\right]_{-h}^{h} + \mathcal{O}\left(\frac{n^2}{N^2}\right).$$

We will match this quantity with its analogue coming from $I_2^+(N+g(n))$ later, which will produce a term of type $\mathcal{X}_1$. For the second term on the right-hand side of (5.32), we apply (†) as we did for the first term to remove $\Psi_Y(t)$:

$$\frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta''(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t = \frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta''(t)F(t,n)\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right),$$

and then factorise $\Psi(t) - 1 = -c_1 it \log|t| + tf(t)$ from $F(t,n)$ to obtain

$$
\begin{aligned}
&-\frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta''(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t \\
&\quad= -\frac{c_1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-2}F(t,n)\,\mathrm{d}t + \mathcal{O}\biggl(\frac{n^2}{N^2}\biggr) \\
&\quad= \frac{c_1^2 i}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\log|t|\Psi(t)^k\,\mathrm{d}t \qquad (5.33) \\
&\qquad -\frac{c_1}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}f(t)\Psi(t)^k\,\mathrm{d}t + \mathcal{O}\biggl(\frac{n^2}{N^2}\biggr).
\end{aligned}
$$

We will match the first term on the right-hand side of (5.33) with its analogue coming from $I_2(N+g(n))$ later, which will produce an integral of type $\mathcal{X}_2$. By (5.12), the second term on the right-hand side of (5.33) is

$$
\begin{aligned}
&-\frac{c_1}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}f(t)\Psi(t)^k\,\mathrm{d}t \\
&\qquad= -\frac{c_1 A}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\Psi(t)^k\,\mathrm{d}t + \mathcal{O}\biggl(\frac{n^2}{N^2}\biggr),
\end{aligned}
$$

where $A \coloneqq Ci + \zeta_1 + \xi_1 \in \mathbb{C}$. Again, we match this term later to produce an integral of type $\mathcal{X}_2$. Next, $\Theta'(t) = \mathcal{O}(|t|^{-1})$ by (5.27) so, using (‡), the third term on the right-hand side of (5.32) is

$$
-\frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)\Psi_Y'(t)F(t,n)\,\mathrm{d}t = \mathcal{O}\biggl(\frac{n^2}{N^2}\int_h^\varepsilon t|\log t|\,\mathrm{d}t\biggr) = \mathcal{O}\biggl(\frac{n^2}{N^2}\biggr).
$$

For the fourth term on the right-hand side of (5.32), applying (†) and computing $F'(t,n)$ gives

$$
\begin{aligned}
&-\frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t \\
&\quad= -\frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)F'(t,n)\,\mathrm{d}t + \mathcal{O}\biggl(\frac{n^2}{N^2}\biggr) \\
&\quad= -\frac{1}{N^2}\sum_{j=1}^{n-1} j\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)\Psi'(t)\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\biggl(\frac{n^2}{N^2}\biggr).
\end{aligned}
$$

By (5.17) and (5.27), we find that

$$
\begin{aligned}
\Theta'(t)\Psi'(t) &= \bigl(-c_1 t^{-1} + \mathcal{O}(|\log|t||)\bigr)\bigl(-c_1 i\log|t| + B + \mathcal{O}\bigl(\bigl|t\log|t|\bigr|\bigr)\bigr) \\
&= c_1^2 it^{-1}\log|t| - c_1 Bt^{-1} + \mathcal{O}\bigl((\log|t|)^2\bigr),
\end{aligned}
$$

where $B \coloneqq (C - c_1)i + \zeta_1 + \xi_1$. Thus, the fourth term is

$$
\begin{aligned}
&-\frac{1}{N^2}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}\Theta'(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t \\
&\quad= -\frac{c_1^2 i}{N^2}\sum_{j=1}^{n-1} j\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\log|t|\Psi(t)^{j-1}\,\mathrm{d}t \\
&\qquad + \frac{c_1 B}{N^2}\sum_{j=1}^{n-1} j\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\biggl(\frac{n^2}{N^2}\biggr).
\end{aligned}
$$

Putting all of the above together therefore yields

$$\begin{aligned}
I_2^+(N) = &-\left[\frac{e^{-itN}}{N^2}\Theta'(t)F(t,n)\right]_{-h}^{h} + \frac{c_1^2 i}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\log|t|\Psi(t)^k\,\mathrm{d}t \\
&- \frac{c_1^2 i}{N^2}\sum_{j=1}^{n-1} j\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\log|t|\Psi(t)^{j-1}\,\mathrm{d}t \\
&- \frac{c_1 A}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\Psi(t)^k\,\mathrm{d}t \\
&+ \frac{c_1 B}{N^2}\sum_{j=1}^{n-1} j\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} e^{-itN}t^{-1}\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right).
\end{aligned}$$

We may deduce an analogous expansion for $I_2^+(N+g(n))$ by replacing N with $N+g(n)$ throughout; therefore, we find by adding and subtracting terms that

$$\begin{aligned}
&I_2^+(N) - I_2^+\big(N+g(n)\big) \\
&\quad= -\left[\frac{e^{-itN} - e^{-it(N+g(n))}}{N^2}\Theta'(t)F(t,n)\right]_{-h}^{h} \\
&\qquad+ \frac{c_1^2 i}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} \big(e^{-itN} - e^{-it(N+g(n))}\big)t^{-1}\log|t|\Psi(t)^k\,\mathrm{d}t \\
&\qquad- \frac{c_1^2 i}{N^2}\sum_{j=1}^{n-1} j\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} \big(e^{-itN} - e^{-it(N+g(n))}\big)t^{-1}\log|t|\Psi(t)^{j-1}\,\mathrm{d}t \\
&\qquad- \frac{c_1 A}{N^2}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} \big(e^{-itN} - e^{-it(N+g(n))}\big)t^{-1}\Psi(t)^k\,\mathrm{d}t \\
&\qquad+ \frac{c_1 B}{N^2}\sum_{j=1}^{n-1} j\int_{[h,\varepsilon]\cup[-\varepsilon,-h]} \big(e^{-itN} - e^{-it(N+g(n))}\big)t^{-1}\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right) \\
&\quad= \mathcal{O}\left(\frac{1}{N^2}|\mathcal{X}_1| + \frac{n^2}{N^2}|\mathcal{X}_2|\right) + \mathcal{O}\left(\frac{n^2}{N^2}\right). \qquad (5.34)
\end{aligned}$$

The range including zero

To treat the second range $I_2^{(0)}(N)$ of $I_2(N)$, first note that

$$\big(\Psi(t)-1\big)\Theta'(t) = \big(-c_1 it\log|t| + \mathcal{O}(|t|)\big)\big(-c_1 t^{-1} + \mathcal{O}(|\log|t||)\big) = c_1^2 i\log|t| + \mathcal{O}(1).$$

Consequently, applying (†) and then factorising $F(t,n)$ gives

$$\begin{aligned}
I_2^{(0)}(N) &:= \frac{1}{iN}\int_{-h}^{h} e^{-itN}\Theta'(t)\Psi_Y(t)F(t,n)\,\mathrm{d}t = \frac{1}{iN}\int_{-h}^{h} e^{-itN}\Theta'(t)F(t,n)\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2(\log N)^2}{N^3}\right) \\
&= \frac{1}{iN}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{-h}^{h} e^{-itN}\big(\Psi(t)-1\big)\Theta'(t)\Psi(t)^k\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2(\log N)^2}{N^3}\right) \\
&= \frac{c_1^2}{N}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{-h}^{h} e^{-itN}\log|t|\Psi(t)^k\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right).
\end{aligned}$$

We can deduce a similar expansion for $I_2^{(0)}(N+g(n))$. Therefore, by adding and subtracting terms, we get

$$\begin{aligned} I_2^{(0)}(N) - I_2^{(0)}\big(N+g(n)\big) &= \frac{c_1^2}{N}\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\int_{-h}^{h}\big(e^{-itN} - e^{-it(N+g(n))}\big)\log|t|\Psi(t)^k\,\mathrm{d}t + \mathcal{O}\Big(\frac{n^2}{N^2}\Big) \\ &= \mathcal{O}\Big(\frac{n^2}{N}|\mathcal{X}_3|\Big) + \mathcal{O}\Big(\frac{n^2}{N^2}\Big). \end{aligned} \tag{5.35}$$

The result follows from (5.31), (5.34) and (5.35). □

5.3.3 Expanding the Sub-integral I_3 of I

Recall that $\Theta(t) := (\Psi(t)-1)/it$ and $F(t,n) := \sum_{j=1}^{n-1}(\Psi(t)^j - 1)$. In this section, we expand the integral I_3 from (5.30).

Proposition 5.13. *As $n \to +\infty$,*

$$I_3 := \frac{1}{iN}\int_{-\varepsilon}^{+\varepsilon} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t = \mathcal{O}\Big(\frac{n^2}{N^2}|\mathcal{X}_2| + \frac{n^2}{N}|\mathcal{X}_3| + \frac{n^2}{N^2}|\mathcal{X}_4|\Big) + \mathcal{O}\Big(\frac{n^2\log n}{N^2}\Big).$$

In order to prove Proposition 5.13, so as not to introduce too many factors of n, we must not differentiate $F'(t,n)$. Consequently, we cannot integrate $I_3(N)$ by parts as we have done with $I_1(N)$ and $I_2(N)$ to obtain the second factor of N^{-1} needed. Instead, we use a modulus-of-continuity argument to produce this factor. For this, we must first split the range of $I_3(N)$:

$$I_3(N) = \frac{1}{iN}\Big(\int_h^{\varepsilon} + \int_{-h}^{h} + \int_{-\varepsilon}^{-h}\Big)e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t =: J_+(N) + J_0(N) + J_-(N).$$

We treat the integral $J_+(N)$ in the next lemma. The calculations used there can be used to deduce the expansions of $J_0(N)$ and $J_-(N)$, which we record in Corollary 5.15 below.

Lemma 5.14. *As $n \to +\infty$,*

$$J_+ = \mathcal{O}\Big(\frac{n^2}{N^2}|\mathcal{X}_2| + \frac{n^2}{N}|\mathcal{X}_3| + \frac{n^2}{N^2}|\mathcal{X}_4|\Big) + o\Big(\frac{n^2(\log n)^2}{N^2}\Big).$$

Proof. The change of coordinates $t \mapsto t-h$ yields

$$J_+(N) := \frac{1}{iN}\int_h^{\varepsilon} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t = -\frac{1}{iN}\int_{2h}^{\varepsilon+h} e^{-itN}\Theta(t-h)\Psi_Y(t-h)F'(t-h,n)\,\mathrm{d}t$$

and so

$$\begin{aligned} 2J_+(N) &= \frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\big(\Theta(t)\Psi_Y(t)F'(t,n) - \Theta(t-h)\Psi_Y(t-h)F'(t-h,n)\big)\,\mathrm{d}t \\ &\quad + \frac{1}{iN}\int_h^{2h} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t + \frac{1}{iN}\int_{\varepsilon-h}^{\varepsilon} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t \\ &=: J_1(N) + J_2(N) + J_3(N). \end{aligned}$$

The integral J_1:

By adding and subtracting terms and using that $\Psi_Y \in \mathcal{C}^2$, we find that

$$\begin{aligned}
&\Theta(t)\Psi_Y(t)F'(t,n) - \Theta(t-h)\Psi_Y(t-h)F'(t-h,n) \\
&\quad = \big(\Theta(t) - \Theta(t-h)\big)\Psi_Y(t)F'(t,n) + \Theta(t-h)\big(\Psi_Y(t) - \Psi_Y(t-h)\big)F'(t-h,n) \\
&\qquad + \Theta(t-h)\Psi_Y(t-h)\big(F'(t,n) - F'(t-h,n)\big) \\
&\quad = \big(\Theta(t) - \Theta(t-h)\big)\Psi_Y(t)F'(t,n) + \Theta(t-h)\Psi_Y(t-h)\big(F'(t,n) - F'(t-h,n)\big) \\
&\qquad + \mathcal{O}\big(n^2 h (\log|t|)^2\big).
\end{aligned}$$

Consequently,

$$\begin{aligned}
J_1(N) &:= \frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\big(\Theta(t)\Psi_Y(t)F'(t,n) - \Theta(t-h)\Psi_Y(t-h)F'(t-h,n)\big)\,\mathrm{d}t \\
&= \frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\big(\Theta(t) - \Theta(t-h)\big)\Psi_Y(t)F'(t,n)\,\mathrm{d}t \\
&\quad + \frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\Theta(t-h)\Psi_Y(t-h)\big(F'(t,n) - F'(t-h,n)\big)\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right).
\end{aligned} \tag{5.36}$$

