COMBINATORIAL AND FOURIER ANALYTIC L^2 METHODS FOR BUFFON'S NEEDLE PROBLEM

By

Matthew R. Bond

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Geometric Measure Theory and Harmonic Aathematics

2011

ABSTRACT

COMBINATORIAL AND FOURIER ANALYTIC L^2 METHODS FOR BUFFON'S NEEDLE PROBLEM

By

Matthew R. Bond

Figure 1: One of Count Buffon's beasts. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.

In recent years, progress has been made on Buffon's needle problem, in which one considers a subset of the plane and asks how likely "Buffon's needle" - a long, straight needle with independent, uniform distributions on its position and orientation - is to intersect said

set. The case in which the set is a small neighborhood of a one-dimensional unrectifiable Cantor-like set has been considered in recent years, and progress has been made, motivated in part by connections to analytic capacity [25].

Call the set E, the radius of the neighborhood ε, and the neighborhood E_ε. Then in some special cases [5][13][18], it has been confirmed that Buffon's needle intersects E_ε with probability at most $C|\log\varepsilon|^{-p}$, for $p > 0$ small enough, $C > 0$ large enough. In the special case of the so-called "four corner" Cantor set and Sierpinski's gasket, the lower bound $\frac{C\log|\log\varepsilon|}{|\log\varepsilon|}$ is known [3], replacing the previously-known lower bound $\frac{C}{|\log\varepsilon|}$ which is good for more general one-dimensional self-similar sets.

In addition, the stronger lower bounds are still good if one "bends the needle" into the shape of a long circular arc, or "Buffon's noodle." The radius one uses can be as small as $|\log\varepsilon|^{\epsilon_0}$, for any $\epsilon_0 > 0$, with the constant C depending on ϵ_0 [6]. It is unknown whether this condition or anything like it is necessary.

Work continues on generalizing the upper bound results.

For Rachel and Erica, my favorite couple ever. They don't have to read this document. They do have to visit me in Vancouver sometime, though.

ACKNOWLEDGMENT

I have received a lot of help from a lot of sources, and it would be a shame to cut this section short.

As early as my first year at MSU - this was 2005 - fellow graduate student Mike Dabkowski already knew that I was an analyst and never believed anything else. Around this time, there was not a large or active group of graduate students interested in analysis at MSU - there was only Alberto Condori, a few years ahead of us all and eager to share what he knew. Though Nick Boros and I were a few years behind and probably learned much more from Alberto than vice versa, I hope he got something out of it, even if nothing more than a few chances to practice some talks.

I do not know all of the details, but undoubtedly Alberto was instrumental in helping Nick and I find our advisor, Alexander Volberg. While it is quite common for graduate students to have to do a bit of begging, searching, and proving themselves before getting an advisor, we were lucky enough to have Dr. Volberg show up to substitute for our analsis class one day and start a habit of meeting with each of us once a week to describe interesting problems to us, give us papers to read, etc. We weren't even done with our qualifying exams yet, but already we had a distinguished and very helpful advisor. Alberto must have been confident enough in our skills from having graded our homework in our first year.

Dr. Volberg is also a vigorous proponent of his students, always looking for colloquia, workshops, etc. for his students to participate in, sharing with the organizers his - I trust well-deserved - high opinion of them. I may not have met half as many mathematicians as I have if it hadn't been for his level of advocacy.

The writers of my letters of recommendation for employment are my advisor, Yang Wang, Michael Lacey, and Ignacio Uriarte-Tuero, and the sponsoring scientist at University of British Columbia, where I begin my postdoctoral appointment next year, is Izabella Łaba. I have not read these letters of course, but they must have been very good letters. Though my papers summarized in this thesis are co-authored with only my advisor, having an appointment lined up - and quite importantly, at such a place and to work with such a person - is a relief and a motivation, making the circumstances around the writing of this thesis much brighter than they may have been otherwise.

I'd like to thank Nikos Pattakos and Alexander Reznikov for coming here and giving Nick and I more peers to talk to about analysis these last couple of years. Though they are a couple years behind us as naively measured by time spent in graduate school, they are very quick to pick up analysis, came knewing a lot already, and have contributed more than their fair share to the student analysis seminar, which now runs again at MSU. Thanks also to Clark Mussellman for recently making what Mike Dabkowski would undoubtledly call "the right choice" - that is, deciding to become an analyst. He has also contributed a couple of nice talks to our seminar.

Thanks also to Ignacio Uriarte-Tuero for co-organizing the student analysis seminar with me the last two years. Though he is not anyone's advisor, he has been consistently helpful to all of us analysis graduate students here at MSU in the few years he's been around. I am sure that he'll be well-suited to advising graduate students when the time comes.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Definitions, notations, results, and background

1.1 Buffon needle probability and Favard length

All sections of this thesis will have a great deal in common. In it, we will consider the **Buffon needle probability**, or **Favard length**, of a measurable[1] set $E \subseteq \mathbb{C}$. This quantity is defined as

$$Fav(E) := \frac{1}{\pi} \int_0^{\pi} |proj_\theta(E)| d\theta, \tag{1.1}$$

where $proj_\theta$ denotes orthogonal projection onto the line forming the angle θ with the positive real axis, and $|F|$ denotes the Lebesgue measure of F regarded a subset of $\mathbb{R}$. Pointwise, one defines $proj_\theta(re^{i\theta'}) := r \cdot \cos(\theta' - \theta)$.

The reason this is sometimes also called Buffon needle probability: after a normalization constant, it is the probability that "Buffon's needle" will intersect E when thrown, where

[1] In fact, we will only consider compact sets

"Buffon's needle" is a straight line which lands with independent, uniformly distributed location and angle with the positive real axis[2]. Favard length is known to be related to analytic capacity, a measure of how well E can "hide singularities of bounded analytic functions" - see [19], [25]. The sets we study in this thesis are of interest both for the analytic capacity problem and for Buffon's needle problem.

1.2 Homogeneous Cantor-like sets

As we consider Buffon's needle problem here, the sets which will play the role of E can either be thought of as "partially constructed" self-similar sets, or small neighborhoods of a self-similar set; they will be equivalent for our purposes, and we will freely conflate the two without harm. In particular, the self-similar sets we study will be **unrectifiable** self-similar sets of Hausdorff dimension one.

Definition 1. *For $s \geq 0$, we say that $H^s(E) < \infty$ if there is a constant M such that for all $\epsilon > 0$, E can be covered by countably many balls B_k of radii r_k smaller than ϵ such that $\sum_k r_k^s \leq M$. The* ***Hausdorff measure*** *$H^s(E)$ is defined to be the infimum over all such possible values of M. The* ***Hausdorff dimension****, $dim(E)$, is given by $dim(E) := \inf\{s : H^s(E) = 0\} = \sup\{s : H^s(E) = \infty\}$. When $0 < H^1(E) < \infty$, a set such that $Fav(E) = 0$ is called* ***purely 1-unrectifiable****, referred to in this thesis simply as* ***unrectifiable*** *or* ***1-unrectifiable****.*

The opposite of an unrectifiable set is a **rectifiable** set (properly speaking, an m-

[2] In the 18th century, "Buffon's needle" was a short, physical needle which was thrown repeatedly at a grid of uniformly spaced lines. By counting the proportion of the time the needle crossed a line, approximate values of π were found from a Monte Carlo type of formula.