First, we find by (5.19) and (5.28) that

$$\begin{aligned}
&\big(\Theta(t) - \Theta(t-h)\big)F'(t,n) \\
&\quad = c_1\left(\log\left(1 - \frac{h}{t}\right) + ht^{-1}\log|t-h|\right)\big(-c_1 i\log|t| + B\big)\sum_{j=1}^{n-1} j\Psi(t)^{j-1} + \mathcal{O}\big(n^2 h(\log|t|)^2\big),
\end{aligned}$$

where $B := (C - c_1)i + \zeta_1 + \xi_1 \in \mathbb{C}$. Then (†) implies that the first of the integrals on the right-hand side of (5.36) is

$$\begin{aligned}
&\frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\big(\Theta(t) - \Theta(t-h)\big)\Psi_Y(t)F'(t,n)\,\mathrm{d}t \\
&\quad = \frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\big(\Theta(t) - \Theta(t-h)\big)F'(t,n)\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right) \\
&\quad = \frac{c_1^2}{N}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon} e^{-itN}\left(\log\left(1 - \frac{h}{t}\right) + ht^{-1}\log(t-h)\right)\big(-c_1 i\log t + B\big)\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right).
\end{aligned}$$

We pair this integral with the corresponding integral coming from $J_1(N+g(n))$ later to produce an integral that is asymptotically equivalent to $\mathcal{X}_2$. For the second integral on the right-hand side of (5.36), we first apply (†) to remove $\Psi_Y(t-h)$:

$$\begin{aligned}
&\frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\Theta(t-h)\Psi_Y(t-h)\big(F'(t,n) - F'(t-h,n)\big)\,\mathrm{d}t \\
&\quad = \frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\Theta(t-h)\big(F'(t,n) - F'(t-h,n)\big)\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2}{N^2}\right)
\end{aligned}$$

5

We find by (5.25) and (5.26) that

$$\begin{aligned}
&\Theta(t-h)\big(F'(t,n)-F'(t-h,n)\big)\\
&\quad=\big(-c_1\log|t-h|-Ai+\mathcal{O}(|t\log|t||)\big)\\
&\qquad\times\Bigg(\bigg(c_1 i\log\Big(1-\frac{h}{t}\Big)+c_2 h\log|t-h|\bigg)\sum_{j=1}^{n-1} j\Psi(t)^{j-1}\\
&\qquad-c_1 i\bigg(c_1 it\log\Big(1-\frac{h}{t}\Big)+Ah\bigg)\log|t-h|\sum_{j=2}^{n-1}\sum_{k=0}^{j-2} j\Psi(t)^k\Psi(t-h)^{j-2-k}\\
&\qquad+\mathcal{O}\bigg(h\sum_{j=2}^{n-1} j^2|\Psi(t)|^{j-2}\bigg)\Bigg)\\
&\quad=\bigg(c_1\big(-c_1 i\log|t-h|+A\big)\log\Big(1-\frac{h}{t}\Big)-c_1c_2h\big(\log|t-h|\big)^2\bigg)\sum_{j=1}^{n-1} j\Psi(t)^{j-1}\\
&\qquad+c_1^2 i\bigg(c_1 it\log\Big(1-\frac{h}{t}\Big)+Ah\bigg)\big(\log|t-h|\big)^2\sum_{j=2}^{n-1}\sum_{k=0}^{j-2} j\Psi(t)^k\Psi(t-h)^{j-2-k}\\
&\qquad+\mathcal{O}\bigg(h|\log|t||\sum_{j=1}^{n-1} j^2|\Psi(t)|^{j-1}\bigg),
\end{aligned}$$

where $A \coloneqq Ci+\zeta_1+\xi_1\in\mathbb{C}$, so the second integral on the right-hand side of (5.36) is

$$\begin{aligned}
&\frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\Theta(t-h)\Psi_Y(t-h)\big(F'(t,n)-F'(t-h,n)\big)\,\mathrm{d}t\\
&\quad=\frac{c_1}{iN}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon} e^{-itN}\log\Big(1-\frac{h}{t}\Big)\big(-c_1 i\log(t-h)+A\big)\Psi(t)^{j-1}\,\mathrm{d}t\\
&\qquad-\frac{c_1c_2h}{iN}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon} e^{-itN}\big(\log(t-h)\big)^2\Psi(t)^{j-1}\,\mathrm{d}t\\
&\qquad+\frac{c_1^2}{N}\sum_{j=1}^{n-1}\sum_{k=0}^{j-2} j\int_{2h}^{\varepsilon} e^{-itN}\bigg(c_1 it\log\Big(1-\frac{h}{t}\Big)+Ah\bigg)\big(\log(t-h)\big)^2\Psi(t)^{j-1}\,\mathrm{d}t\\
&\qquad+\mathcal{O}\bigg(\frac{1}{N^2}\sum_{j=1}^{n-1} j^2\int_{2h}^{\varepsilon}|\log t||\Psi(t)|^{j-1}\,\mathrm{d}t+\frac{n^2}{N^2}\bigg).
\end{aligned}$$

We match the first and third terms with their counterparts coming from the expansion of $J_1\big(N+g(n)\big)$ later; the former will be asymptotic to $\mathcal{X}_2$ and the latter will be asymptotic $\mathcal{X}_4$. By Lemma 5.4 with $\ell(x)=(\log x)^2$, the second term is

$$\begin{aligned}
-\frac{c_1c_2h}{iN}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon} e^{-itN}\big(\log(t-h)\big)^2\Psi(t)^{j-1}\,\mathrm{d}t&=\mathcal{O}\bigg(\frac{1}{N^2}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon}(\log t)^2|\Psi(t)|^{j-1}\,\mathrm{d}t\bigg)\\
&=\mathcal{O}\bigg(\frac{1}{N^2}\sum_{j=1}^{n-1}(\log j)^2\bigg)=\mathcal{O}\bigg(\frac{n(\log n)^2}{N^2}\bigg),
\end{aligned}$$

5

and the error term simplifies, this time by Lemma 5.4 with $\ell(x) = |\log x|$, to give

$$\mathcal{O}\left(\frac{1}{N^2}\sum_{j=1}^{n-1} j^2 \int_{2h}^{\varepsilon} |\log t||\Psi(t)|^{j-1}\,\mathrm{d}t + \frac{n^2}{N^2}\right) = \mathcal{O}\left(\frac{1}{N^2}\sum_{j=1}^{n-1} j\log j + \frac{n^2}{N^2}\right) = \mathcal{O}\left(\frac{n^2\log n}{N^2}\right).$$

Therefore, the second integral on the right-hand side of (5.36) is

$$\begin{aligned}
&\frac{1}{iN}\int_{2h}^{\varepsilon} e^{-itN}\Theta(t-h)\Psi_Y(t-h)\big(F'(t,n) - F'(t-h,n)\big)\,\mathrm{d}t \\
&\quad= \frac{c_1}{iN}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon} e^{-itN}\log\left(1-\frac{h}{t}\right)\big(-c_1 i\log(t-h) + A\big)\Psi(t)^{j-1}\,\mathrm{d}t \\
&\qquad+ \frac{c_1^2}{N}\sum_{j=1}^{n-1}\sum_{k=0}^{j-2} j\int_{2h}^{\varepsilon} e^{-itN}\left(c_1 it\log\left(1-\frac{h}{t}\right) + Ah\right)\big(\log(t-h)\big)^2\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\left(\frac{n^2\log n}{N^2}\right).
\end{aligned}$$

Altogether, we have shown that

$$\begin{aligned}
J_1(N) = {} & \frac{c_1^2}{N}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon} e^{-itN}\left(\log\left(1-\frac{h}{t}\right) + ht^{-1}\log(t-h)\right)\big(-c_1 i\log t + B\big)\Psi(t)^{j-1}\,\mathrm{d}t \\
& + \frac{c_1}{iN}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon} e^{-itN}\log\left(1-\frac{h}{t}\right)\big(-c_1 i\log(t-h) + A\big)\Psi(t)^{j-1}\,\mathrm{d}t \\
& + \frac{c_1^2}{N}\sum_{j=1}^{n-1}\sum_{k=0}^{j-2} j\int_{2h}^{\varepsilon} e^{-itN}\left(c_1 it\log\left(1-\frac{h}{t}\right) + Ah\right)\big(\log(t-h)\big)^2\Psi(t)^{j-1}\,\mathrm{d}t \\
& + \mathcal{O}\left(\frac{n^2\log n}{N^2}\right)
\end{aligned}$$

We can find an analogous expansion for $J_1\big(N+g(n)\big)$ so, since $\log(1-h/t) = h\big(1+o(1)\big)/t$, adding and subtracting terms to $J_1 = J_1(N) - J_1\big(N+g(n)\big)$ yields

$$\begin{aligned}
J_1 = {} & \frac{c_1^2}{N}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon}\big(e^{-itN} - e^{it(N+g(n))}\big)\left(\log\left(1-\frac{h}{t}\right) + ht^{-1}\log(t-h)\right)\big(B - c_1 i\log t\big)\Psi(t)^{j-1}\,\mathrm{d}t \\
& + \frac{c_1}{iN}\sum_{j=1}^{n-1} j\int_{2h}^{\varepsilon}\big(e^{-itN} - e^{it(N+g(n))}\big)\log\left(1-\frac{h}{t}\right)\big(A - c_1 i\log(t-h)\big)\Psi(t)^{j-1}\,\mathrm{d}t \\
& + \frac{c_1^2}{N}\sum_{j=1}^{n-1}\sum_{k=0}^{j-2} j\int_{2h}^{\varepsilon}\big(e^{-itN} - e^{it(N+g(n))}\big)\left(c_1 it\log\left(1-\frac{h}{t}\right) + Ah\right)\big(\log(t-h)\big)^2\Psi(t)^{j-1}\,\mathrm{d}t \\
& + \mathcal{O}\left(\frac{n^2\log n}{N^2}\right) \\
= {} & \mathcal{O}\left(\frac{n^2}{N^2}|\mathcal{X}_2| + \frac{n^2}{N^2}|\mathcal{X}_4|\right) + \mathcal{O}\left(\frac{n^2\log n}{N^2}\right).
\end{aligned}$$

The integral J_2:

By (5.19) and (5.26), we find that

$$\Theta(t)F'(t,n) = \big(-c_1\log|t| - Ai + \mathcal{O}\big(\big|t\log|t|\big|\big)\big)\big(-c_1 i\log|t| + B + \mathcal{O}\big(\big|t\log|t|\big|\big)\big)\sum_{j=1}^{n-1} j\Psi(t)^{j-1}$$
$$= c_1\big(c_1 i(\log|t|)^2 - (A+B)\log|t| + \mathcal{O}(1)\big)\sum_{j=1}^{n-1} j\Psi(t)^{j-1}.$$

Thus, (†) yields

$$J_2(N) := \frac{1}{iN}\int_h^{2h} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t = \frac{1}{iN}\int_h^{2h} e^{-itN}\Theta(t)F'(t,n)\,\mathrm{d}t + \mathcal{O}\Big(\frac{n^2(\log N)^2}{N^4}\Big)$$
$$= \frac{c_1}{iN}\sum_{j=1}^{n-1} j\int_h^{2h} e^{-itN}\big(c_1 i(\log t)^2 - (A+B)\log t\big)\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\Big(\frac{n^2}{N^2}\Big).$$

In the same way, we can find an analogous expansion for $J_2\big(N+g(n)\big)$ and so

$$J_2 = J_2(N) - J_2\big(N+g(n)\big)$$
$$= \frac{c_1}{iN}\sum_{j=1}^{n-1} j\int_h^{2h}\big(e^{-itN} - e^{-it(N+g(n))}\big)\big(c_1 i(\log t)^2 - (A+B)\log t\big)\Psi(t)^{j-1}\,\mathrm{d}t + \mathcal{O}\Big(\frac{n^2}{N^2}\Big)$$
$$= \mathcal{O}\Big(\frac{n^2}{N}|\mathcal{X}_3|\Big) + \mathcal{O}\Big(\frac{n^2}{N^2}\Big).$$

The integral J_3:

We find by (5.19) and (5.26) that

$$J_3(N) := \frac{1}{iN}\int_{\varepsilon-h}^{\varepsilon} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t = \mathcal{O}\Big(\frac{n^2}{N}\int_{\varepsilon-h}^{\varepsilon}(\log t)^2\,\mathrm{d}t\Big) = \mathcal{O}\Big(\frac{n^2}{N^2}\Big).$$

Similarly, $J_3\big(N+g(n)\big) = \mathcal{O}(n^2/N^2)$ and so $J_3 = J_3(N) - J_3\big(N+g(n)\big) = \mathcal{O}(n^2/N^2)$.
Putting the estimates for J_1, J_2 and J_3 together gives the result. □

Corollary 5.15. *As $n \to +\infty$,*

$$J_- := \frac{1}{iN}\int_{-\varepsilon}^{-h} e^{-itN}\Theta(t)\Psi_Y(t)F'(t,n)\,\mathrm{d}t = \mathcal{O}\Big(\frac{n^2}{N^2}|\mathcal{X}_2| + \frac{n^2}{N}|\mathcal{X}_3| + \frac{n^2}{N^2}|\mathcal{X}_4|\Big) + \mathcal{O}\Big(\frac{n^2\log n}{N^2}\Big)$$

and

$$J_0 = \mathcal{O}\Big(\frac{n^2}{N}|\mathcal{X}_3|\Big) + \mathcal{O}\Big(\frac{n^2}{N^2}\Big).$$

Proof. The estimate for J_- can be deduced from J_+ by tracking the sign changes throughout the proof of Lemma 5.14. The estimate for J_0 follows by the same reasoning as is used to treat J_2 in Lemma 5.14. □

Proof of Proposition 5.13. This follows from Lemma 5.14 and Corollary 5.15. □

Proof of Proposition 5.9. This follows from Lemma 5.10, Lemma 5.11, Lemma 5.12 and Proposition 5.13. □

5.4 Large Deviations: Proving Theorem 5.1

In this section, we prove Theorem 5.1. Given what we have shown so far in §5.2.1 and §5.3, it remains to prove the following proposition concerning the quantities from (5.29).