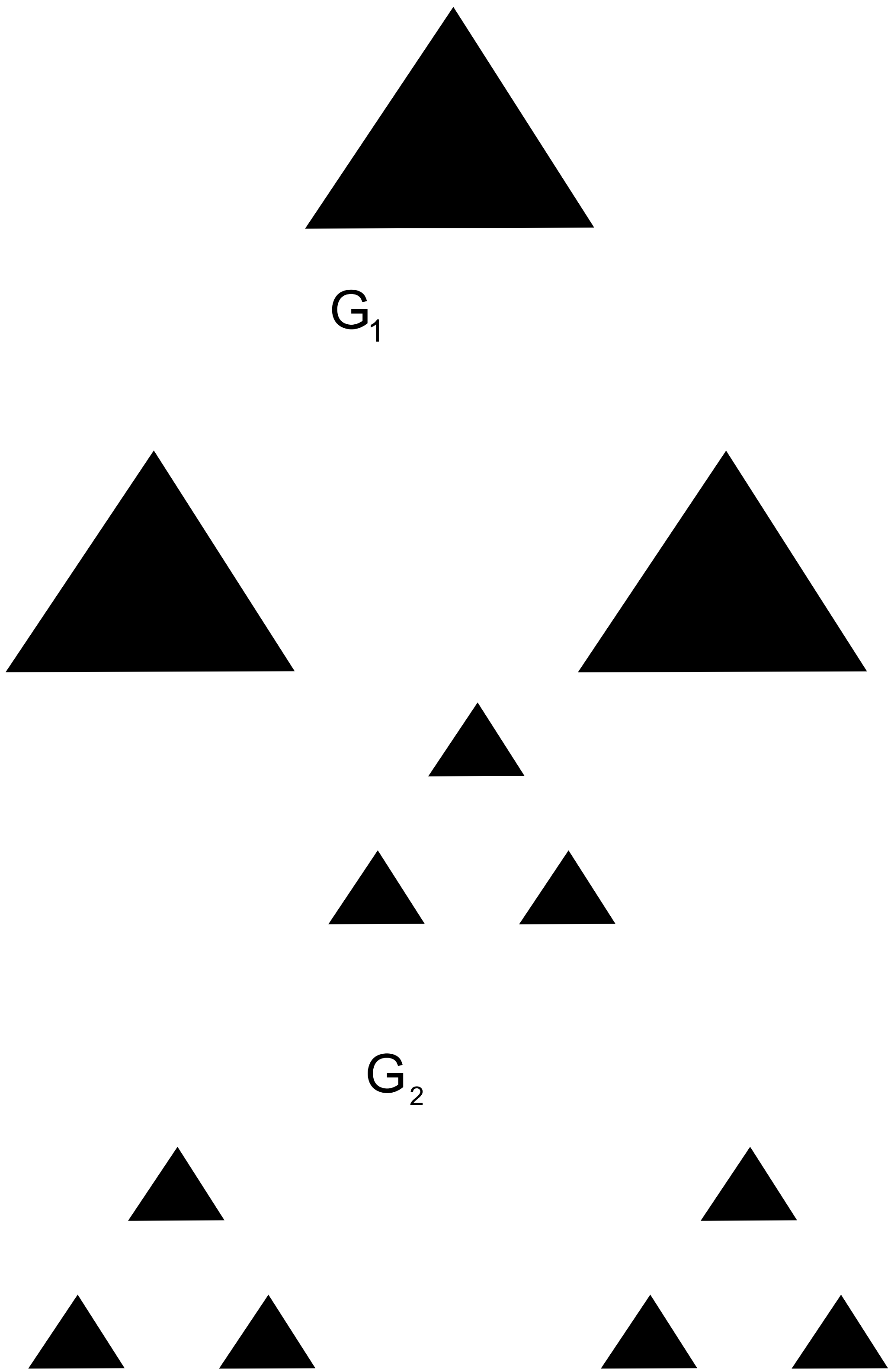

Figure 1.1: $\mathcal{G}_1$ and $\mathcal{G}_2$, stages 1 and 2 of the construction of Sierpinski's gasket.

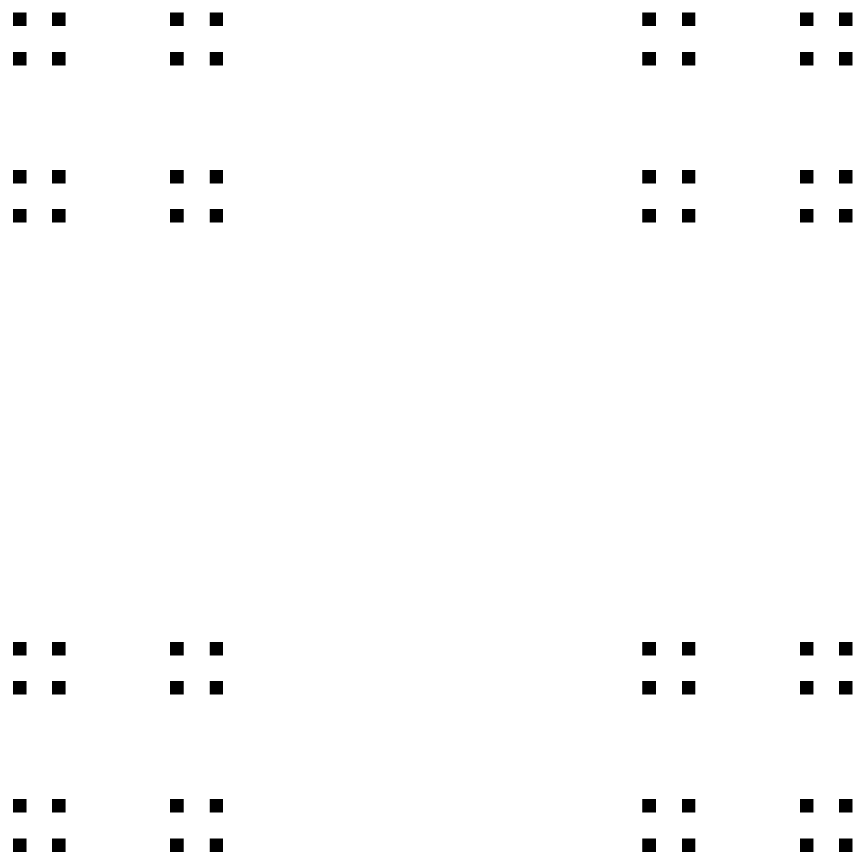

Figure 1.2: $\mathcal{K}_3$, stage 3 of the construction of the square Cantor set.

rectifiable set), such as an m-dimensional smooth manifold in $\mathbb{R}^n$, where $s = m \in \mathbb{N}$. H^m agrees with the usual notions of length, area, volume, etc. for $m = 1, 2, 3$, etc. when E is a smooth m-manifold. Therefore H^s generalizes such notions, as it is well known to be a Borel measure on $\mathbb{R}^n$, and s is allowed to be non-integer. For $m \in \mathbb{N}$, an m-rectifiable set is any countable union of Lipschitz images of $\mathbb{R}^m$ and H^m null sets. For equivalent definitions of rectifiability, see [16]; this thesis is concerned with unrectifiable sets of Hausdorff dimension 1.

Because of work done by Besicovitch, it is known that 1-unrectifiable sets $E \subset \mathbb{C}$ are those such that at least two orthogonal projections have zero Lebesgue measure, or equivalently, every Lipschitz curve meets E in a set of zero H^1-measure [16]. In fact, $dim(E) = 1$ is the critical case for Buffon's needle problem: if $dim(E) > 1$, $Fav(E) > 0$, and if $dim(E) < 1$, $Fav(E) = 0$; hence the role played by $H^1(E)$ in the definition of rectifiability. In general, if $H^m(E) < \infty$, E decomposes into m-rectifiable and m-unrectifiable parts.

A standard example of a 1-unrectifiable set is $\mathcal{K}$, the four-corner Cantor set. It is the unique compact invariant set[3] of the function system $S_k(z) = \frac{1}{4}z + c_k$, where $c_1 = (0, 0)$, $c_2 = (3/4, 0)$, $c_3 = (0, 3/4)$, $c_4 = (3/4, 3/4)$. Note also that

$$\mathcal{K} = \bigcap_n \mathcal{K}_n, \text{ where } \mathcal{K}_0 = [0,1] \times [0,1] \text{ and } \mathcal{K}_{n+1} = \bigcup_{k=1}^{4} S_k(\mathcal{K}_n).$$

We can do this with other function systems, too. Consider also Sierpinski's gasket, $\mathcal{G}$, the unique compact invariant set of the function system $S_k(z) = \frac{1}{3}z + r^{2\pi i(\frac{1}{2}+\frac{k}{3})}$, $k = -1, 0, 1$.

The most general case we will consider here: $S_k(z) = \frac{1}{L}z + z_k$, $k = 1, ..., L$. In such a case, the unique compact invariant set is called $\mathcal{J}$; this case contains both cases above. If

[3] E is an invariant set if $E = \bigcup_k S_k(E)$; such a compact set exists and is unique [16].

the centers z_k are not all collinear, then $\mathcal{J}$ is unrectifiable. By the word **homogeneous**, it is meant that instead of $S_k(z) = \frac{1}{L}z + z_k$, one could have had $S_k(z) = r_k z + z_k$ such that $\sum_{k=1}^{L} r_k = 1$, but in this thesis, we limit ourselves to the **homogeneous** case $r_k = \frac{1}{L}$.

Above $\mathcal{K}_n$ was defined as the union of all possible images of the convex hull of $\mathcal{K}$ under n-fold compositions of the similarity maps S_k; define $\mathcal{G}_n$ and $\mathcal{J}_n$ analogously, or see the following formal definition:

Definition 2. *Let $k = 1, 2, ..., L$. Let $\Sigma_n = \{1, 2, ..., L\}^n$. For any $\mathbf{v} = (v_1, v_2, ..., v_n) \in \Sigma_n$ and any $k \in \{1, 2, ..., L\}$, let $(\mathbf{v}, k) := (v_1, v_2, ..., v_n, k) \in \Sigma_{n+1}$. Let $S_{(k)} := S_k$, and for $v \in \Sigma_n$, let $S_{(\mathbf{v},k)} : \mathbb{C} \to \mathbb{C}$ be given by $S_{(\mathbf{v},k)} := S_k \circ S_{\mathbf{v}}$. Let $\mathcal{J}_0$ be the convex hull of $\mathcal{J}$. Let $\mathcal{J}_{\mathbf{v}} := S_{\mathbf{v}}(\mathcal{J}_0)$, and let $\mathcal{J}_n := \bigcup_{\mathbf{v}\in\Sigma_n} \mathcal{J}_{\mathbf{v}}$.*

For example, $\mathcal{K}_0 = [0, 1] \times [0, 1]$, and earlier we saw a picture of $\mathcal{K}_3$. $\mathcal{G}_0$ is a certain closed triangle which is "filled in" rather than "empty."

Remark 1. *For our intents and purposes, we more or less identify $\mathcal{J}_n$ with an appropriate neighborhood of $\mathcal{J}$. Define $B_\varepsilon(E) := \{z : dist(z, E) < \varepsilon\}$. Temporarily define $\tilde{\mathcal{J}}_n$ to be $B_{L^{-n}}(\mathcal{J})$. Then $c \cdot Fav(\tilde{\mathcal{J}}_n) \leq Fav(\mathcal{J}_n) \leq C \cdot Fav(\tilde{\mathcal{J}}_n)$. The reason for this is simply the fact that in either case, either of $\tilde{\mathcal{J}}_n$ and $\mathcal{J}_n$ can be covered by several translations of the other* [4]*. As such, we will no longer bother to distinguish between the two. This is also the reason why Buffon's needle problem for sets like $\mathcal{J}_n$ is often phrased for simplicity, "How likely is Buffon's needle to land* ***near*** *$\mathcal{J}$?"[20] rather than "How likely is Buffon's needle to intersect $\mathcal{J}_n$?" (See Figure 1.3)*

Remark 2. *$\mathcal{K}$,$\mathcal{G}$, and $\mathcal{J}$ were chosen for the notation as follows: $\mathcal{K}$ is K is for "Cantor"*[5]*;*

[4] the number of translates, and thus the constants c and C, depend on the eccentricity of the convex hull of $\{z_k\}_{k=1}^{L}$

[5] In [18], $\mathcal{C}$ has been used for the usual Cantor subset of $[0, 1]$

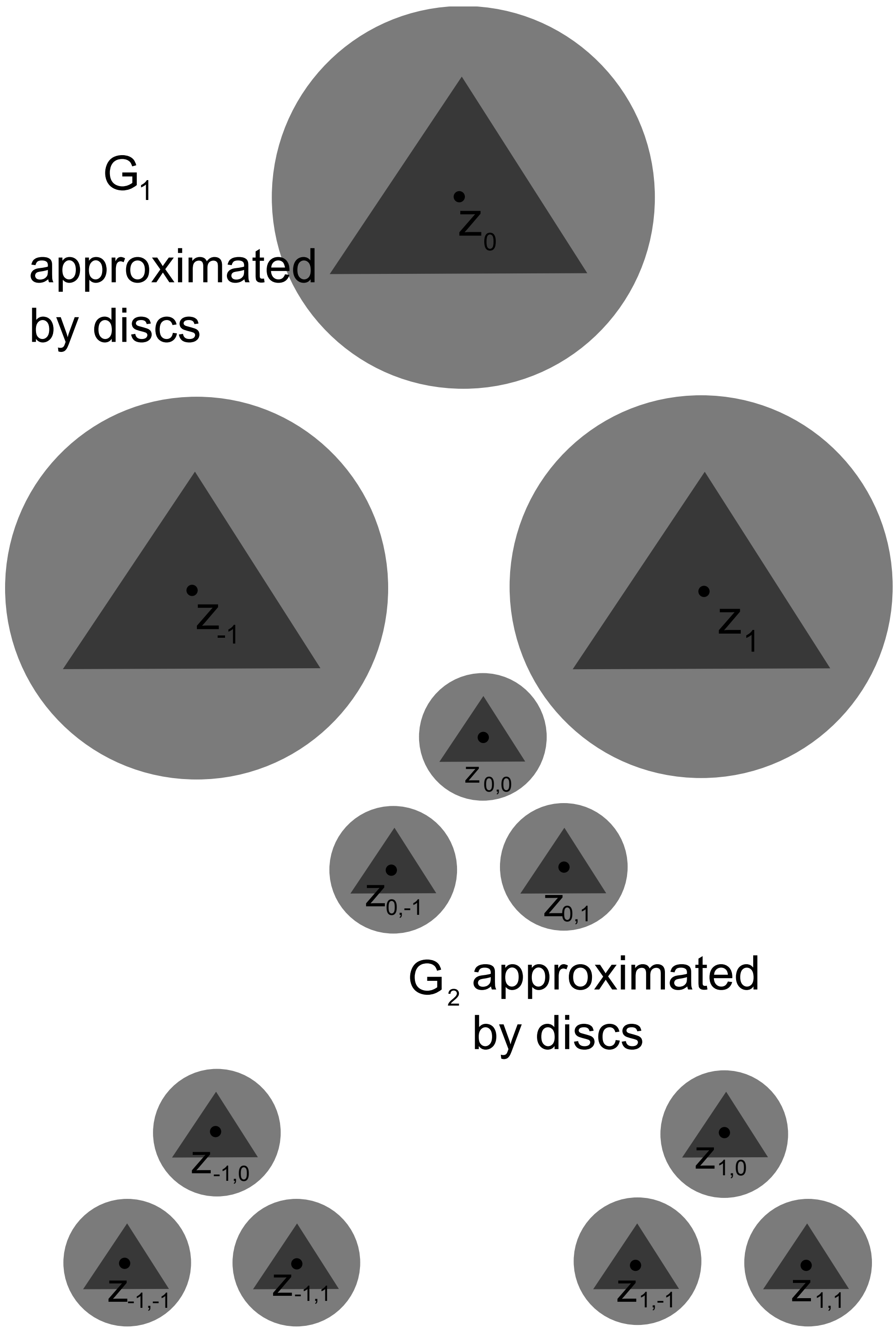

Figure 1.3: Several translations of the triangles cover the discs, so the lengths of the orthogonal projections are comparable.

$\mathcal{G}$ is G is for "Gasket"; since $\mathcal{G}$ is taken, $\mathcal{J}$ is J is for "General" – the phonics of the situation don't make it possible to reasonably misspell "gasket" as "jasket" or with any other first letter. So it goes.

1.3 Results for Buffon's needle problem

Let $A_n \lesssim B_n$ mean that there exists a constant C such that $A_n \leq CB_n$, where C must not depend on n.

Some known results:

Theorem 1. *[3], 2008*

$Fav(\mathcal{K}_n) \gtrsim \frac{\log n}{n}$. The same proof also shows that $Fav(\mathcal{G}_n) \gtrsim \frac{\log n}{n}$.

Theorem 2. *[18], 2008*

$Fav(\mathcal{K}_n) \lesssim \frac{1}{n^p}$ for any fixed $0 < p < \frac{1}{6}$, where the implied constant may depend on p.

Theorem 3. *[13], 2010*

Let $\mathcal{J}$ be as above. Additionally, suppose $\mathcal{J}$ is a product set, and let the coordinates of the z_k be rational. Suppose also that there exists a direction θ_0 such that $|proj_{\theta_0}(\mathcal{J})| > 0$. Then $Fav(\mathcal{J}_n) \lesssim \frac{1}{n^p}$, where p depends only on θ_0.

This thesis contains a generalization of [18] and [13] to $\mathcal{G}_n$. In addition, a weaker estimate is proved for $\mathcal{J}_n$. Current work between myself, Volberg, and Łaba continues toward proving the power estimate for $\mathcal{J}_n$ (without one or more of the additional conditions in [13]). Volberg and I have published work in this direction in [7], where the strong result was proved entirely for $\mathcal{G}_n$.

Theorem 14:

For some $c > 0$, $Fav(\mathcal{G}_n) \lesssim \frac{1}{n^c}$

Theorem 13:

For some $\epsilon_0 > 0$ depending on only the z_k defining $\mathcal{J}_0$, $Fav\mathcal{J}_n \lesssim e^{-\epsilon_0\sqrt{\log n}}$

This thesis also contains a generalization of [3], not for more general sets, but for a generalized notion of Favard length, called “Buffon noodle probability” or “circular Favard length”. The results for this problem, which we will prove in Chpaters 2 and 3, are also found in [6]. We refrain from stating the results here since they employ specialized notation for describing the “noodle”.

1.4 Counting function

All results concerning Buffon's needle problem for $\mathcal{J}_n$ employ the functions $f_{n,\theta} : \mathbb{R} \to \mathbb{R}$, called either the “counting function” or the “projection multiplicity function”. $\theta \in [0, \pi]$, $n \in \mathbb{N}$. Recall Definition 2 and B_ε from Remark 1.

$$f_{n,\theta} := \sum_{\mathbf{v}\in\Sigma_n} \chi_{proj_\theta(\mathcal{J}_\mathbf{v})} \tag{1.2}$$

In light of Remark 1, there exists an alternate form of f_n that is equivalent for our purposes. Recall: $S_k(z) = \frac{1}{L}z + z_k$. Then one simply redefines $\mathcal{J}_\mathbf{v} := B_{L^{-n}}(\sum_{k=1}^n L^{-k} z_k)$; i.e., for simplicity, one rescales a little and then approximates by discs. (See Figure 1.3)

We will not make use of the following fact, but it is interesting to notice the role of a function like $f_{n,\theta}$ in defining the **integralgeometric measure** of a set. Define $g_{n,\theta} = \sum_{p\in E} \chi_{proj_\theta(p)}$; i.e., $g_{n,\theta}$ counts the number of points in E “Buffon's needle” intersects if the “needle” goes through $xe^{i\theta}$ and is perpendicular to the line $re^{i\theta}$, $r \in \mathbb{R}$. Then the **integralgeometric** measure $I_1^1(E)$ is given by $I_1^1(E) = \int_0^\pi \int_{\mathbb{R}} g_{n,\theta}(x)\, dx\, d\theta$. With proper normalization constant, I_1^1 gives the length of any smooth curve $E \subset \mathbb{C}$, just like H^1. However, H^1 is positive on $\mathcal{J}$ (if the maps S_k satisfy the **open set condition**, [16]), whereas I_1^1 vanishes. Generalizations of H^s and I_1^1 appear in [16] and others.

1.5 Heuristics and napkin sketches

In Chapters 2 and 3, we will prove a lower bound in Buffon's noodle problem, for circular arcs and more general "noodles," respectively. In chapters 4 and 5, we will prove upper bounds in Buffon's needle problem. The lower bounds have a much simpler proof, remaining relatively painless even in the presence of an additional complication, the "bend" in the needle. The small bend in the needle is an unwelcome distraction for now, so forget it as we briefly discuss heuristics; in Chapters 2 and 3, we'll bend the needle as much as possible without damaging our argument.