Proposition 5.16. *Fix $n \in \mathbb{N}$, and let g be a positive multiple of $2\pi N^2$. Put $h := \pi/N$. Then*

(i) $\mathcal{X}_1 := \left[\left(e^{-itN} - e^{-it(N+g(n))}\right)\Theta'(t)F(t,n)\right]_{-h}^{h} = 0.$

(ii) $\mathcal{X}_2 := \int_h^{\varepsilon}(e^{-itN} - e^{-it(N+g(n))})t^{-1}(\log t)^2\Psi(t)^n \,\mathrm{d}t = o\big((\log N)^2\big).$

(iii) $\mathcal{X}_3 := \int_0^{h}(e^{-itN} - e^{-it(N+g(n))})(\log t)^2\Psi(t)^n \,\mathrm{d}t = o\big(N^{-1}(\log N)^2\big).$

(iv) $\mathcal{X}_4 := n\int_h^{\varepsilon}\left(e^{-itN} - e^{-it(N+g(n))}\right)(\log t)^2\Psi(t)^n \,\mathrm{d}t = \mathcal{O}\big((\log n)^2\big).$

Proof.

(i) Since $g(n)$ is a multiple of $2\pi N$, we have $e^{\pm ihN} - e^{\pm ih(N+g(n))} = 0$ and so $\mathcal{X}_1 = 0$.

(ii) By the change of variables $\sigma = tN$, we find that

$$\begin{aligned}\mathcal{X}_2 &= (\log N)^2\int_{\pi}^{\varepsilon N}\frac{e^{-i\sigma} - e^{-i\sigma(1+g(n)/N)}}{\sigma}\left(\frac{\log\sigma/N}{\log N}\right)^2\Psi\left(\frac{\sigma}{N}\right)^n \mathrm{d}\sigma\\ &= (\log N)^2\int_{\pi}^{\varepsilon N}\frac{e^{-i\sigma}(1 - e^{-i\sigma g(n)/N})}{\sigma}\left(\frac{\log\sigma/N}{\log N}\right)^2\Psi\left(\frac{\sigma}{N}\right)^n \mathrm{d}\sigma.\end{aligned}$$

By Potter's bounds (see e.g. [BGT87]), we have for any $0 < \delta < 1$ that

$$\left|\frac{e^{-i\sigma}(1 - e^{-i\sigma g(n)/N})}{\sigma}\left(\frac{\log\sigma/N}{\log N}\right)^2\Psi\left(\frac{\sigma}{N}\right)^n\right| \le 2e^{-k\sigma/N}\sigma^{-(1-\delta)},$$

which is integrable over $[\pi, +\infty]$ so, since $\left|\frac{\log\sigma/N}{\log N}\right| \to 1$ and $g(n)/N \to +\infty$ as $n \to +\infty$, we may apply the dominated convergence theorem to see that

$$\begin{aligned}&\int_{\pi}^{\varepsilon N}\frac{e^{-i\sigma}(1 - e^{-i\sigma\frac{g(n)}{N}})}{\sigma}\frac{\log\sigma/N}{\log N}\Psi\left(\frac{\sigma}{N}\right)^n \mathrm{d}\sigma\\ &\quad= \big(1 + o(1)\big)\int_{\pi}^{+\infty}\frac{e^{-i\sigma}(1 - e^{-i\sigma\frac{g(n)}{N}})}{\sigma}\frac{\log\sigma/N}{\log N}\Psi\left(\frac{\sigma}{N}\right)^n \mathrm{d}\sigma = o(1).\end{aligned}$$

Therefore, $\mathcal{X}_2 = o\big((\log N)^2\big)$.

(iii) This follows using the same method as in (ii).

(iv) This estimate follows from Lemma 5.4 with $\ell(x) = (\log x)^2$. □

Proof of Theorem 5.1. Choose $g : \mathbb{R} \to (0, +\infty)$ to be a positive multiple of $2\pi N^2$. Then putting Proposition 5.9 and Proposition 5.16 together yields

$$I = \mathcal{O}\left(\frac{1}{N^2}|\mathcal{X}_1| + \frac{n^2}{N^2}|\mathcal{X}_2| + \frac{n^2}{N}|\mathcal{X}_3| + \frac{n^2}{N^2}|\mathcal{X}_4|\right) + \mathcal{O}\left(\frac{n^2\log n}{N^2}\right) = o\left(\frac{n^2(\log N)^2}{N^2}\right),$$

Therefore, we have by Proposition 5.7 that

$$\mathbb{P}(S_n > N) = n\mathbb{P}(X_1 > N) + o\left(\frac{n^2(\log N)^2 + N^{2/\mathfrak{p}}}{N^2}\right).$$ □

5

5.5 Quasi-large Deviations: Proving Theorem 5.2

In this section, we prove Theorem 5.2. Given what we have shown so far in §5.2.1 and §5.3, it remains to prove the following proposition concerning the quantities from (5.29).

Proposition 5.17. *Recall that $n\log n = o(N)$. Let $g : \mathbb{N} \to (0,+\infty)$ be an increasing function satisfying* (5.4) *and $g(n) \to +\infty$ as $n \to +\infty$. Put $h := \pi/N$. Then*

(i) $\mathcal{X}_1 := \big[\big(e^{-itN} - e^{-it(N+g(n))}\big)\Theta'(t)F(t,n)\big]_{-h}^{h} = o\big(n^2(\log n)^2\big)$.

(ii) $\mathcal{X}_2 := \int_h^\varepsilon (e^{-itN} - e^{-it(N+g(n))})t^{-1}(\log t)^2\Psi(t)^n\,\mathrm{d}t = o\big((\log n)^2\big)$.

(iii) $\mathcal{X}_3 := \int_0^h (e^{-itN} - e^{-it(N+g(n))})(\log t)^2\Psi(t)^n\,\mathrm{d}t = o\big((\log n)^2N^{-1}\big)$.

(iv) $\mathcal{X}_4 := n\int_h^\varepsilon \big(e^{-itN} - e^{-it(N+g(n))}\big)(\log t)^2\Psi(t)^n\,\mathrm{d}t = o\big((\log n)^2\big)$.

Proof.

(i) Observe that

$$\Theta'(h) - \Theta'(-h) = -c_1\big(h^{-1} - (-h)^{-1}\big) - i\big(f'(h) - f'(-h)\big) = \mathcal{O}(N) \qquad \text{and}$$

$$F(h,n) - F(-h,n) = \big(\Psi(h) - \Psi(-h)\big)\sum_{j=1}^{n-1}\sum_{k=0}^{j-1}\Psi(h)^k\Psi(-h)^{j-1-k} = \mathcal{O}\left(\frac{n^2\log N}{N}\right)$$

so we have

$$\begin{aligned}
\mathcal{X}_1 &= -\big[\big(e^{-itN} - e^{-it(N+g(n))}\big)\Theta'(t)F(t,n)\big]_{-h}^{h} + \mathcal{O}(n^2)\\
&= \big(1 - e^{-ihg(n)}\big)\Theta'(h)F(h,n) - \big(1 - e^{ihg(n)}\big)\Theta'(-h)F(-h,n) + \mathcal{O}(n^2)\\
&= \big(e^{ihg(n)} - e^{-ihg(n)}\big)\Theta'(h)F(h,n) + \big(1 - e^{ihg(n)}\big)\big(\Theta'(h) - \Theta'(-h)\big)F(h,n)\\
&\qquad + \big(1 - e^{ihg(n)}\big)\Theta'(-h)\big(F(h,n) - F(-h,n)\big) + \mathcal{O}(n^2)\\
&= \mathcal{O}\left(hg(n)\cdot N\cdot\frac{n^2\log N}{N}\right) = \mathcal{O}\left(\frac{n^2 g(n)\log N}{N}\right) = o\big(n^2(\log n)^2\big)
\end{aligned}$$

by the choice of $g(n)$.

(ii) We first split $\mathcal{X}_2$ into two parts:

$$\mathcal{X}_2 = \int_h^\varepsilon e^{-itN}t^{-1}(\log t)^2\Psi(t)^n\,\mathrm{d}t - \int_h^\varepsilon e^{-it(N+g(n))}t^{-1}(\log t)^2\Psi(t)^n\,\mathrm{d}t.$$

Using the change of variables $\sigma = tN$ on the first part, $\sigma = t(N+g(n))$ on the second and then adding and subtracting terms yields

$$\begin{aligned}
\mathcal{X}_2 &= \int_{hN}^{\varepsilon N} e^{-i\sigma}\sigma^{-1}\left(\log\frac{\sigma}{N}\right)^2\Psi\left(\frac{\sigma}{N}\right)^n\mathrm{d}\sigma\\
&\quad - \int_{h(N+g(n))}^{\varepsilon(N+g(n))} e^{-i\sigma}\sigma^{-1}\left(\log\frac{\sigma}{N+g(n)}\right)^2\Psi\left(\frac{\sigma}{N+g(n)}\right)^n\mathrm{d}\sigma\\
&= \int_\pi^{\varepsilon N} e^{-i\sigma}\sigma^{-1}\left(\left(\log\frac{\sigma}{N}\right)^2 - \left(\log\frac{\sigma}{N+g(n)}\right)^2\right)\Psi\left(\frac{\sigma}{N+g(n)}\right)^n\mathrm{d}\sigma \qquad (5.37)\\
&\quad + \int_\pi^{\varepsilon N} e^{-i\sigma}\sigma^{-1}\left(\log\frac{\sigma}{N}\right)^2\left(\Psi\left(\frac{\sigma}{N}\right)^n - \Psi\left(\frac{\sigma}{N+g(n)}\right)^n\right)\mathrm{d}\sigma\\
&\quad + \left(\int_{\varepsilon N}^{\varepsilon(N+g(n))} - \int_{hN}^{h(N+g(n))}\right)e^{-i\sigma}\sigma^{-1}\left(\log\frac{\sigma}{N+g(n)}\right)^2\Psi\left(\frac{\sigma}{N+g(n)}\right)^n\mathrm{d}\sigma.
\end{aligned}$$

Since

$$\Big(\log\frac{\sigma}{N}\Big)^2-\Big(\log\frac{\sigma}{N+g(n)}\Big)^2=\log\Big(1+\frac{g(n)}{N}\Big)\log\frac{\sigma^2}{N(N+g(n))}=\mathcal{O}\Big(\frac{g(n)}{N}\log\frac{\sigma}{N}\Big),$$

using that $g(n)=o(N)$ to expand the logarithm, the first of the terms on the right-hand side of (5.37) is

$$\int_\pi^{\varepsilon N}e^{-i\sigma}\sigma^{-1}\bigg(\Big(\log\frac{\sigma}{N}\Big)^2-\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\bigg)\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma$$
$$=\mathcal{O}\bigg(\frac{g(n)}{N}\int_\pi^{\varepsilon N}\sigma^{-1}\log\frac{\sigma}{N}\,\mathrm{d}\sigma\bigg)=\mathcal{O}\bigg(\frac{g(n)(\log N)^2}{N}\bigg)=o\big((\log n)^2\big). \tag{5.38}$$

Next, we find after undoing the change of coordinates and applying (5.23) that

$$\int_\pi^{\varepsilon N}e^{-i\sigma}\sigma^{-1}\Big(\log\frac{\sigma}{N}\Big)^2\bigg(\Psi\Big(\frac{\sigma}{N}\Big)^n-\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\bigg)d\sigma$$
$$=\int_h^{\varepsilon}e^{-itN}t^{-1}(\log t)^2\bigg(\Psi(t)^n-\Psi\Big(t-\frac{tg(n)}{N+g(n)}\Big)^n\bigg)\,\mathrm{d}t=\mathcal{O}\bigg(\frac{ng(n)}{N+g(n)}\int_h^\varepsilon(\log t)^2\,\mathrm{d}t\bigg)$$
$$=\mathcal{O}\bigg(\frac{ng(n)}{N+g(n)}\bigg)=o\big((\log n)^2\big). \tag{5.39}$$

Last, we have

$$\bigg(\int_{\varepsilon N}^{\varepsilon(N+g(n))}-\int_{hN}^{h(N+g(n))}\bigg)e^{-i\sigma}\sigma^{-1}\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma$$
$$=\mathcal{O}\bigg(\varepsilon g(n)\cdot(\varepsilon N)^{-1}\Big(\log\frac{N+g(n)}{\varepsilon N}\Big)^2+hg(n)\cdot(hN)^{-1}\Big(\log\frac{N+g(n)}{hN}\Big)^2\bigg)$$
$$=\mathcal{O}\bigg(\frac{g(n)(\log N)^2}{N}\bigg)=o\big((\log n)^2\big). \tag{5.40}$$

Putting (5.37), (5.38), (5.39) and (5.40) together gives (ii).