Note that $||f_{n,\theta}||_1 = C$ for all n and θ; we can rescale and say $C = 1$. As n increases, however, $||f_{n,\theta}||_p$ is, in fact, an unbounded function for almost all θ for $p > 1$. This growth occurs because the L^1 mass concentrates on smaller sets as n increases; the effect is quite dramatic for the case $\theta = 0$ and $\mathcal{J}_n = \mathcal{K}_n$, and the squares stack up perfectly, the number of squares in each stack being 2^n. However, $\mathcal{K}_n$ also has $\tan(\theta) = 1/2$ as a clear counterexample (see Figure 2.2), and for $\mathcal{G}_n$, the (perhaps surprising) truth is that the exceptional θ such that $|proj_\theta(\mathcal{G})| > 0$ form a dense subset of $[0, \pi]$[12], see also [14] (the "obvious" examples for $\mathcal{G}_n$ are $\theta = 0, 2\pi/3, 4\pi/3$). As such, the quantitative Buffon's needle problem is inherently a bit finnicky.

Micro-theorem: If $|proj_\theta(\mathcal{J})| = 0$, then $||f_{n,\theta}||_p \to \infty$ as $n \to \infty$.

Proof: Since the $\mathcal{J}_n$ are compact and nested, $|proj_\theta(\mathcal{J}_n)| \to |proj_\theta(\mathcal{J})| = 0$. The result follows from Holder's inequality:

$$1 = ||f_{n,\theta}||_1 \leq ||\chi_{\text{supp}(f_{n,\theta})}||_q ||f_{n,\theta}||_p = |proj_\theta(\mathcal{J}_n)|^{1/q} ||f_{n,\theta}||_p. \tag{1.3}$$

□

If $\mathcal{J}_n$ had no self-similar structure, it would not be possible to state much in the way of a converse to the micro-theorem[6]. However, there is a converse.

Micro-theorem converse: If $||f_{N_0,\theta}||_\infty > K$ for some N_0, then $|proj_\theta(\mathcal{J}_n)| < \frac{C}{K}$ for some n large enough.

Sketch of proof: (See next figure) Fix θ. $\mathcal{J}_{N_0}$ has a stack of K discs above θ. So say that these K out of L^{N_0} discs are green, and label also its descendents green. Consider $\mathcal{J}_{j \cdot N_0}$. Each disc of $\mathcal{J}_{jN_0}$ is replaced by a rescaled copy of $\mathcal{J}_{N_0}$ when forming the set $\mathcal{J}_{(j+1)N_0}$. In particular, each white disc gives birth to a stack of K discs we label green, and $L^{N_0} - K$ white discs, and green discs give birth to only green discs. In this way, the total proportion of white discs is $\left(1 - \frac{K}{L^{N_0}}\right)^j$, and this proportion does not exceed the measure of their unified projection. In particular, the union of projected white discs has measure that approaches zero as $j \to \infty$.

On the other hand, the green discs do not unify to any more than C/K in the projection at any stage n, either. This ultimately follows from the Hardy-Littlewood theorem: If we sit at x, directly below a green disc at some stage JN_0, then find the smallest j such that

[6] Suppose no self-similar structure were available, and suppose still that $||f_{n,\theta}|| = 1$. One can use the Chebyshev's inequality to split to two level sets and show that $1 \cdot |\text{supp}(f_{n,\theta})| + (K-1)|\{x : f_{n,\theta} \geq K\}| \leq 1$. The resulting bound on $|\text{supp}(f_{n,\theta})|$ is not that strong if the height varies a lot. One could expand this to include many more level sets and try again, but then the problem seems more difficult than the one we started with.

in generation jN_0, this ancestor has turned green for the first time; taking an interval of width $2 \cdot L^{-jN_0}$ centered at x, we obtain an average value of $f_{jN_0,\theta}$ of size at least $K/2$. As this interval contains all projections of all children, and the union of the children equals the parent in L^1 mass, this estimate on the average remains valid; that is, all green discs live above places where $\mathcal{M}f_{n,\theta} \geq K/2$. ($\mathcal{M}$ is the usual Hardy-Littlewood maximal operator) Thus |union of projections of green discs| $\leq |\{x : \mathcal{M}f_{n,\theta}(x) > K/2\}| \leq \frac{C}{K}||f_{n,\theta}||_1 \leq \frac{C}{K}$.

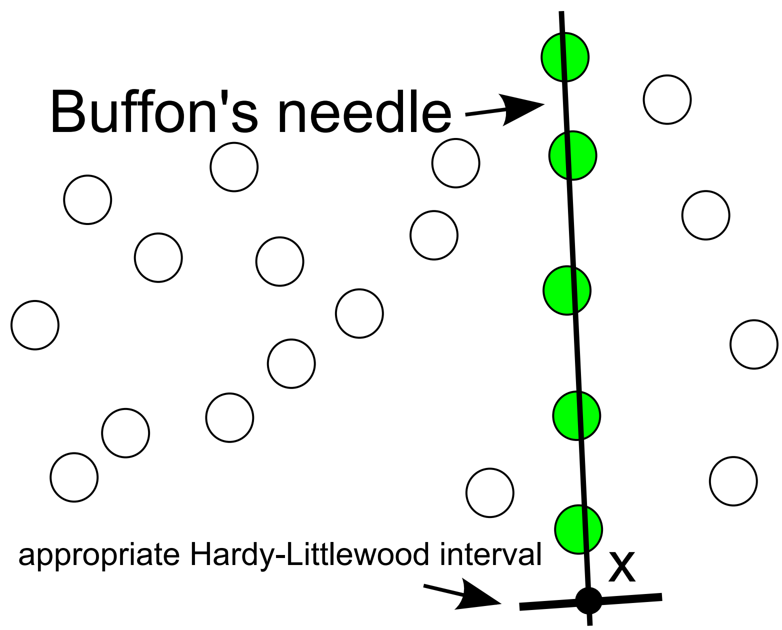

Figure 1.4: Discs turn green when the stack is tall for the first time; averages of $f_{n,\theta}$ on the illustrated interval will remain bounded below as n increases.

□

Note that if $||f_{n,\theta}||_\infty \to \infty$, the above theorem implies that the measure of $proj_\theta(\mathcal{J}_n) \to 0$ as $n \to \infty$ by monotonicity. [18] uses a sharpened form of this micro-theorem converse. The L^∞ condition is replaced with an L^2 condition, so that one finds many stacks of size K at various different generations, rather than just one stack in a single generation. By doing this, one can start out with a much larger proportion of green discs, leading to a much more rapid exponential rate of conversion of discs from white to green. Go green, indeed.

In all cases, we will use $p = 2$. Note that the micro-theorem (lower bound) was easier; as such, the bound it proves is easier to obtain. One needs only set the problem up with the aid of just one additional insight, and the rest is counting. The insight is simply partitioning the Favard length integral into well-chosen θ-intervals $I_1, I_2, ..., I_{\log\, n}$ and integrating the inequality $|proj_\theta(\mathcal{J}n)| \geq ||f_{n,\theta}||_2^{-2}$ in θ (this comes from (1.3)); a single integral with no θ partitioning exactly leads to inferior estimate $Fav(\mathcal{K}n) \gtrsim \frac{1}{n}$.

The upper bound is more finnicky, relying on some somewhat delicate Fourier analysis. The Fourier transform $\hat{f}_{n,\theta}$ is (away from ∞ equivalent to) a self-similar exponential polynomial, and ultimately, it is the bad behavior of its zeroes when $L > 4$ that delays us from proving more general results for now.

Chapter 2

The lower bound in Buffon's noodle problem - circular noodle case

In this chapter, we will state and prove Theorems 4 and 5.

In [6], a related **circular Favard length**, or **Buffon noodle probability**, was studied. To get circular Favard length Fav_σ instead of usual Favard length Fav, orthognal projection along the line is replaced by projection along a circular arc tangent to the line. Specifically, define the **noodles**

$$F_r(y) := r - \sqrt{r^2 - y^2} \tag{2.1}$$

Also define $\sigma_0(x, y) := (x - F_r(y), y)$, and $\sigma_\theta := R_{-\theta} \circ \sigma_0 \circ R_\theta$, where R_θ is clockwise rotation by the angle θ. [1](Also Figure 2.1. σ_θ depends on r, but r will be stated in each context and always refers to this implicit parameter wherever it appears.)

[1]Note that if we replace σ with the identity map, we are in the setting of [3]. We will often appeal to the $\sigma = Id$ case for intuition, while noting that the content of [6] is that the arguments of [3] carry over into [6] when $c_\epsilon n^\epsilon \leq r < \infty$ with the only difference being a change in the universal constants.

By definition, any $g : \mathbb{R} \to \mathbb{R}$ is a **noodle**, but we will use this language only for functions playing a role like that played by F_r in the definition of σ_θ.