(iii) As in (ii), we split the integral $\mathcal{X}_3$ into two parts and then use the change of variables $\sigma=tN$ on the first part and $\sigma=t(N+g(n))$ on the second part:

$$\begin{aligned}\mathcal{X}_3=&\frac{1}{N}\int_0^{hN}e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\Psi\Big(\frac{\sigma}{N}\Big)^n\,\mathrm{d}\sigma\\&-\frac{1}{N+g(n)}\int_0^{h(N+g(n))}e^{-i\sigma}\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma\\=&\frac{g(n)}{N(N+g(n))}\int_0^\pi e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\Psi\Big(\frac{\sigma}{N}\Big)^n\,\mathrm{d}\sigma\\&+\frac{1}{N+g(n)}\int_0^\pi e^{-i\sigma}\bigg(\Big(\log\frac{\sigma}{N}\Big)^2-\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\bigg)\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma\\&+\frac{1}{N+g(n)}\int_0^\pi e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\bigg(\Psi\Big(\frac{\sigma}{N}\Big)^n-\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\bigg)d\sigma\\&+\frac{1}{N+g(n)}\int_{hN}^{h(N+g(n))}e^{-i\sigma}\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma.\end{aligned}$$

The first integral is

$$\frac{g(n)}{N(N+g(n))}\int_0^\pi e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\Psi\Big(\frac{\sigma}{N}\Big)^n\,\mathrm{d}\sigma=\mathcal{O}\bigg(\frac{g(n)(\log N)^2}{N(N+g(n))}\bigg)=o\bigg(\frac{(\log n)^2}{N}\bigg).$$

Following the derivation of (5.38), the second integral is

$$\begin{aligned}&\frac{1}{N+g(n)}\int_0^\pi e^{-i\sigma}\bigg(\Big(\log\frac{\sigma}{N}\Big)^2-\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\bigg)\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma\\&\quad=\mathcal{O}\bigg(\frac{g(n)}{N(N+g(n))}\int_0^\pi\log\frac{\sigma}{N}\,\mathrm{d}\sigma\bigg)=\mathcal{O}\bigg(\frac{g(n)\log N}{N(N+g(n))}\bigg)=o(N^{-1}).\end{aligned}$$

Following the derivation of (5.39), the third integral is

$$\begin{aligned}&\frac{1}{N+g(n)}\int_0^\pi e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\bigg(\Psi\Big(\frac{\sigma}{N}\Big)^n-\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\bigg)d\sigma\\&\quad=\mathcal{O}\bigg(\frac{ng(n)N}{(N+g(n))^2}\int_0^h t(\log t)^2\,\mathrm{d}t\bigg)=\mathcal{O}\bigg(\frac{ng(n)(\log N)^2}{N(N+g(n))^2}\bigg)=o\bigg(\frac{(\log n)^2}{N}\bigg).\end{aligned}$$

The fourth integral is

$$\begin{aligned}&\frac{1}{N+g(n)}\int_{hN}^{h(N+g(n))}e^{-i\sigma}\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma\\&\quad=\mathcal{O}\bigg(\frac{hg(n)}{N+g(n)}\log\frac{N+g(n)}{hN}\bigg)=\mathcal{O}\bigg(\frac{g(n)\log\big(N+g(n)\big)}{N(N+g(n))}\bigg)=o\bigg(\frac{(\log n)^2}{N}\bigg).\end{aligned}$$

Putting the last five displayed equations together completes the proof of (iii).

(iv) Following the method in (ii) (and similarly to (iii)), we find that

$$\begin{aligned}\mathcal{X}_4&=n\int_h^\varepsilon e^{-itN}(\log t)^2\Psi(t)^n\,\mathrm{d}t-n\int_h^\varepsilon e^{-it(N+g(n))}(\log t)^2\Psi(t)^n\,\mathrm{d}t\\&=\frac{ng(n)}{N(N+g(n))}\int_\pi^{\varepsilon N}e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\Psi\Big(\frac{\sigma}{N}\Big)^n\,\mathrm{d}\sigma\\&\quad+\frac{n}{N+g(n)}\int_\pi^{\varepsilon N}e^{-i\sigma}\bigg(\Big(\log\frac{\sigma}{N}\Big)^2-\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\bigg)\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma\\&\quad+\frac{n}{N+g(n)}\int_\pi^{\varepsilon N}e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\bigg(\Psi\Big(\frac{\sigma}{N}\Big)^n-\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\bigg)d\sigma\\&\quad+\frac{n}{N+g(n)}\bigg(\int_{hN}^{\varepsilon N}-\int_{h(N+g(n))}^{\varepsilon(N+g(n))}\bigg)e^{-i\sigma}\Big(\log\frac{\sigma}{N+g(n)}\Big)^2\Psi\Big(\frac{\sigma}{N+g(n)}\Big)^n\,\mathrm{d}\sigma.\end{aligned}$$

By Lemma 5.4 with $\ell(x)=(\log x)^2$, we find that the first term is

$$\begin{aligned}\frac{ng(n)}{N(N+g(n))}\int_\pi^{\varepsilon N}e^{-i\sigma}\Big(\log\frac{\sigma}{N}\Big)^2\Psi\Big(\frac{\sigma}{N}\Big)^n\,\mathrm{d}\sigma&=\frac{ng(n)}{N+g(n)}\int_h^\varepsilon e^{-itN}(\log t)^2\Psi(t)^n\,\mathrm{d}t\\&=\mathcal{O}\bigg(\frac{g(n)(\log n)^2}{N+g(n)}\bigg)=o\big((\log n)^2\big).\end{aligned}$$

Following the derivation of (5.38), the second term is

$$\begin{aligned}
&\frac{n}{N+g(n)}\int_{\pi}^{\varepsilon N} e^{-i\sigma}\left(\left(\log\frac{\sigma}{N}\right)^2-\left(\log\frac{\sigma}{N+g(n)}\right)^2\right)\Psi\left(\frac{\sigma}{N+g(n)}\right)^n \mathrm{d}\sigma\\
&\quad=\mathcal{O}\left(\frac{ng(n)}{N(N+g(n))}\int_{\pi}^{\varepsilon N}\log\frac{\sigma}{N}\,\mathrm{d}\sigma\right)=\mathcal{O}\left(\frac{ng(n)\log N}{N(N+g(n))}\right)=o\big((\log n)^2\big).
\end{aligned}$$

By (5.23) and Lemma 5.4 with $\ell(x)=(\log x)^2$, we have that the third term is

$$\begin{aligned}
&\frac{n}{N+g(n)}\int_{\pi}^{\varepsilon N} e^{-i\sigma}\left(\log\frac{\sigma}{N}\right)^2\left(\Psi\left(\frac{\sigma}{N}\right)^n-\Psi\left(\frac{\sigma}{N+g(n)}\right)^n\right)d\sigma\\
&\quad=\mathcal{O}\left(\frac{n^2g(n)}{N(N+g(n))^2}\int_{\pi}^{\varepsilon N}\left(\log\frac{\sigma}{N}\right)^2\left|\Psi\left(\frac{\sigma}{N}\right)\right|^{n-1}\mathrm{d}\sigma\right)\\
&\quad=\mathcal{O}\left(\frac{n^2g(n)}{(N+g(n))^2}\int_{h}^{\varepsilon}(\log t)^2|\Psi(t)|^{n-1}\,\mathrm{d}t\right)=\mathcal{O}\left(\frac{ng(n)(\log n)^2}{(N+g(n))^2}\right)=o\big((\log n)^2\big).
\end{aligned}$$

Finally, the fourth term is

$$\begin{aligned}
&\frac{n}{N+g(n)}\left(\int_{hN}^{\varepsilon N}-\int_{h(N+g(n))}^{\varepsilon(N+g(n))}\right)e^{-i\sigma}\left(\log\frac{\sigma}{N+g(n)}\right)^2\Psi\left(\frac{\sigma}{N+g(n)}\right)^n\mathrm{d}\sigma\\
&=\frac{n}{N+g(n)}\left(\int_{hN}^{h(N+g(n))}-\int_{\varepsilon N}^{\varepsilon(N+g(n))}\right)e^{-i\sigma}\left(\log\frac{\sigma}{N+g(n)}\right)^2\Psi\left(\frac{\sigma}{N+g(n)}\right)^n\mathrm{d}\sigma\\
&=\mathcal{O}\left(\frac{n}{N+g(n)}\cdot hg(n)\cdot\left(\log\frac{N+g(n)}{hN}\right)^2+\frac{n}{N+g(n)}\cdot\varepsilon g(n)\cdot\left(\log\frac{N+g(n)}{\varepsilon N}\right)^2\right)\\
&=\mathcal{O}\left(\frac{ng(n)(\log N)^2}{N(N+g(n))}+\frac{ng(n)^3}{N^2(N+g(n))}\right)=o\big((\log n)^2\big).
\end{aligned}$$

Putting the above five displayed equations together completes the proof of (iv). □

Proof of Theorem 5.2. Putting Propositions 5.9 and 5.17 together yields

$$I=\mathcal{O}\left(\frac{1}{N^2}|\mathcal{X}_1|+\frac{n^2}{N^2}|\mathcal{X}_2|+\frac{n^2}{N}|\mathcal{X}_3|+\frac{n^2}{N^2}|\mathcal{X}_4|\right)+\mathcal{O}\left(\frac{n^2\log n}{N^2}\right)=o\left(\frac{n^2(\log n)^2}{N^2}\right).$$

Therefore, we have by Proposition 5.7 that

$$\mathbb{P}\big(S_n\in[N,N+g(n)]\big)=n\mathbb{P}\big(X_1\in[N,N+g(n)]\big)+o\left(\frac{n^2(\log n)^2+N^{2/\mathfrak{p}}}{N^2}\right).$$ □

5.6 Applications

5.6.1 Arithmetic Mean of Lüroth Series Expansion Digits

In this section, we apply Theorems 5.1 and 5.2 to the setting of Lüroth series expansions as outlined in the introduction of this chapter. Recall the definition of the Lüroth map L and the digits $d_n(x)$, $n\in\mathbb{N}$, of each $x\in[0,1]\setminus\mathbb{Q}$. For a slowly varying function ℓ, recall that $\ell^{\#}$ denotes the de Bruijn conjugate of ℓ. For each $j\in\mathbb{N}$, write d_j for the function $x\mapsto d_j(x)$. We first establish the tail behaviour of d_1 when composed with certain regularly varying functions; this is an analogue of Lemma 4.23 for Lüroth series expansions.

Lemma 5.18. *Suppose that* $\psi : (0, +\infty) \to (0, +\infty)$ *is invertible on the interval* $[x_0, +\infty)$ *for some* $x_0 > 0$*. Then* $\lambda(\psi \circ d_1 \geq x) = \lceil \psi^{-1}(x) \rceil^{-1}$ *for all* $x \geq x_0$*.*

Proof. Fix $x \geq x_0$. As the Lüroth map and the Gauss map are defined on the same partition $\{(\frac{1}{n+1}, \frac{1}{n}] : n \in \mathbb{N}\} \cup \{0\}$, we may use the argument in Lemma 4.23 to see that

$$\lambda(\psi \circ d_1 \geq x) = \sum_{\psi(n) \geq x} \lambda\left(\left(\frac{1}{n+1}, \frac{1}{n}\right]\right) = \sum_{n \geq \psi^{-1}(x)} \lambda\left(\left(\frac{1}{n+1}, \frac{1}{n}\right]\right)$$
$$= \lambda\left(\left(0, \frac{1}{\lceil \psi^{-1}(x) \rceil}\right]\right) = \frac{1}{\lceil \psi^{-1}(x) \rceil}. \qquad \square$$

Proof of Corollary 5.3. Define $\psi : (0, +\infty) \to (0, +\infty)$ by $\psi(x) = x$. Then Lemma 5.18 gives

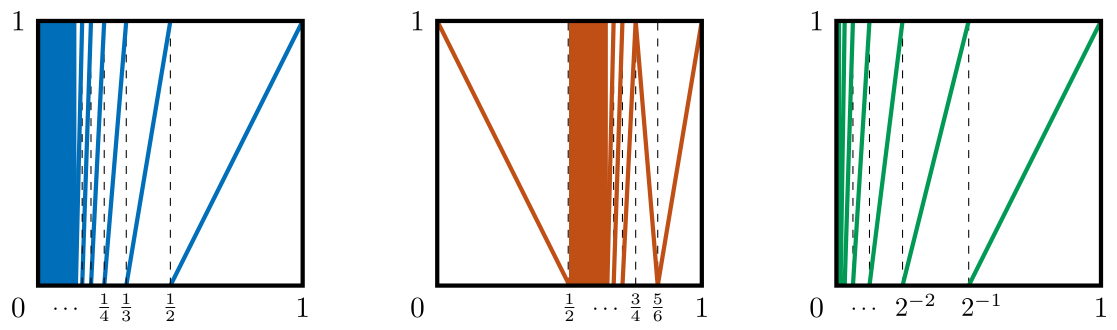
$$\lambda(d_1 \geq x) = \lceil x \rceil^{-1} = \frac{1}{x} - \frac{\lceil x \rceil - x}{x \lceil x \rceil} = \sum_{j \in \mathbb{N}} \frac{c_j}{x^j}$$

for some sequence $(c_j)_{n \in \mathbb{N}}$ with $c_1 = 1$. Thus, (5.3) holds so, since the d_j, $j \in \mathbb{N}$, are i.i.d. random variables (see e.g. [BBDK96]), we may apply Theorems 5.1 and 5.2 to obtain the two statements in Corollary 5.3, also using that $c_j / N^j = o(N^{-(2-2/\mathfrak{p})})$ for all $j \geq 2$. $\square$

5.6.2 Arithmetic Mean of GLS Expansion Digits

Recall from Part I that GLS maps are piecewise linear full-branched unit interval maps; see Figure 3.1 in Part I and Figure 5.2 below for examples. In this section, we will generalise Corollary 5.3 (and Lemma 5.18) to infinite branched GLS maps under a certain condition on the branch-support partition.