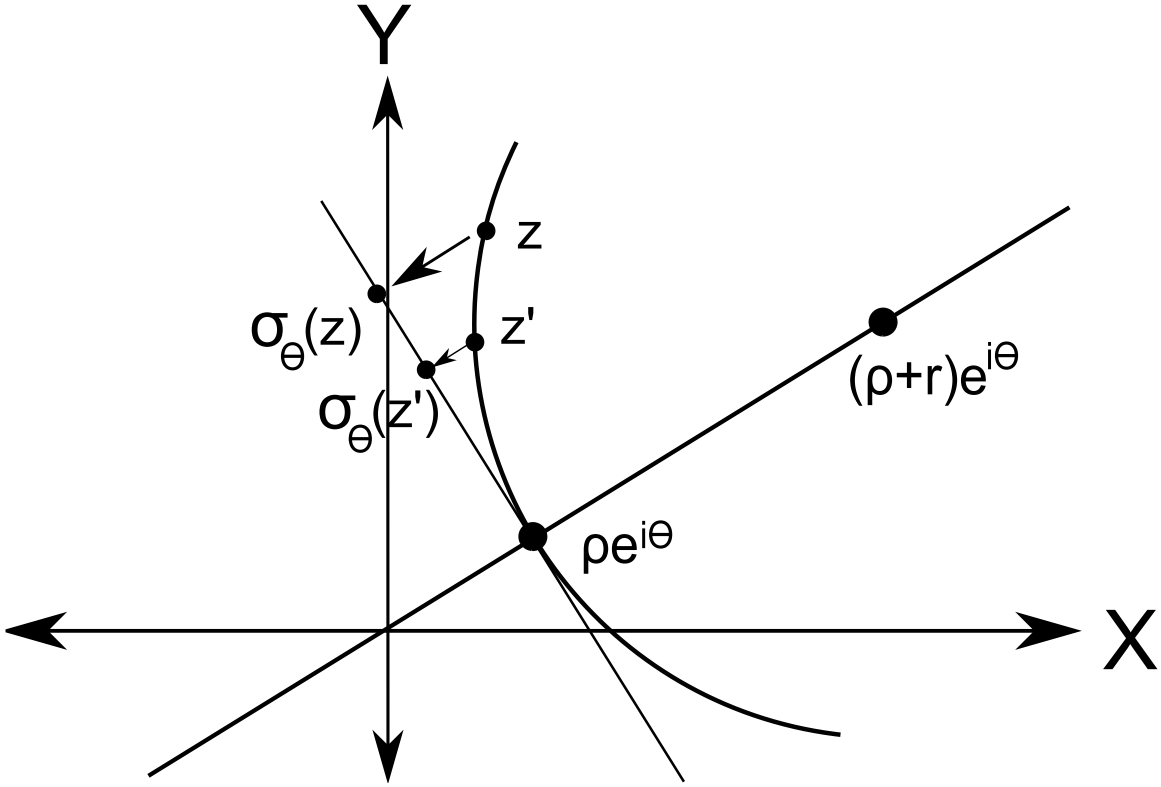

Figure 2.1: And illustration of the action of σ_θ.

Finally, let

$$Fav_\sigma(\mathcal{K}_n) := \frac{1}{\pi}\int_0^\pi |Proj_\theta(\sigma_\theta(\mathcal{K}_n))|\, d\theta$$

Remark 3. *Note that $Fav_\sigma(\mathcal{K}_n)$ is the $d\rho d\theta$ measure of the set of centers of circles of radius r that intersect $\mathcal{K}_n$, where such centers are parameterized by $z = (\rho + r)e^{i\theta}$. In addition to considering the $d\rho\, d\theta$ measure of this set, we may also naturally be interested in the $(r + \rho)\, d\rho\, d\theta$ measure of this set - that is, its area. Indeed, since r is much larger than the diameter of $\mathcal{K}_n$, $\rho + r \approx r$. This is the key convenience that makes our estimate for the circular noodle much easier and sharper by the arguments given here.*

Specifically, if $A \subseteq \{z \in \mathbb{C} : |z| \in (cr, Cr)\}$ is measurable and $|A|$ denotes its area, then

$|A| \approx r \int_0^{2\pi} \int_{\mathbb{R}} \chi_{\{\rho' : (\rho'+r)e^{i\theta} \in A\}}(\rho) d\rho d\theta$. *If we let* $A = \{z : z + re^{i\theta} \in \mathcal{K}_n$ *for some* $\theta \in \mathbb{R}\}$*, then this says, "The area of all points distance* r *away from* $\mathcal{K}_n \approx r\cdot$*the noodle probability of* $\mathcal{K}_n$*." Our main application, however, will be to a setting in which* A *is a set of circle centers like in Figure 2.1 - that is, the circle centered at* $z \in A$ *intersects two or more squares of* $\mathcal{K}_n$*.*

We will modify $f_{n,\theta}$ according to this problem. For any Cantor square $Q \subset \mathcal{K}_n$, let $\chi_{Q,\theta} := \chi_{Proj_\theta(\sigma_\theta(Q))}$.

$$f_{n,\theta,\sigma} := \sum_{\text{Cantor squares } Q \subset \mathcal{K}_n} \chi_{Q,\theta}.$$

$proj_\theta(\sigma_\theta(\mathcal{K}_n)) = \text{supp}(f_{n,\theta,\sigma})$, which we will also call $E_{n,\theta,\sigma}$.

Note that

$$\int_I |E_{n,\theta,\sigma}| \geq \frac{(\int_I \int_{\mathbb{R}} f_{n,\theta,\sigma} dx d\theta)^2}{(\int_I \int_{\mathbb{R}} f_{n,\theta,\sigma}^2 dx d\theta)}. \tag{2.2}$$

(This is (1.3) with a bend in the needle) The idea is to pick $\approx \log n$ many disjoint intervals I_j such that each such estimate gives

$$\int_{I_j} |E_{n,\theta,\sigma}| \, d\theta \geq \frac{C}{n}. \tag{2.3}$$

Summing over $j = 1, 2, ..., C_\epsilon \log n$, the result will be

Theorem 4. *For each* $c > 0$*, there exists* $C > 0$ *such that whenever* $r \geq cn^\epsilon$*,* $Fav_\sigma(\mathcal{K}_n) \gtrsim C_\epsilon \frac{\log n}{n}$*. Further, we may interpret* $Fav(\mathcal{K}_n)$ *to be* $Fav_\sigma(\mathcal{K}_n)$ *in the case* $r = \infty$*.*

If $r << n^\epsilon$, then we can still say something. We will prove the above theorem, but the following generalized theorem is proved by carefully examining for which values of j the estimate (2.5) holds in this general case. The lower bound on r is enough to make sure for Lemma 2.6 to holds.

Theorem 5. *For all $n \in \mathbb{N}$ and for all $r \lesssim n$,*

$$Fav_{\sigma}(\mathcal{K}_n) \gtrsim \frac{\log(r)}{n}$$

whenever $10 \leq r \leq n$.

Good intervals I_j can be found near $\theta = \arctan(1/2)$, because on this direction, $\mathcal{K}_n$ orthogonally projects onto a single connected interval, and the projected squares intersect only on their endpoints. These almost-disjoint projected intervals induce a 4-adic structure on the interval. Let us rotate the axes and redefine the old $\arctan(1/2)$ direction to be our new $\theta = 0$ direction. (See figure)

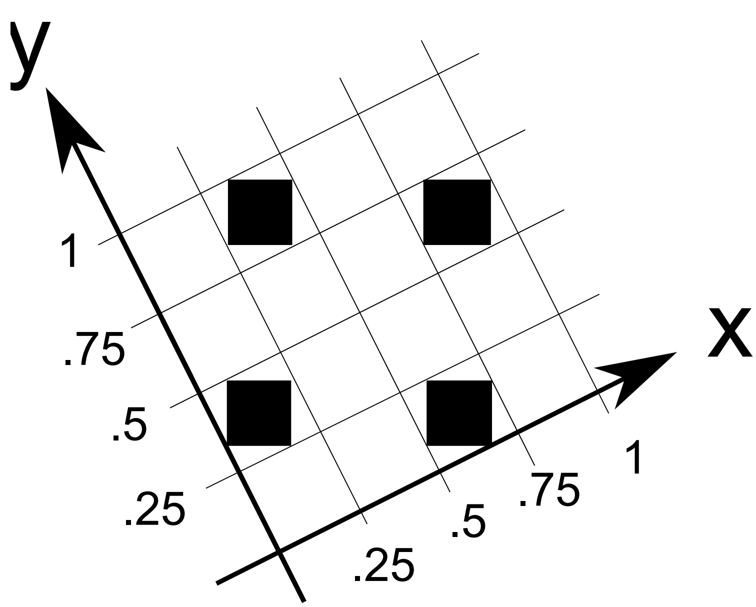

Figure 2.2: The projection to the x-axis is the entire interval; the same interval is covered by $proj_{\theta}(\mathcal{K}_n)$ for all n by self-similarity.

Definition 3. *Let* $I_j := [\arctan(4^{-j-1}), \arctan(4^{-j})]$, $3 < j < C_{\epsilon} \log n$.

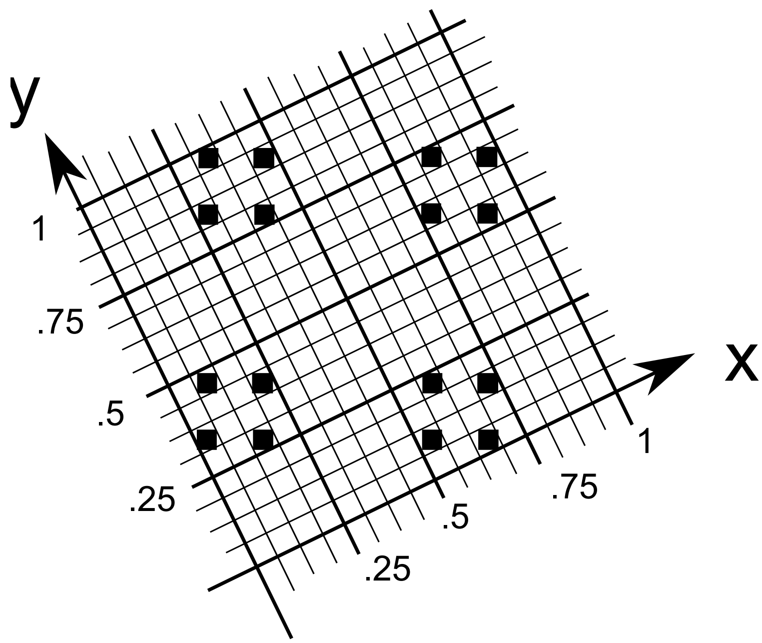

Figure 2.3: $\mathcal{K}_2$ in the adjusted coordinate system.