Suppose that T is an infinite branched GLS map on the partition $\{P_n : n \in \mathbb{N}\}$ of $[0, 1]$ into non-degenerate sub-intervals of $[0, 1]$ (up to some countable set, which has zero Lebesgue measure). Recall that $N_n := 1/\lambda(P_n)$, $n \in \mathbb{N}$, denotes the unsigned slope of the n^{th} branch of T. Let $d_n(y)$ denote the n^{th} digit appearing in the T-expansion of $y \in [0, 1] \setminus \mathbb{Q}$, and write d_1 for the function $y \mapsto d_1(y)$.



(a) The Lüroth map (b) Altered Lüroth example (c) Non-Lüroth example

Figure 5.2: Examples of GLS maps with infinitely many branches that satisfy the condition (5.41). (a) shows the usual Lüroth map. (b) shows a GLS map obtained by changing the signs of slopes of, and locations of, the branches of the Lüroth map. (c) shows a GLS map whose n^{th} branch has support $[2^{-n}, 2^{-(n-1)}]$, $n \in \mathbb{N}$.

Corollary 5.19. *Let T be a GLS map on the digit set $\mathbb{N}$, and let $\psi : (0, +\infty) \to (0, +\infty)$ be an invertible function. Assume that there is a sequence $(c_j)_{j\in\mathbb{N}}$ of constants with $c_1 > 0$ such that*

$$\sum_{n\geq\psi^{-1}(x)} N_n^{-1} = \sum_{j\in\mathbb{N}} \frac{c_j}{x^j} \tag{5.41}$$

for all x sufficiently large. Set $S_n(\psi \circ d_1) = \sum_{1\leq j\leq n} \psi \circ d_1 \circ T^{j-1}$. Then the following statements hold for any $\mathfrak{p} \geq 2$.

(i) *For all x such that $N := x + b_n$ satisfies $n \log N = o(N)$ as $n \to +\infty$,*

$$\lambda\big(S_n(\psi \circ d_1) - b_n \geq x\big) = n\lambda(\psi \circ d_1 \geq N) + o\bigg(\frac{n^2(\log N)^2 + N^{2/\mathfrak{p}}}{N^2}\bigg).$$

In particular, if (5.41) *holds for $\psi(x) = x$, then*

$$\lambda\bigg(\bigg\{y \in [0, 1] \setminus \mathbb{Q} : \frac{d_1(y) + \cdots + d_n(y)}{n} \geq x\bigg\}\bigg) = \frac{n}{N} + o\bigg(\frac{n^2(\log N)^2 + N^{2/\mathfrak{p}}}{N^2}\bigg).$$

(ii) *As $x/a_n \to +\infty$,*

$$\lambda\big(S_n(\psi \circ d_1) - b_n \in [x, x + g(n)]\big) = n\lambda(\psi \circ d_1 \geq N) + o\bigg(\frac{n^2(\log n)^2 + N^{2/\mathfrak{p}}}{N^2}\bigg).$$

In particular, if (5.41) *holds for $\psi(x) = x$, then*

$$\lambda\bigg(\bigg\{y \in [0, 1] \setminus \mathbb{Q} : \frac{d_1(y) + \cdots + d_n(y)}{n} \in [x, x + g(n)]\bigg\}\bigg) = \frac{ng(n)}{N(N + g(n))} + o\bigg(\frac{n^2(\log n)^2 + N^{2/\mathfrak{p}}}{N^2}\bigg)$$

for any choice of function $g : \mathbb{R} \to (0, +\infty)$ satisfying (5.4).

Proof. Following the proof of Lemma 5.18 (or Lemma 4.23) and then using the condition (5.41), we see that there is a sequence $(c_j)_{j\in\mathbb{N}}$ with $c_1 > 0$ such that

$$\lambda(\psi \circ d_1 \geq x) = \sum_{\psi(n)\geq x} \lambda(P_n) = \sum_{n\geq\psi^{-1}(x)} \lambda(P_n) = \sum_{n\geq\psi^{-1}(x)} N_n^{-1} = \sum_{j\in\mathbb{N}} \frac{c_j}{x^j}$$

for all x sufficiently large. Therefore, we may apply Theorems 5.1 and 5.2 to the random variables $X_j := \psi \circ d_1 \circ T^{j-1} = \psi \circ d_j$, $j \in \mathbb{N}$, to get the main statements of (i) and (ii) respectively. The statements on the arithmetic mean follow by taking $\psi(x) = x$. □

The condition (5.41) in Corollary 5.19 holds in various settings. For example, (5.41) holds for the Lüroth map and $\psi(x) = x$ by the calculation in the proof of Lemma 5.18, using that the branches of the Lüroth map are ordered by increasing (unsigned) slope from right to left, which allowed us to obtain Corollary 5.3. By the convention that the branches of GLS maps are ordered by increasing (unsigned) slope size, condition (5.41) continues to hold for any GLS map obtained from the Lüroth map by changing the left-to-right order of the branches or by changing the signs of the slopes; see Figure 5.2(b) for an example of a GLS map obtained in this way. 5

For an example with a GLS not related to the Lüroth map, consider the GLS map T such that the n^{th} branch of T has positive slope and is supported on $[2^{-n}, 2^{-(n-1)}]$ for each $n \in \mathbb{N}$

(one could further change the left-to-right order of these branches or change the signs of their slopes); see Figure 5.2(c). Put $\psi(x) = 2^x$. Then

$$\sum_{n \geq \psi^{-1}(x)} N_n^{-1} = \sum_{n \geq \log_2 x} 2^{-n} = 2^{-\lceil \log_2 x \rceil}$$

so there is indeed a sequence $(c_j)_{j \in \mathbb{N}}$ with $c_1 > 0$ such that (5.41) holds for this choice of GLS T and invertible function ψ.

5

Bibliography

[Aar97] J. Aaronson. *An introduction to infinite ergodic theory.* Vol. 50. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1997, pp. xii+284.

[AD98] J. Aaronson and M. Denker. "Characteristic functions of random variables attracted to 1-stable laws". In: *Ann. Probab.* 26.1 (1998), pp. 399–415.

[AD01] J. Aaronson and M. Denker. "Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps". In: *Stoch. Dyn.* 1.2 (2001), pp. 193–237.

[ADSZ04] J. Aaronson, M. Denker, O. Sarig, and R. Zweimüller. "Aperiodicity of cocycles and conditional local limit theorems". In: *Stoch. Dyn.* 4.1 (2004), pp. 31–62.

[AS64] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Vol. No. 55. National Bureau of Standards Applied Mathematics Series. U. S. Government Printing Office, Washington, DC, 1964, pp. xiv+1046.

[Arr15] A. Arroyo. "Generalised Lüroth expansions and a family of Minkowski's question-mark functions". In: *C. R. Math. Acad. Sci. Paris* 353.10 (2015), pp. 943–946.

[Bak21] S. Baker. "Iterated function systems with super-exponentially close cylinders". In: *Adv. Math.* 379 (2021), Paper No. 107548, 13.

[BD17] S. Baker and K. Dajani. "Induced random β-transformation". In: *Acta Arith.* 178.1 (2017), pp. 1–14.

[BKSN19] H.N. Balakrishnan, A. Kathpalia, S. Saha, and N. Nagaraj. "ChaosNet: a chaos based artificial neural network architecture for classification". In: *Chaos* 29.11 (2019), pp. 113125, 17.

[BJKR21] B. Bárány, T. Jordan, A. Käenmäki, and M. Rams. "Birkhoff and Lyapunov spectra on planar self-affine sets". In: *Int. Math. Res. Not. IMRN* 10 (2021), pp. 7966–8005.

[BK17] B. Bárány and A. Käenmäki. "Ledrappier–Young formula and exact dimensionality of self-affine measures". In: *Advances in Mathematics* 318 (2017), pp. 88–129.

[BK21] B. Bárány and A. Käenmäki. "Super-exponential condensation without exact overlaps". In: *Advances in Mathematics* 379 (2021), p. 107549.

[BM08] J. Barral and M. Mensi. "Multifractal analysis of Birkhoff averages on 'self-affine' symbolic spaces". In: *Nonlinearity* 21.10 (2008), pp. 2409–2425.

[BG11] L. Barreira and K. Gelfert. "Dimension estimates in smooth dynamics: a survey of recent results". In: *Ergodic Theory and Dynamical Systems* 31.3 (2011), pp. 641–671.

[BI09] L. Barreira and G. Iommi. "Frequency of digits in the Lüroth expansion". In: *J. Number Theory* 129.6 (2009), pp. 1479–1490.

[BS01] L. Barreira and B. Saussol. "Variational principles and mixed multifractal spectra". In: *Trans. Amer. Math. Soc.* 353.10 (2001), pp. 3919–3944.

[BSS02a] L. Barreira, B. Saussol, and J. Schmeling. "Distribution of frequencies of digits via multifractal analysis". In: *J. Number Theory* 97.2 (2002), pp. 410–438.

[BSS02b] L. Barreira, B. Saussol, and J. Schmeling. "Higher-dimensional multifractal analysis". In: *J. Math. Pures Appl. (9)* 81.1 (2002), pp. 67–91.

[BBDK96] J. Barrionuevo, R.M. Burton, K. Dajani, and C. Kraaikamp. "Ergodic properties of generalized Lüroth series". In: *Acta Arith.* 74.4 (1996), pp. 311–327.

[Bed84] T. Bedford. "Crinkly curves, Markov partitions and dimension in self-similar sets". PhD thesis. University of Warwick, 1984.

[Ber19] Q. Berger. “Notes on random walks in the Cauchy domain of attraction”. In: *Probab. Theory Related Fields* 175.1-2 (2019), pp. 1–44.

[BDKK+24] V. Berthé, K. Dajani, C. Kalle, E. Krawczyk, H. Kuru, and A. Thevis. “Rational approximations, multidimensional continued fractions, and lattice reduction”. In: *Women in numbers Europe IV*. Vol. 32. Assoc. Women Math. Ser. Springer, Cham, 2024, pp. 111–154.

[Bes35] A.S. Besicovitch. “On the sum of digits of real numbers represented in the dyadic system”. In: *Math. Ann.* 110.1 (1935), pp. 321–330.

[Bil65] P. Billingsley. *Ergodic theory and information*. John Wiley & Sons, Inc., New York-London-Sydney, 1965, pp. xiii+195.

[Bil12] P. Billingsley. *Probability and measure*. Anniversary. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, 2012, pp. xviii+624.

[BGT87] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*. Vol. 27. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1987, pp. xx+491.

[BP17] C. J. Bishop and Y. Peres. *Fractals in probability and analysis*. Vol. 162. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2017, pp. ix+402.

[BG09] J. Bochi and N. Gourmelon. “Some characterizations of domination”. In: *Math. Z.* 263.1 (2009), pp. 221–231.

[BK24] A. Boonstra and C. Kalle. “Constructions of normal numbers with infinitely many digits”. https://arxiv.org/pdf/2402.14500. 2024.

[Bow73] R. Bowen. “Topological entropy for noncompact sets”. In: *Transactions of the American Mathematical Society* 184 (1973), pp. 125–136.

[Caj81] H. Cajar. *Billingsley dimension in probability spaces*. Vol. 892. Lecture Notes in Mathematics. Springer-Verlag, Berlin-New York, 1981, pp. i+106.

[Can69] G. Cantor. “Über die einfachen Zahlensysteme”. In: *Z. Math. Phys.* 14 (1869), pp. 121–128.

[CFH08] Y. Cao, D. Feng, and W. Huang. “The thermodynamic formalism for sub-additive potentials”. In: *Discrete Contin Dyn Syst Ser A* 20.3 (2008), pp. 639–657.