Then $I_{C_\epsilon \log n}$ will be the closest direction to 0, and it's reasonable to think that on average, each time j decreases by 1, I_j will grow by the factor 4, and for $\theta \in I_j$, $|E_{n,\theta,\sigma}|$ will decay no more than by a factor of 1/4, resulting in the persistence of (2.3). For individual θ, this is reasoning is completely invalid, but in the "average" sense as formulated by the integral $d\theta$ in (2.3), the reasoning is sound. (2.3) is, indeed, a theorem, which we will now prove:

Proposition 6. *For* $3 < j < C_\epsilon \log n$, $\int_{I_j} |E_{n,\theta,\sigma}| d\theta \gtrsim \frac{1}{n}$.

Recall (2.2). Trivially, $[\int_{I_j} \int f_{n,\theta,\sigma} dx\, d\theta]^2 \leq |I_j|^2 \cdot 1 \leq C4^{-2j}$, while

$$f_{n,\theta,\sigma}^2 = \sum_{Q,Q'} \chi_{Q,\theta}\chi_{Q',\theta} = \sum_{Q \neq Q'} \chi_{Q,\theta}\chi_{Q',\theta} + \sum_Q \chi_{Q,\theta}^2 .$$

Integrating over $I_j \times \mathbb{R}$, the latter diagonal sum becomes $C4^{-j} \lesssim n4^{-2j}$ (the inequality uses $j < \epsilon \log n < \log n$). When estimating the other integral, things become combinatorial

- most of these terms are identically 0 in $I_j \times \mathbb{R}$. It remains only to show

Proposition 7. *For $3 < j < C_\epsilon \log n$,*

$$\int_{I_j \times \mathbb{R}} \sum_{Q \neq Q'} \chi_{Q,\theta} \chi_{Q',\theta} d\,x\,d\,\theta \lesssim n 4^{-2j}$$

Definition 4. *$A_{j,k}$ is the set of pairs $P = (Q, Q')$ of Cantor squares such that there exists $\theta \in [0, \pi]$ such that the σ_θ images of the centers $z = x + iy$ and $z' = x' + iy'$ of Q and Q' have distance $4^{-k-1} \leq |y_{\sigma_\theta(z)} - y_{\sigma_\theta(z')}| \leq 4^{-k}$ and satisfy the condition on horizontal spacing*

$$4^{-j-1} \leq \left| \frac{x_{\sigma_\theta(q)} - x_{\sigma_\theta(q')}}{y_{\sigma_\theta(q)} - y_{\sigma_\theta(q')}} \right| \leq 4^{-j}. \tag{2.4}$$

We can think of 4^{-j} as being $\tan(\theta)$ for θ as in Figure 2.1. The terms in the sum of Proposition 7 are supported on the integration region only when $(Q, Q') \in A_{j-1,k}, A_{j,k}$, or $A_{j+1,k}$.

In [3], it was proved[2] that

$$|A_{j,k}| \leq 4^{2n-k-2j} \tag{2.5}$$

when $r = \infty$. The proof is very direct counting argument; roughly, if $(Q, Q') \in A_{j,k}$, then the most recent common ancestor of Q and Q' must have been of generation k, and Q and Q' must have been as close as possible in the x direction for the next j generations. That is, of the $4n$ bits of information needed to specify a pair $P \in A_{j,k}$, all but $k + 2j + c$ of them are free to vary, where c is an absolute constant.

[2] Actually, the bound and its proof on $|A_{j,k}|$ are entirely two-sided, but we do not need this fact.

To get the same $|A_{j,k}|$ estimate for $n^\epsilon \lesssim r < \infty$ as shown in [6], it suffices to compare the two cases with an application of the following lemma:

Lemma 8. *Let $\varepsilon > 0$ be small enough. Let $T : \mathbb{C} \to \mathbb{C}$ be such that $Lip(T - Id) < \varepsilon$. Then $\forall z, w \in \mathbb{C}$,*

$$|arg(z - w) - arg(T(z) - T(w))| < 2\varepsilon (mod\, 2\pi)$$

Proof. Write $z - w = \rho e^{i\theta}$, and let $\alpha := arg(z - w) - arg(T(z) - T(w))$.

$$arg(T(z) - T(w)) = arg((T - Id)(z) - (T - Id)(w) + (z - w)) = arg(\lambda \rho e^{i\beta} + \rho e^{i\theta})$$

for some $\lambda < \varepsilon, \beta \in [0, 2\pi]$. So $arg(T(z) - T(w)) = arg(\lambda e^{i\beta} + e^{i\theta})$ Then $|\alpha| \le \hat{\alpha}$, where $\tan(\hat{\alpha}) = \frac{\varepsilon}{1-\varepsilon} \Rightarrow |\alpha| < 2\varepsilon$. □

This is where the condition $r \gtrsim n^\epsilon$ is used: to make Lemma 8 sufficient for the purposes of relation 2.4. Since σ_θ is just σ_0 conjugated by an isometry, the Lipschitz constant for σ_θ (restricted to $\mathcal{K}_n$) is uniformly bounded by the size of the derivative of F_{r_n} on $[-2, 2]$.

Definition 5. *For any $P = (Q, Q') \in A_{j,k}$, let*

$$\nu_P := \int_0^\pi \int_{\mathbb{R}} \chi_{Q,\theta} \chi_{Q',\theta} dx d\theta.$$

We need the estimate

$$\nu_P \le C 4^{k-2n}, \tag{2.6}$$

since the integrand is supported only for angles belonging to I_{j-1}, I_j, and I_{j+1}. So we fix j and sum over k to get

$$\int_{I_j \times \mathbb{R}} \sum_{Q \neq Q'} \chi_{Q,\theta} \chi_{Q',\theta} d\theta dx \leq$$

$$\sum_{k=1}^{n-j+1} \max\{\nu_P : P \in A_{j',k} \text{ for some } j' = j-1, j, j+1\}(|A_{j-1,k}| + |A_{j,k}| + |A_{j+1,k}|)$$

$$\leq Cn4^{-2j}.$$

Here we used (2.5) and (2.6). The estimate (2.6) is elementary when $r = \infty$. It is true more generally than that, though.

Lemma 9. ***ν_P lemma for circles***

For any j, k pair P and $r \gtrsim n^{\epsilon}$, $\nu_P \lesssim C_{\epsilon} 4^{k-2n}$.

Proof. It may be useful to consult Figure 2.1 and Remark 3 now. If an arc of radius r intersects two Cantor squares, then the arc must be centered inside the intersection of two annuli whose radii are $r \pm 4^{-n}$, and whose centers are the centers of the two Cantor squares. So we want to prove that the area A of this intersection of annuli satisfies $A \leq Cr4^{k-2n}$.

Without the loss of generality, the squares are centered on the x-axis at 0 and at rx_0. We have $rx_0 \approx 4^{-k}$ and we define η by $r\eta = 4^{-n}$. So we need to show that $A \leq Cr^2\eta^2/x_0$. We can scale the problem by r. Thus if we let $r = 1$, then we need only show that if $x_0 < 1/2$, then $A \leq C\eta^2/x_0$. It will not hurt to let the inner radius be 1 rather than $1 - \eta^3$. Let $R = 1 + \eta$.

The area A is taken from the region bounded by $y = y_1 = \sqrt{1 - x^2}$, $y = y_1^+ = \sqrt{R^2 - x^2}$, $y = y_2 = \sqrt{1 - (x - x_0)^2}$, and $y = y_2^+ = \sqrt{R^2 - (x - x_0)^2}$.

[3] One may divide the annulus along the circle with radius one. The inner annulus can be rescaled to have inner radius 1, and the constants change negligibly

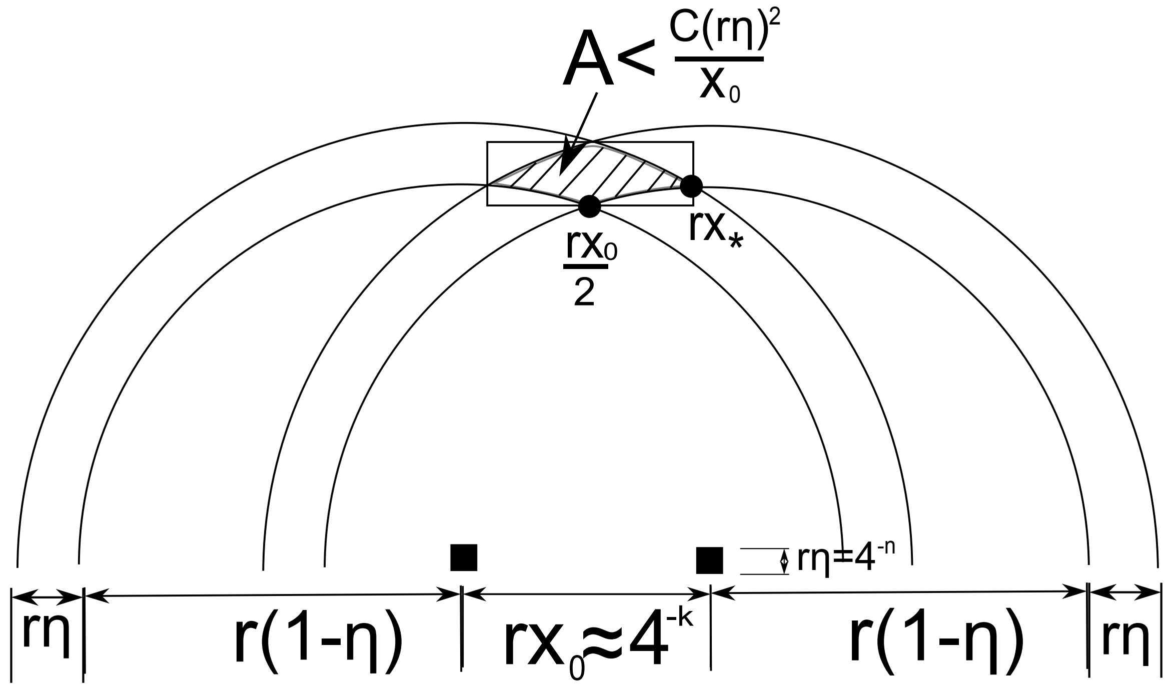

Figure 2.4: Only where the annuli intersect will we find centers of circles of radius r which intersect both Cantor squares. Approximation by a rectangle is sufficient to give the desired estimate.

$y_1^+ = y_2$ at a point we will call $x_* = \frac{1}{2}x_0 + \frac{1}{2x_0}\eta(2+\tau)$. So a rectangle which contains the area A has width $2(x_* - \frac{x_0}{2}) = \frac{1}{2x_0}\eta(2+\eta)$, and height $y_1^+(\frac{x_0}{2}) - y_1(\frac{x_0}{2})$. So we need only show that $y_1^+(\frac{x_0}{2}) - y_1(\frac{x_0}{2}) \leq C\eta$. To do this, we use the Mean Value Theorem on the function $s(x) = \sqrt{x}$.

$$y_1^+(\frac{x_0}{2}) - y_2(\frac{x_0}{2})) = \sqrt{R^2 - (\frac{x_0}{2})^2} - \sqrt{1 - (\frac{x_0}{2})^2}$$

$$\leq s'(1 - (\frac{x_0}{2})^2)(2\eta + \eta^2) \leq C\frac{\eta}{\sqrt{1-(\frac{x_0}{2})^2}} \leq C'\eta$$

Thus $A \leq C\eta^2/x_0$, as desired, so that $\nu_P \leq 4^{k-2n}$. □

This completes all proofs for this chapter. □

Chapter 3

The lower bound in Buffon's noodle problem - general noodles

3.1 General Buffon noodle probabilities and some preliminary reductions.

A notation for this chapter: $Proj_\theta(E)(x) := \chi_{proj_\theta(E)}(x)$. Aside from mathematical grammar and context, one can also tell what is referred to by $proj$ and $Proj$ by paying attention to capitalization.

In the previous chapter, our noodles were the functions F_{r_n}, playing a certain role in the expression σ_θ.

Let us define general noodle probabilities now. Because an arbitrary noodle does not have as many symmetries as a circular arc, a general noodle probability will need to integrate over three independent parameters: two real variables for where the noodle lands, and one for the orientation of the noodle. This serves two purposes: first, it better conforms to our intuition about what it means to randomly toss a possibly asymmetric noodle. Second, an extra variable of intergration allows us to more readily partition regions of integration into ones possessing symmetry. Our parameterization will have three real variables, ρ and θ like before, and a third parameter τ for translation orthogonal to the axis in the θ direction. In the case where the noodle is a circular arc, the two-parameter definition is equivalent for our purposes. It is clear that such a translation by τ of a circle is again a circle, and the information about whether this circle intersects a set can be transformed into an equivalent question in the two-parameter setting of the previous chapter. This is clearly not possible for noodles with less symmetry.

Let $g_\tau(y) := g(y-\tau)$. (If we have a family g_n of noodles, then we can write $g_{n,\tau}(y) := (g_n)_\tau(y) = g_n(y-\tau)$.) For a probability distribution P on $\mathbb{R}^2 \times S^1$, a set $E \subset \mathbb{C}$, and noodle g, we can define

$$Bu^g(E) = \int Proj_\theta \sigma_\theta^{g_\tau}(E)(x) dP(x,\tau,\theta).$$

We can choose an $L > 10$, say, and let P be normalized Lebesgue measure on $(-2,2) \times (-L,L) \times (0,2\pi)$, under which

$$Bu^g(E) = \frac{1}{16\pi L}\int_0^{2\pi}\int_{-L}^{L} |proj_\theta(\sigma_\theta^{g_\tau}(E))| d\tau d\theta = \frac{1}{16\pi L}\int_{-L}^{L} Fav_{\sigma_\theta^{g_\tau}}(E) d\tau.$$

(Note: $proj$ was lower-case, so $|proj(...)|$ denoted Lebesgue measure rather than pointwise absolute value of a function.)

Having done this, we will say that a noodles g_n are **undercooked** if $Bu^{g_n}(\mathcal{K}_n) \gtrsim \frac{\log n}{n}$. We call such a family of noodles **undercooked** because they are sufficiently close to being straight lines. It is not clear whether the “undercooking” condition is necessary or an artifact of the proof; on the other hand, it is clear that nearly-linear noodles are **undercooked** by the definition specified for some appropriate notion of “nearly linear”. We will prove one such result in this chapeter:

Theorem 10. *If* $||g_n'(y)||_\infty^4 \cdot ||g_n''||_\infty \leq 4^{-n}$ *and* $||g_n'(y)|| \leq \frac{1}{100n}$*, then the noodles* g_n *are undercooked.*

Remark 4. *In particular, this theorem implies that the* F_{r_n} *are undercooked if* $r_n \geq 4^{n/5}$*, which is a much stronger condition than that required by Theorem 4. Another example is* $g_n(y) = 4^{-n/2}\sin(4^{n/4}y)$*.*

Remark 5. *Using methods like those of Lemma 9 combined with the methods of this chapter, it may be possible to weaken the first condition of Theorem 10 in favor of conditions that require convexity and/or a condition on* $||g'''||_\infty$*. One would estimate noodle probability by estimating the distortion caused by thinking of noodle segments as segments of circular arcs rather than thinking of these segments as being “nearly linear”.*

Theorem 4 of Chapter 2 could be stated as follows:

Theorem 11. *The functions* F_{r_n}*, where* $F_{r_n}(y) := r_n - \sqrt{r_n^2 - y^2}$*, define undercooked noodles if* $r_n \gtrsim n^\epsilon$ *for some* $\epsilon > 0$*.*

The proof will be essentially the same, with the difference being that the corresponding ν_P lemma will be more tortuous. Define

$$\nu_{P,\sigma g} = \int_{-L}^{L}\int_0^{2\pi} |proj_\theta(\sigma_\theta^{g_\tau}(Q)) \cap proj_\theta(\sigma_\theta^{g_\tau}(Q'))| d\theta d\tau\,. \qquad (3.1)$$

Lemma 12.

$$\nu_{P,\sigma^g} \lesssim 4^{k-2n}.$$

There will be two main parts of the proof of the above ν_P lemma. If one of the two squares, Q, were centered at the origin and τ were fixed, the computation would merely amount to finding how often a needle close to the origin intersected the other square, Q'. We claim that this assumption can be justified if one folliates the domain appropriately and then changes variables. In fact, one can further assume that Q' lies on the negative $y-axis$ and that $\tau = 0$. Having done this, we will linearly approximate g_n and use the structure of the **shear group**(Section 3.2) to get our desired estimate. The idea is that we pick one of the two squares Q and partition the integration domain accoding to which point along the noodle punctures the center of Q. One can imagine dropping the noodle so it intersects Q, gluing this point of intersection in place, and then rotating the noodle around this point, asking how often the noodle hits the other square, Q'. Each of these positionings of the noodle can be expressed uniquely by a triple (τ, θ, x). If a particular point on the needle crosses the center of Q in a particular point along the noodle, then under this restriction, one thinks of θ as free and of x and τ as functions of θ.

Let us state the formulas. Fix a j,k pair Q, Q'. We will describe the portion of the domain of integration in which the noodle hits the center of a square Q at the same point $-\tau_0$ of the noodle. That is, if Q has center $z = \rho e^{i\theta_0}$, consider $\tilde{g} := g - g(-\tau_0)$ and $\sigma_\theta^{\tilde{g}\tau_0}$. For each θ, we need to find the unique x_θ and τ_θ such that the line centered at $x_\theta e^{i\theta}$ and with positive axis in the $\theta + \pi/2$ direction intersects z at $y = \tau_\theta - \tau_0$. In fact,

$$x_\theta = |z| \cos(\theta - \theta_0),$$

$$\text{and } \tau_\theta = \tau_0 - |z|\sin(\theta - \theta_0).$$

Then when computing

$$\int_0^{2\pi}\int_{x_\theta - a}^{x_\theta + a} Proj_\theta(\sigma_\theta^{\tilde{g}\tau_\theta}(Q))(x) Proj_\theta(\sigma_\theta^{\tilde{g}\tau_\theta}(Q'))(x)dxd\theta,$$

Without the loss of generality $z = 0$. That is,

$$\int_0^{2\pi}\int_{x_\theta - a}^{x_\theta + a} Proj_\theta(\sigma_\theta^{\tilde{g}\tau_\theta}(Q))(x) Proj_\theta(\sigma_\theta^{\tilde{g}\tau_\theta}(Q'))(x)dxd\theta$$
$$= \int_0^{2\pi}\int_{-a}^{a} Proj_\theta(\sigma_\theta^{\tilde{g}}(Q - z))(x) Proj_\theta(\sigma_\theta^{\tilde{g}}(Q' - z))(x)dxd\theta.$$

For $z =$ center of Q, and for fixed τ_0, define

$$D = \{\tau = \tau_0 - |z|\sin(\theta - \theta_0), |x - |z|\cos(\theta - \theta_0)| \le C4^{-n}, \theta \in (0, 2\pi)\}.$$

Then if

$$I_D(\tau_0) := \int_D Proj_\theta(\sigma_\theta^{\tilde{g}\tau_\theta}(Q'))(x)dxd\theta,$$

$\nu_{P,\sigma g} \le \int_{-L}^{L} I_D(\tau_0)d\tau_0$.