[CD19] F. Caravenna and R. Doney. “Local large deviations and the strong renewal theorem”. In: *Electron. J. Probab.* 24 (2019), Paper No. 72, 48.

[Che21] C. Chen. “Self-similar sets with super-exponential close cylinders”. In: *Annales Fennici Mathematici* 46.2 (2021), pp. 727–738.

[Cli14] V. Climenhaga. “The thermodynamic approach to multifractal analysis”. In: *Ergodic Theory and Dynamical Systems* 34.5 (2014), pp. 1409–1450.

[DK07] K. Dajani and C. Kalle. “Random β-expansions with deleted digits”. In: *Discrete Contin. Dyn. Syst.* 18.1 (2007), pp. 199–217.

[DK20] K. Dajani and C. Kalle. “Invariant measures, matching and the frequency of 0 for signed binary expansions”. In: *Publ. Res. Inst. Math. Sci.* 56.4 (2020), pp. 701–742.

[DK21] K. Dajani and C. Kalle. *A first course in ergodic theory*. Chapman & Hall/CRC, Boca Raton, FL, 2021, pp. xiii+253.

[DK03] K. Dajani and C. Kraaikamp. “Random β-expansions”. In: *Ergodic Theory Dynam. Systems* 23.2 (2003), pp. 461–479.

[DO18] K. Dajani and M. Oomen. “Random N-continued fraction expansions”. In: *J. Approx. Theory* 227 (2018), pp. 1–26.

[DV05] K. Dajani and M. de Vries. “Measures of mazimal entropy for random β-expansions”. In: *J. Eur. Math. Soc.* 7.1 (2005), pp. 51–68.

[DV07] K. Dajani and M. de Vries. “Invariant densities for random β-expansions”. In: *J. Eur. Math. Soc.* 9.1 (2007), pp. 157–176.

[DDGV06] I. Daubechies, R. DeVore, C.S. Güntürk, and V. Vaishampayan. “A/D conversion with imperfect quantizers”. In: *IEEE Trans. Inform. Theory* 52.3 (2006), pp. 874–885.

[Dou94] P. Doukhan. *Mixing*. Vol. 85. Lecture Notes in Statistics. Properties and examples. Springer-Verlag, New York, 1994, pp. xii+142.

[Egg49] H. G. Eggleston. “The fractional dimension of a set defined by decimal properties”. In: *Quart. J. Math. Oxford Ser.* 20 (1949), pp. 31–36.

[Fal14] K. Falconer. *Fractal geometry.* Third. Mathematical foundations and applications. John Wiley & Sons, Ltd., Chichester, 2014, pp. xxx+368.

[FFW01] A. Fan, D. Feng, and J. Wu. “Recurrence, dimension and entropy”. In: *J. London Math. Soc. (2)* 64.1 (2001), pp. 229–244.

[FJLR15] A. Fan, T. Jordan, L. Liao, and M. Rams. “Multifractal analysis for expanding interval maps with infinitely many branches”. In: *Trans. Amer. Math. Soc.* 367.3 (2015), pp. 1847–1870.

[FLMW10] A. Fan, L. Liao, J. Ma, and B. Wang. “Dimension of Besicovitch-Eggleston sets in countable symbolic space”. In: *Nonlinearity* 23.5 (2010), pp. 1185–1197.

[FH09] D. Feng and H. Hu. “Dimension theory of iterated function systems”. In: *Communications on Pure and Applied Mathematics* 62.11 (2009), pp. 1435–1500.

[FH10] D.-J. Feng and W. Huang. “Lyapunov spectrum of asymptotically sub-additive potentials”. In: *Comm. Math. Phys.* 297.1 (2010), pp. 1–43.

[Gan96] N. Gantert. “Large deviations for a heavy-tailed mixing sequence”. Pre-print TU Berlin. 1996.

[GK49] B. V. Gnedenko and A. N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables.* Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow-Leningrad, 1949, p. 264.

[GKKS25] S. van Golden, C. Kalle, S. Kombrink, and T. Samuel. “Dimensions of infinitely generated self-affine sets and restricted digit sets for signed Lüroth expansions”. In: *Nonlinearity* 38 (2025).

[Gou10a] S. Gouëzel. “Almost sure invariance principle for dynamical systems by spectral methods”. In: *Ann. Probab.* 38.4 (2010), pp. 1639–1671.

[Gou10b] S. Gouëzel. “Characterization of weak convergence of Birkhoff sums for Gibbs-Markov maps”. In: *Israel J. Math.* 180 (2010), pp. 1–41.

[GL16] Y. Gui and W. Li. “Hausdorff dimensions of sets related to Lüroth expansion”. In: *Acta Math. Hungar.* 150.2 (2016), pp. 286–302.

[Hey67] C. C. Heyde. “On large deviation problems for sums of random variables which are not attracted to the normal law”. In: *Ann. Math. Statist.* 38 (1967), pp. 1575–1578.

[Hoc14] M. Hochman. “On self-similar sets with overlaps and inverse theorems for entropy”. In: *Ann. Math.* 180.773 (2014), pp. 773–822.

[Hol00] Frank den Hollander. *Large deviations.* Vol. 14. Fields Institute Monographs. American Mathematical Society, Providence, RI, 2000, pp. x+143.

[HK24] Y. Huang and C. Kalle. “Rational approximation with generalised α-Lüroth expansions”. In: *Acta Arith.* 215.4 (2024), pp. 327–353.

[Hut81] J. E. Hutchinson. “Fractals and self-similarity”. In: *Indiana Univ. Math. J.* 30.5 (1981), pp. 713–747.

[IL65] I. A. Ibragimov and J. V. Linnik. *Independent and Stationary sequences of random variables.* Izdat. “Nauka”, Moscow, 1965, p. 524.

[IKM24] J. Imbierski, C. Kalle, and R. Mohammadpour. “Birkhoff spectrum for diagonally self-affine sets and digit frequencies for GLS systems with redundancy”. In: *Proceedings of the Royal Society of Edinburgh: Section A Mathematics* (2024), pp. 1–26.

[IT24] J. Imbierski and D. Terhesiu. “Stable large deviations for deterministic dynamical systems”. In: *Stoch. Dyn.* 24.4 (2024), Paper No. 2450023, 42.

[IJ15] G. Iommi and T. Jordan. “Multifractal analysis of Birkhoff averages for countable Markov maps”. In: *Ergodic Theory Dynam. Systems* 35.8 (2015), pp. 2559–2586.

[JV69] H. Jager and C. de Vroedt. “Lüroth series and their ergodic properties”. In: *Nederl. Akad. Wetensch. Proc. Ser. A 72=Indag. Math.* 31 (1969), pp. 31–42.

[JM16] Y. Jitsumatsu and K. Matsumura. “A β-ary to binary conversion for random number generation using a β encoder”. In: *NOLTA, IEICE* (2016), pp. 38–55.

[JMS18] T. Jordan, S. Munday, and T. Sahlsten. “Stability and perturbations of countable Markov maps”. In: *Nonlinearity* 31.4 (2018), pp. 1351–1377.

[JR21] T. Jordan and M. Rams. “Birkhoff spectrum for piecewise monotone interval maps”. In: *Fund. Math.* 252.2 (2021), pp. 203–223.

[JS07] T. Jordan and K Simon. “Multifractal analysis of Birkhoff averages for some self-affine IFS”. In: *Dynamical Systems* 22.4 (2007), pp. 469–483.

[Jur21] N. Jurga. “Dimension spectrum of infinite self-affine iterated function systems”. In: *Selecta Math. (N.S.)* 27.3 (2021), Paper No. 49, 23.

[KR14] A. Käenmäki and H. W. J. Reeve. “Multifractal analysis of Birkhoff averages for typical infinitely generated self-affine sets”. In: *J. Fractal Geom.* 1.1 (2014), pp. 83–152.

[KKV17] C. Kalle, T. Kempton, and E. Verbitskiy. “The random continued fraction transformation”. In: *Nonlinearity* 30.3 (2017), pp. 1182–1203.

[KM22a] C. Kalle and M. Maggioni. “Invariant densities for random systems of the interval”. In: *Ergodic Theory Dynam. Systems* 42.1 (2022), pp. 141–179.

[KM22b] C. Kalle and M. Maggioni. “On approximation by random Lüroth expansions”. In: *Int. J. Number Theory* 18.5 (2022), pp. 1013–1046.

[KMTV22] C. Kalle, V. Matache, M. Tsujii, and E. Verbitskiy. “Invariant densities for random continued fractions”. In: *J. Math. Anal. Appl.* 512.2 (2022), Paper No. 126163, 28.

[Kat76] Y. Katznelson. *An introduction to harmonic analysis.* corrected. Dover Publications, Inc., New York, 1976, pp. xiv+264.

[KKSS15] J. Kautzsch, M. Kesseböhmer, T. Samuel, and B. O. Stratmann. “On the asymptotics of the α-Farey transfer operator”. In: *Nonlinearity* 28.1 (2015), pp. 143–166.

[Kem14] K. Kempton. “On the invariant density of the random β-transformation”. In: *Acta Math. Hungar.* 142.2 (2014), pp. 403–419.

[KT22] P. Kevei and D. Terhesiu. “Strong renewal theorem and local limit theorem in the absence of regular variation”. In: *J. Theoret. Probab.* 35.2 (2022), pp. 1013–1048.

[Kle20] A. Klenke. *Probability theory—a comprehensive course.* Third. Universitext. Springer, Cham, 2020, pp. xiv+716.

[KT93] K. Koyama and Y. Tsuruoka. “Speeding up elliptic cryptosystems by using a signed binary window method”. In: *Advances in cryptology—CRYPTO ’92 (Santa Barbara, CA, 1992).* Vol. 740. Lecture Notes in Comput. Sci. Springer, Berlin, 1993, pp. 345–357.

[LG92] S. P. Lalley and D. Gatzouras. “Hausdorff and box dimensions of certain self-affine fractals”. In: *Indiana Univ. Math. J.* 41.2 (1992), pp. 533–568.

[Lür83] J. Lüroth. “Ueber eine eindeutige Entwickelung von Zahlen in eine unendliche Reihe”. In: *Math. Ann.* 21.3 (1883), pp. 411–423.

[Mar54] J. M. Marstrand. “The dimension of Cartesian product sets”. In: *Proc. Cambridge Philos. Soc.* 50 (1954), pp. 198–202.

[McM84] C. McMullen. “The Hausdorff dimension of general Sierpiński carpets”. In: *Nagoya Math. J.* 96 (1984), pp. 1–9.

[MPT24] I. Melbourne, F. Pène, and D. Terhesiu. “Local large deviations for periodic infinite horizon Lorentz gases”. In: *J. Anal. Math.* 152.1 (2024), pp. 283–316.

[MT17] I. Melbourne and D. Terhesiu. “Operator renewal theory for continuous time dynamical systems with finite and infinite measure”. In: *Monatsh. Math.* 182.2 (2017), pp. 377–431.

[MT22] I. Melbourne and D. Terhesiu. “Analytic proof of multivariate stable local large deviations and application to deterministic dynamical systems”. In: *Electron. J. Probab.* 27 (2022), Paper No. 21, 17.

[MW13] T. Mikosch and O. Wintenberger. “Precise large deviations for dependent regularly varying sequences”. In: *Probab. Theory Related Fields* 156.3-4 (2013), pp. 851–887.

[Moh22] R. Mohammadpour. “Lyapunov spectrum properties and continuity of the lower joint spectral radius”. In: *J. Stat. Phys.* 187.3 (2022), Paper No. 23, 29.

[Moh23] R. Mohammadpour. “On the multifractal formalism of Lyapunov exponents: a survey of recent results”. In: *New trends in Lyapunov exponents, CIM Series in Mathematical Sciences, Springer, Cham.* CIM Ser. Math. Sci. (2023), pp. 119–139.

[Moh24] R. Mohammadpour. “Entropy spectrum of Lyapunov exponents for typical cocycles”. In: *Journal of Dynamics and Differential Equations* (2024).

[MO90] F. Morain and J. Olivos. “Speeding up the computations on an elliptic curve using addition-subtraction chains”. In: *RAIRO Inform. Théor. Appl.* 24.6 (1990), pp. 531–543.

[Mun11] S. Munday. “A note on Diophantine fractals for α-Lüroth systems”. In: *Integers* 11B (2011), Paper No. A10, 14.

[Nag69] A. V. Nagaev. “Integral limit theorems with regard to large deviations when Cramér’s condition is not satisfied. I, II”. In: *Teor. Verojatnost. i Primenen.* 14 (1969), pp. 51–63, 203–216.

[Nag79] S. V. Nagaev. “Large deviations of sums of independent random variables”. In: *Ann. Probab.* 7.5 (1979), pp. 745–789.

[NT24] Y. Nakajima and H. Takahasi. “Hausdorff dimension of sets with restricted, slowly growing partial quotients in the semi-regular continued fraction”. In: *J. Math. Soc. Japan Advance Publication* (2024).

[Nie99] O.A. Nielsen. “The Hausdorff and packing dimensions of some sets related to Sierpiński carpets”. In: *Canad. J. Math.* 51.5 (1999), pp. 1073–1088.

[Pes97] Y. Pesin. *Dimension theory in dynamical systems.* Chicago Lectures in Mathematics. Contemporary views and applications. University of Chicago Press, Chicago, IL, 1997, pp. xii+304.

[PW01] Y. Pesin and H. Weiss. “The multifractal analysis of Birkhoff averages and large deviations”. In: *Global analysis of dynamical systems.* Inst. Phys., Bristol, 2001, pp. 419–431.

[PW99] M. Pollicott and H. Weiss. “Multifractal analysis of Lyapunov exponent for continued fraction and Manneville-Pomeau transformations and applications to Diophantine approximation”. In: *Comm. Math. Phys.* 207.1 (1999), pp. 145–171.

[PS72] G. Pólya and G. Szegö. *Problems and Theorems in Analysis.* Vol. 1: Series, Integral Calculus, Theory of Functions. Springer-Verlag, New York and Berlin, 1972.

[Rap22] A. Rapaport. “Proof of the exact overlaps conjecture for systems with algebraic contractions”. In: *Ann. Sci. Éc. Norm. Supér. (4)* 55.5 (2022), pp. 1357–1377.

[Ree11] H. W. J. Reeve. “Multifractal analysis for Birkhoff averages on Lalley-Gatzouras repellers”. In: *Fund. Math.* 212.1 (2011), pp. 71–93.

[RHN23] A.S. Remya Ajai, N.B. Harikrishnan, and N. Nagaraj. “Analysis of logistic map based neurons in neurochaos learning architectures for data classification”. In: *Chaos Solitons Fractals* 170 (2023), Paper No. 113347, 11.

[Roz89] L. V. Rozovskii. “Probabilities of large deviations of sums of independent random variables with a common distribution function that belongs to the domain of attraction of a normal law”. In: *Teor. Veroyatnost. i Primenen.* 34.4 (1989), pp. 686–705.

[Rus23] T. Rush. “Multifractal analysis for Markov interval maps with countably many branches”. In: *Nonlinearity* 36.4 (2023).

[Suz19] S. Suzuki. “Invariant density functions of random β-transformations”. In: *Ergodic Theory Dynam. Systems* 39.4 (2019), pp. 1099–1120.

[Tak19] H. Takahasi. “Large deviation principle for arithmetic functions in continued fraction expansion”. In: *Monatsh. Math.* 190.1 (2019), pp. 137–152.

[Tie23] Y. Tierce. “The natural extension of the random beta-transformation”. In: *Ergodic Theory Dynam. Systems* 43.11 (2023), pp. 3861–3896.

[Weg68] H. Wegmann. “Über den Dimensionsbegriff in Wahrscheinlichkeitsräumen von P. Billingsley. I, II”. In: *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 9 (1968).

[You84] L.-S. Young. “Dimension, entropy and Lyapunov exponents in differentiable dynamical systems”. In: *Phys. A* 124.1-3 (1984). Mathematical physics, VII (Boulder, Colo., 1983).

[Zwe98] R. Zweimüller. “Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points”. In: *Nonlinearity* 11.5 (1998), pp. 1263–1276.

Samenvatting

Een *dynamisch systeem in discrete tijd* is een abstracte ruimte samen met een collectie regels die beschrijven hoe de ruimte zich kan ontwikkelen tussen opvolgende discrete tijdsstappen. Deze abstracte ruimte heet de *toestandsruimte* van het systeem, gezien dat deze de verscheidene toestanden beschrijft waar het systeem zich in kan bevinden. De ontwikkeling van de toestandsruimte met betrekking tot de collectie regels wordt de *dynamica* van het systeem genoemd. In dit proefschrift hebben we de onwaarschijnlijkheid gekwantificeerd van een aantal probabilistisch onwaarschijnlijke gebeurtenissen die gedefinieerd zijn vanuit bepaalde families van dynamische systemen. Er waren verscheidene manieren waarop de onwaarschijnlijkheid waarmee een dynamisch gedefinieerde gebeurtenis optreedt gekwantificeerd kon worden. De meest natuurlijke was om een kans toe te wijzen dat zo'n gebeurtenis optrad, al gebruikten we ook andere ideeën om meer informatie te verkrijgen in de situatie dat de kans op een gebeurtenis nul was. In dit laatste geval waren de gebeurtenissen die we hebben beschouwd voorbeelden van *fractalen*. Dit zijn verzamelingen die vaak gekarakteriseerd worden door een meetkundig gecompliceerde rand en een herhalende structuur. In plaats van kansen hebben we daarom gebruik gemaakt van de *Hausdorff-dimensie* – een functie die de meetkundige complexiteit van een fractaal kwantificeert – om deze gebeurtenissen van elkaar te onderscheiden.

Dit proefschrift is opgedeeld in twee delen, elk bestaande uit twee hoofdstukken.

In Deel I zijn 'niveauverzamelingen voor Birkhoff-gemiddelden' bestudeerd voor verscheidene dynamische systemen. Een *Birkhoff-gemiddelde* is het gemiddelde over de tijd van een of andere vastgestelde functie langs de baan van een dynamisch systeem en is daarom één van de meest belangrijke grootheden die verschijnen in de *ergodentheorie* – de studie van statistische eigenschappen van dynamische systemen beschreven door het gedrag van zulke gemiddelden. Gegeven een vaste functie Φ, ook wel een *potentiaal* genoemd, en een vast getal $\boldsymbol{\alpha}$ uit de verzameling van alle mogelijke Birkhoff-gemiddelden is de *Birkhoff-gemiddelde* $\boldsymbol{\alpha}$*-niveauverzameling* behorende bij Φ gedefinieerd als de verzameling van alle toestanden in de toestandsruimte waarvan het Birkhoff-gemiddelde (m.b.t. Φ) langs de baan van de toestand gelijk is aan $\boldsymbol{\alpha}$. Gezien het belang van Birkhoff-gemiddelden in de ergodentheorie is het een relevante vraag om te stellen hoe groot de niveauverzamelingen van Birkhoff-gemiddelden kunnen zijn. Voor veel dynamische systemen, waaronder *ergodische* dynamische systemen (zie §1.3.4) en de systemen die we in Deel I hebben beschouwd, hebben bijna alle toestanden in de toestandsruimte hetzelfde Birkhoff-gemiddelde (m.b.t. één of andere vaste potentiaal), wat betekent dat er een unieke niveauverzameling is met een kans van 1. Er kunnen ook toestanden zijn met een Birkhoff-gemiddelde dat hiervan verschilt en dus aan een andere niveauverzameling toebehoren. Gezien deze andere niveauverzamelingen allemaal een kans van 0 moeten hebben, hebben we van deze verzamelingen de grootte verder gekwantificeerd met behulp van de Hausdorff-dimensie. De Hausdorff-dimensie van verzamelingen van Birkhoff-gemiddelden zijn voor verschillende potentialen bestudeerd in verscheidene contexten; zie bijvoorbeeld [PW01; BSS02a; JS07; BM08; FLMW10; KR14; IJ15; BJKR21; Jur21; JR21; Rus23].

In Hoofdstuk 2 zijn de Birkhoff-gemiddelden gegeven met betrekking tot een type dynamische systemen genaamd *geïtereerde functiesystemen*, waar de dynamica gegeven wordt door een collectie functies van dezelfde toestandsruimte naar zichzelf die de afstanden tussen tweetallen toestanden niet vergroten. In de eerste helft van Hoofdstuk 2 zijn de specifieke geïtereerde functiesystemen die we beschouwd hebben *Lalley-Gatzouras systemen* [LG92], die bestaan uit eindige collecties van lineaire afbeeldingen op het eenheidsvierkant. In Stelling 2.1 hebben we een formule verkregen voor de Hausdorff-dimensies van de niveauverzamelingen voor Birkhoff-gemiddelden behorende bij deze systemen met betrekking tot een arbitraire continue potentiaal, onder één van twee alternatieve voorwaarden die elk te maken hadden met hoe weinig de iteraties van het geïtereerde functiesysteem overlapten. Deze situatie is ook in een meer algemene context behandeld in [Ree11]; gezien de extra voorwaarden op ons systeem waren we echter in staat meer informatie te verkrijgen dan daar (via andere methoden) gegeven is. Ook hebben we in Propositie 2.8 de verzameling van waarden $\boldsymbol{\alpha}$ verkregen die hoorden bij niet-lege Birkhoff-gemiddelde niveauverzamelingen. De grootste uitdaging in deze (tweedimensionale) context was dat de systemen die we behandelden *niet-conform* waren, wat wil zeggen dat de mate waarin de functies punten langs elk van de twee assen van het eenheidsvierkant waarop onze functies gedefinieerd zijn samentrokken van elkaar verschilden. Hierdoor konden veel van de gebruikelijke methoden die in de fractale meetkunde gebruikt worden niet toegepast worden.

In de rest van Hoofdstuk 2, en vervolgens in Hoofdstuk 3, beschouwden we dynamische systemen genaamd *getalsystemen*, die gebruikt worden om getalsontwikkelingen, d.w.z. symbolische representaties, te genereren en toe te wijzen aan elk getal in een toestandsruimte. Met behulp van deze getalsystemen hebben we een familie van gebeurtenissen definieerd door de verschillende frequenties te beschouwen waarin symbolen, ook wel *cijfers* genoemd, konden verschijnen in de getalsontwikkelingen van de getallen in de toestandsruimte. Deze gebeurtenissen zijn een type verzamelingen genaamd *Besicovitch-Eggleston verzamelingen*, vernoemd naar het baanbrekende werk van Besicovitch [Bes35] en Eggleston [Egg49]. Gegeven dat de frequentie waarin een cijfer voorkomt in een getalsontwikkeling geschreven kan worden als een Birkhoff-gemiddelde met betrekking tot een zekere potentiaal (zie §1.3.4), zijn Besicovitch-Eggleston verzamelingen specifieke voorbeelden van niveauverzamelingen voor Birkhoff-gemiddelden. Dit betekent dat deze verzamelingen behoren tot de bovengenoemde situatie waarin er een unieke niveauverzameling is met een kans van 1 en de andere niveauverzamelingen een kans van 0 hebben; we hebben deze laatste verzamelingen verder bestudeerd met behulp van de Hausdorff-dimensie. Er is een overvloed aan literatuur over de Hausdorff-dimensie van verzamelingen die bepaald worden door eigenschappen van cijfers in getalsontwikkelingen; zie bijvoorbeeld [Bes35; Egg49; PW99; BSS02a; BI09; FLMW10; Mun11; FJLR15; NT24]. Het meest relevant voor dit proefschrift was [FLMW10, Theorem 1.2], waar de auteurs de Hausdorff-dimensie verkrijgen van Besicovitch-Eggleston verzamelingen gedefinieerd voor dynamische systemen die symbolische representaties genereren waarin de cijfers uit een aftelbaar oneindige poel symbolen kunnen komen. Als een toepassing verkrijgen ze [FLMW10, Theorem 1.1] door het bovengenoemde resultaat over te dragen naar getalsystemen die gegeven worden door stuksgewijs lineair en stuksgewijs bijectieve afbeeldingen op het eenheidsvierkant – de zogenaamde *gegeneraliseerde Lüroth ontwikkeling* (in het Engels afgekort tot GLS) afbeeldingen geïntroduceerd in [BBDK96]. Deze afbeeldingen zijn in het bijzonder in Deel I veel gebruikt.

Een belangrijke eigenschap van de getalsystemen die we in de tweede helft van Hoofdstuk 2 en in Hoofdstuk 3 beschouwden was dat deze *niet-autonoom* waren, dat wil zeggen dat van deze systemen de dynamica gegeven werd op een tijdsafhankelijke manier; de regels die de tijdsevolutie van het systeem bepaalden konden van de ene tijdsstap naar de volgende veranderen. Tijdsafhankelijke dynamische systemen zijn belangrijk om te beschouwen, omdat deze beter het gedrag omvatten van fenomenen uit de echte wereld, die niet vaak beschreven kunnen worden door tijdsonafhankelijke processen. Er bestaan veel resultaten over getalsontwikkelingen gegenereerd door niet-autonome en andere types tijdsafhankelijke dynamische systemen, zie bijvoorbeeld [DK03; DV05; DK07; Kem14; KKV17; DO18; DK20; KM22b; NT24; GKKS25; BDKK+24]. De tweede helft van Hoofdstuk 2 werd gewijd aan het bewijzen van een resultaat (namelijk Stelling 2.2) dat vergelijkbaar is met [FLMW10, Theorem 1.1], maar dan voor niet-autonome getalsystemen opgebouwd uit GLS afbeeldingen waarvan de bijbehorende getalsontwikkelingen hun cijfers haalden uit een eindige poel cijfers. Ook verkregen we in Propositie 2.9 de verzameling van waarden $\boldsymbol{\alpha}$ waarvan de bijbehorende Birkhoff-gemiddelde niveauverzamelingen voor deze niet-autonome systemen niet-leeg waren. In tegenstelling tot de situatie van Stelling 2.1, zoals hierboven beschreven, was deze niet-autonome constructie in Stelling 2.2 *conform*, waardoor we de standaardmethoden van de fractale meetkunde hebben kunnen toegepassen; de niet-autonomie betekende echter weer dat fundamentele resultaten zoals de Birkhoff-ergodenstelling (zie Stelling 1.7) niet gebruikt konden worden. Een standaardmethode die in Hoofdstuk 2 is gebruikt was om de onderliggende 'symbolische dynamica' van het getalsysteem te beschouwen. Een getalsysteem heeft een verzameling van bijbehorende symbolen (namelijk de cijfers) waarvan je oneindige rijtjes kunt construeren door oneindig van deze cijfers achter elkaar te zetten. Je kunt dan een symbolisch dynamisch systeem definiëren op de ruimte van al deze oneindige rijtjes door de *verschuivingsafbeelding* te beschouwen, die werkt door het eerste cijfer in een rijtje te verwijderen en de posities van alle overgebleven cijfers met één naar links te verschuiven. Via de symbolische representaties die de getalsontwikkelingen beschrijven, kun je de getallen in de toestandsruimte associëren met oneindige rijtjes van symbolen uit de symbolenruimte en vice versa. Op deze manier kunnen eigenschappen van een getalsysteem gevonden worden door de symbolische dynamica te bestuderen en eigenschappen die we vinden over te dragen naar het originele getalsysteem.

In Hoofdstuk 3 generaliseerden we Stelling 2.2 uit Hoofdstuk 2 door toe te staan dat de mogelijke poel cijfers aftelbaar oneindig is onder aanvullende, kleine voorwaarden op het systeem. De generalisatie die we verkregen in Stelling 3.3 omvat de autonome gevallen die beschouwd zijn in [FLMW10, Theorem 1.1]. We verkregen ook Stelling 3.2, een uitbreiding van Propositie 2.9 naar dit geval. Gezien de onderliggende symbolische dynamica van het systeem op een oneindige verzameling symbolen werkte, konden veel van de argumenten gebruikt in het eindige geval uit Hoofdstuk 2 niet overgedragen worden. Als gevolg hiervan konden we niet uitgaan van de methoden die in Hoofdstuk 2 gebruikt zijn, maar pasten we in plaats daarvan een aantal technieken aan die gebruikt zijn in de autonome situatie zoals in [FLMW10], die ook op een oneindige verzameling symbolen werkt.

In Deel II beschouwden we een uitgebreide klasse van abstracte deterministische dynamische systemen, waaronder systemen die van groot belang zijn zoals *Gibbs-Markov afbeeldingen* en *AFU afbeeldingen* (zie §1.2). Dit bevatte ook een aantal van de GLS getalsystemen die in Deel I gebruikt zijn om de niet-autonome systemen te beschouwen. We hebben deze uitgebreide klasse dynamische systemen gebruikt om stochastische variabelen te definiëren en de kans te analyseren dat de waarden van de (voldoende gecentreerde en geschaalde) partiële sommen van deze stochasten binnen een onwaarschijnlijk bereik vielen. Het bestuderen van dit soort gedrag van partiële sommen van stochastische variabelen wordt de *theorie van grote afwijking* genoemd en definieert in een klassieke situatie een onwaarschijnlijk bereik als alles wat niet beschreven wordt door de *centrale limietstelling*. We hebben *stabiele grote afwijkingen* beschouwd, die een onwaarschijnlijk bereik definiëren als dat wat niet beschreven wordt door de *stabiele limietstelling* – een generalisatie van de klassieke centrale limietstelling die stelt dat de verdelingen van partiële sommen onder bepaalde voorwaarden convergeren naar een *α-stabiele verdeling*, waar $\alpha \in (0, 2]$ de *stabiliteitsparameter* van de stabiele verdeling genoemd wordt en het vervallen naar nul van de uiteinden van de verdeling beheerst. Zie Figuur 1.6 voor twee voorbeelden van α-stabiele verdelingen: de Cauchy-verdeling (in het geval $\alpha = 1$) en de normaalverdeling (in het geval $\alpha = 2$). De volledige stelling voor stabiele grote afwijkingen van onafhankelijk en identiek verdeelde stochasten is een samenstelling van de resultaten in [Hey67; Nag69; Nag79; Roz89; Ber19], die we hebben samengevat in Stelling 1.8. In de afhankelijk stabiele situatie zijn er maar een klein aantal resultaten betreft grote afwijkingen; zie bijvoorbeeld [Gan96; MW13; Tak19] en de korte discussie hiervan in §4.2.

In Deel II waren we geïnteresseerd in grote afwijkingen in de context van α-stabiele verdelingen wanneer de klassieke centrale limietstelling niet van toepassing was. In het bijzonder presenteerden we een nieuw resultaat, namelijk Stelling 4.11, in de studie van grote afwijkingen in de context van α-stabiele stochastische variabelen met $\alpha \in (0, 1) \cup (1, 2]$ (met oneindige variantie) in Hoofdstuk 4. Huidige resultaten in de literatuur die dit poogden te behandelen, sloegen er of niet in het volledige bereik van grote afwijkingen te beschouwen, zoals in Stelling 1.8, of hadden voorwaarden die bijzonder moeilijk te verifiëren waren voor dynamische systemen. We verkregen Stelling 4.11 door eerst het geval van onafhankelijk en identiek verdeelde stochasten in Stelling 1.8 opnieuw te bewijzen op een manier die beter aansloot bij dynamisch gedefinieerde stochastische variabelen, en vervolgens dit bewijs en de aannames over de klasse van dynamische systemen te gebruiken om het afhankelijke geval af te leiden. Stelling 1.8 geldt voor het merendeel van het bereik van grote afwijkingen en voor alle stabiliteitsparameters $\alpha \in (0, 1) \cup (1, 2]$, waarbij het zeer ingewikkelde geval $\alpha = 1$ wordt vermeden, waarin de verdelingen van de geschaalde en gecentreerde partiële sommen convergeren naar een Cauchy verdeling. Dit lijkt het eerste resultaat te zijn over stabiele grote afwijkingen voor dynamische systemen dat in de buurt komt van de algemeenheid van Stelling 1.8 met verifieerbare voorwaarden. We pasten de resultaten toe op de klasse van Gibbs-Markov dynamische systemen, waaronder de kettingbreukafbeelding (zie §1.3.1 en Figuur 1.3(c)), waardoor we Gevolg 4.3 konden verkrijgen – een uitspraak over de grote afwijkingen van verschillende soorten gemiddelden van cijfers in kettingbreuken, evenals andere dynamische systemen zoals AFU afbeeldingen; zie §4.4.

De methode in het bewijs van Stelling 4.11 suggereerde voor het eerst een manier om foutmarges te krijgen voor algemene stabiele grote afwijkingen. In Hoofdstuk 5 vonden we daardoor foutmarges voor grote afwijkingen in de context van Hoofdstuk 4 in het specifieke geval van $\alpha = 1$, dat wil zeggen wanneer de verdelingen van de partiële sommen naar een Cauchy verdeling convergeren. Dit geval was veel ingewikkelder om te behandelen vanwege het inherent pathologische gedrag van de Cauchy verdeling, wat betekende dat de resultaten met de algemeenheid van die in Hoofdstuk 4 zeer moeilijk te verkrijgen waren. In plaats daarvan gebruikten we de methode uit Hoofdstuk 4 en beperkten we ons tot de onafhankelijk en identiek verdeelde context om in Stelling 5.1 foutmarges te krijgen voor grote afwijkingen in de specifieke context van Cauchy-verdeelde onafhankelijk en identiek verdeelde stochastische variabelen. Het resultaat dat we verkregen geldt voor het merendeel van de theorie van grote afwijkingen zoals gegeven in Stelling 1.8. We verkregen ook in Stelling 5.2 foutmarges over het volledige bereik door het type afwijkingen enigszins aan te passen door een grote bovengrens op de partiële sommen op te leggen; we noemen deze afwijkingen *quasi-grote afwijkingen*. Als toepassingen van de resultaten verkregen in dit hoofdstuk geven we in Gevolg 5.3 en Gevolg 5.19 de foutmarges voor de stabiele grote en quasi-grote afwijkingen en gaven we foutmarges voor de rekenkundige gemiddelden van cijfers die voorkomen in Lüroth ontwikkelingen en enkele andere GLS ontwikkelingen.

Acknowledgements

Who else to acknowledge first but my excellent supervisors Charlene and Dalia. Thank you both so much for your guidance and encouraging words during my PhD, for your supportiveness and for making me feel welcome from the very beginning of my move to the Netherlands – especially to Charlene for inviting me over for a meal on Christmas Day when I could not go home to see family due to COVID-19 restrictions. Thank you both for always pushing me further, which has allowed me to obtain results that I can be proud of. To Charlene, I will miss our many afternoon chats over hot chocolate about maths and unrelated things. Thank you for your patience despite the millions of typos and mistakes in the many drafts of my work! This thesis would certainly not be in the shape it is in now without you. To Dalia, your animated enthusiasm made it fun to learn an area of maths that I originally had little experience with. Thank you for building my confidence throughout – and thank you also for your fantastically quick response time out of hours!

Thank you to everyone on the committee who took time out of their busy schedules to be a part of my thesis defence.

To my collaborator Reza, thank you so much for the time and effort that you put into our work, and a special thank you for inviting me to Sweden to give my first seminar.

To my master's degree project supervisor and dynamical systems lecturer Dr Thomas Jordan, thank you for your inspiring lectures, advice on pursuing a PhD position and your thoughtfulness for sending me the information about the open position in Leiden. This thesis and my time in Leiden would not have happened without your initial guidance.

To the many people I met over the years at Leiden University and at the numerous conferences that I attended, particularly those at Fractals and Related Fields (Porquerolles 2022), Numeration (Liège 2023) and Women in Dynamics (Pisa 2024), and also to the PhD students of Birmingham university, thank you for your friendliness and fascinating conversations.

To my paranymph and colleague Sven, a huge thank you for the fun that we've had in and out of conferences, and for being there during my times of anxiety about my work. I often think back to the time that we were nearly stranded in rural Belgium – let's see if the same happens in Japan! Thank you also for your immense help with translating the samenvatting.

A big thank you to my paranymph and housemate Luuk, and to the other housemates I've had in the Netherlands including Antoni, Lieuwe, Paulina and Tom, who have put up with my English humour over the last four years. You all made living in the Netherlands a fun and unique experience that I won't ever forget.

To my family and especially my mum Michelle, dad Tony and brother Ryan, thank you for your constant support throughout my education and for an upbringing that allowed me to get to where I am now. Living abroad was not as easy as I thought it would be, but you were always there for phone calls when I needed them. An additional thank you to Ryan for designing the stylish artwork on the front cover of this thesis.

To my Bristol housemates Alex, Rudy and Sam, and friends Aidan, Claire, Rowan and Susanna, thanks for the many memorable get-togethers we've had in various places across Europe and for the online games and TV-watching sessions to distract me from my work.

To Hector and Lucas, thank you for the very many fun times in Brighton, for your hospitality, friendship and support, and for the many online games. Having a Brighton trip planned always gave me something to look forward to.

To Doug, thank you for always finding time in your busy schedule to watch our favourite YouTube videos and films together.

Finally, thank you to everyone who came to my thesis ceremony to celebrate with me, especially to those who came from abroad and who took time off work, and thank you to those who watched online as well.

Curriculum Vitae

Jonny Imbierski was born in 1998 in the town of Reading, England. He left The Forest School Winnersh sixth form centre with five AS-levels and four A-levels in June of 2016 in Maths, Further Maths, Chemistry, Biology and English Literature (AS). In September of 2016, Jonny began studying mathematics at the University of Bristol, graduating in June of 2020 with a first-class integrated master's degree (MSCi) including two first-class dissertations: the first supervised by Dr A. Booker on the Number Theory of Elliptic Curves (2019) and the second supervised by Dr T. Jordan on Dynamical Systems and Chaos (2020).

In August 2020, Jonny moved to the Netherlands to begin his PhD research projects at the Mathematics Institute of Leiden University. His doctoral thesis *Digits & Deviations of Dynamical Systems* covers the results obtained on two projects studying the digit frequencies of number systems under supervision of Dr C. Kalle and two projects on the large deviations of deterministic dynamical systems under supervision of Dr D. Terhesiu. These projects resulted in two published articles. Jonny gave a seminar at Uppsala University in Sweden on his research and attended numerous conferences and workshops across Europe, also presenting his research at the annual *Leiden and Birmingham Dynamics in Pure and Applied Mathematics* conference in Birmingham, UK. Jonny served as a teaching assistant in the master's degree courses *Ergodic Theory* (2020) in Amsterdam and *Measure Theoretic Probability* (2021 & 2022) in Utrecht as well as for the bachelor's degree course *Introduction to Measure Theory* (2021) in Leiden.

Alongside study for his integrated master's degree and PhD, Jonny additionally worked several hours a week as a private GCSE and A-level further maths tutor.