Putting this all together, we are seeking to prove that if in addition to the hypotheses of Theorem 10, $g(0) = 0$ and Q' is approximately at distance 4^{-k} from the origin, then $\int_0^{2\pi}\int_{-4^{-n}}^{4^{-n}} Proj_\theta(\sigma_\theta^{g}(Q'))(x)dxd\theta \le C4^{k-2n}$. Here we use that the σ-projection of a small square centered at the origin is essentially an interval around the origin regardless of

θ.

3.2 Some useful facts about shear groups

A few facts about the **shear groups** $\Sigma_\theta := \{\sigma_\theta^g : \; g : \mathbb{R} \to \mathbb{R} \text{ measurable}\}$ need to be stated (the operation is composition of the maps σ_θ^g). Below, g and h will be arbitrary noodles. Recall:

$$\sigma_0^g(x, y) := (x - g(y), y),$$

$$\sigma_\theta^g := R_{-\theta} \circ \sigma_0^g \circ R_\theta.$$

First, there is this simple fact for arbitrary functions g and h:

$$\sigma_\theta^g \circ \sigma_\theta^h = \sigma_\theta^{g+h} \tag{3.2}$$

Next, we show how shears by linear noodles behave. We can let E_θ be a family of subsets of $\mathbb{C}$, but for our application, we will fix $E_\theta = \mathcal{K}_n$. For $g(y) = b$, we get

$$proj_\theta(\sigma_\theta^g(E_\theta)) = proj_\theta(E_\theta) - b \tag{3.3}$$

For $g(y) = my$, $\alpha := \arctan m$, we get

$$proj_\theta(\sigma_\theta^g(E_\theta)) = \left(\frac{proj_{\theta-\alpha}(E_\theta)}{\cos(\alpha)}\right) = \sqrt{1+m^2}\,proj_{\theta-\alpha}(E_\theta) \tag{3.4}$$

Remember that the lower-case *proj* denotes a set, not a characteristic function. That is, keep in mind that we defined $Proj_\theta(E) := \chi_{proj_\theta(E)}$

For for $g(y) = my + b$ and given $x \in \mathbb{R}$, we can see that $x \in proj_\theta(\sigma_\theta^g(E_\theta))$ if and only if $\frac{x+b}{\sqrt{1+m^2}} \in proj_{\theta-\alpha}(E_\theta)$. Thus for any measurable $A \subset \mathbb{R}$,

$$\int_0^{2\pi} \int_A Proj_\theta(\sigma_\theta^g(E_\theta))(x)dxd\theta = \sqrt{1+m^2} \int_0^{2\pi} \int_{\frac{1}{\sqrt{1+m^2}}(A+b)} Proj_\theta(E_{\theta+\alpha})(x)dxd\theta \tag{3.5}$$

3.3 Proof of the ν_P Lemma for general noodles

Now to prove Lemma 12. Lemma 8 will be used several times without explicit mention. Lipschitz constants are clearly gotten from Taylor estimates on g.

The rough idea of this proof: the set of parameters for which two squares are simultaneously punctured by the needle may be translated considerably in parameter space by the shears, but it cannot be dilated by too much. Since a shear with small curvature is well-approximated by a linear shear, the result will follow. Let $\lambda' = ||g_n'(y)||_\infty$ and $\lambda'' = ||g_n''(y)||_\infty$. Let Q and Q' be centered at $(0,0)$ and $(0,-L)$, respectively, where $L \approx 4^{-m}$. Note that

$$\nu_{P,\sigma_\theta^g} \leq C4^{-n}|\{\theta : \sigma_\theta^g Q \cap \sigma_\theta^g Q'\}| \leq C4^{-n}(4^{k-n} + \lambda').$$

We need this quantity to be $< C4^{k-2n}$. This task is already done if $\lambda' \leq 4^{k-n}$, so assume the opposite. Now for such θ we have $|\theta| < C(4^{k-n} + \lambda') < C\lambda'$.

For these θ, rotation $R_\theta(Q)$ is in the band $L-\delta \leq y \leq L+\delta$, for $\delta = 4^{-n}+L(1-\cos(C\lambda'))$, giving $\delta \leq C\max\{4^{-n}, L\lambda'^2\}$. Now transform the integral using the shear group. Let $l(y)$ linearly approximate $g(y)$ at $y = L - \delta$, with $l(y) = my + b$. Note that $|b| \leq CL\lambda'$. Let

$\epsilon(y) := g(y) - l(y)$ on $[L-\delta, L+\delta]$ and extend ϵ continuously to be constant elsewhere. Then, with $b' := b/\sqrt{1+m^2}$:

$$\nu_{P,\sigma^g} = \int |proj_\theta(\sigma^g_\theta(Q')) \cap proj_\theta(\sigma^g_\theta(Q))| d\theta \leq \int_0^{2\pi} \int_{-4^{-n}}^{4^{-n}} Proj_\theta(\sigma^g_\theta(Q'))(x)\, dx d\theta$$

$$= \int_0^{2\pi} \int_{[-4^{-n}, 4^{-n}]} Proj_\theta(\sigma^l_\theta(\sigma^\epsilon_\theta(Q')))\, dx d\theta \leq$$

$$C \int_0^{2\pi} \int_{[b'-2\cdot 4^{-n}, b'+2\cdot 4^{-n}]} Proj_{\theta-\alpha}(\sigma^\epsilon_\theta(Q'))\, dx d\theta\,.$$

Changing variable, we see that this is at most

$$C \int_0^{2\pi} \int_{[b'-2\cdot 4^{-n}, b'+2\cdot 4^{-n}]} Proj_\theta(\sigma^\epsilon_{\theta+\alpha}(Q'))\, dx d\theta\,.$$

Let $\Gamma := \{\theta : proj_\theta(\sigma^\epsilon_{\theta+\alpha}(Q')) \cap [b' - 2\cdot 4^{-n}, b' + 2 \cdot 4{-}n] \neq \emptyset\}$, and let $z := (0, -L)$. If $\theta \in \Gamma$, then $proj_\theta(\sigma^\epsilon_{\theta+\alpha}(z)) \in [b' - 3\cdot 4^{-n}, b' + 3 \cdot 4^{-n}]$.

Since $|\epsilon'(y)| < C\delta\lambda''$, it follows that $|\epsilon(y)| < C\,\delta^2\lambda'' < C\,L^2\lambda'^4\lambda'' < C4^{-n}$. So $|\sigma^\epsilon_{\theta+\alpha}(z) - z| < c\,4^{-n}$, and hence $|proj_\theta(\sigma^\epsilon_{\theta+\alpha}(z)) - proj_\theta(z)| \leq C\,4^{-n}\ \forall\theta \in \Gamma$. So

$$\Gamma \subseteq \{\theta : proj_\theta(z) \in [b' - C4^{-n}, b' + C4^{-n}]\} = \{\theta : L\sin\theta \in [b' - C4^{-n}, b' + C4^{-n}]\},$$

which implies:

$$|\Gamma| \leq C|\{\theta : \sin\theta \in [b/L - C4^{k-n}, b/L + C4^{k-n}]\}|. \tag{3.6}$$

Since $b < CL\lambda'$ and $4^{k-n} \leq \lambda' \leq \frac{C}{n}$, $\sin\theta \approx \theta$, and we get $|\Gamma| \leq C4^{k-n}$, completing the proof of the ν_P Lemma. $\square$

Theorem 3.1 follows as well.

Chapter 4

The upper bound in Buffon's needle problem - Sierpinski's gasket

Here we will prove Theorem 14. The argument is elaborated in slightly more detail in [5]. It may be instructive to compare the general case $\mathcal{J}_n$ with the special case $\mathcal{G}_n$ we consider here. When a theorem for $\mathcal{G}_n$ is treated as a special case to later be proved for $\mathcal{J}_n$, the correspondence will be noted by theorem number and then omitted, both to minimize repetition and to prevent readers from missing the forest for the trees. The following preamble serves equally well for the gasket and for the general case.

In Chapter 1, we saw that the growth of $||f_{n,\theta}||_p \to \infty$ was equivalent to the decay of $|proj_\theta(\mathcal{J}_n)| \to 0$. We stated and proved a **micro-theorem** and its converse. Chapters 2 and 3 used the idea of the micro-theorem, and here we employ a stronger form of the micro-theorem converse, Theorem 27.

$\int_0^\pi |proj_\theta(\mathcal{J}_n)|\, d\,\theta \to 0$, as guaranteed by the Besivotich theorem, is only an average - we also noted that [12] and [14] show that the exceptional angles θ where $|proj_\theta(\mathcal{J})| > 0$

can be dense in $[0,\pi]$. The "set of bad angles at stage n", E_n (or just E), necessarily has small measure when n is large. Since E is small and $|proj_\theta(\mathcal{J}_{n^2})|$ is small for $\theta \in E^c$, the integral can be split according to the cases $\theta \in E$ and $\theta \in E^c$:

$$\pi \cdot Fav(\mathcal{J}_{n^2}) = \int_E |proj_\theta(\mathcal{J}_{n^2})| d\theta + \int_{E^c} |proj_\theta(\mathcal{J}_{n^2})| d\theta$$

$$\leq |E| + (\pi - |E|) \cdot \sup_{\theta \in E^c} |proj_\theta(\mathcal{J}_{n^2})| << 1$$

(The exponent 2 is chosen somewhat arbitrarily.)

Quntitative control over E has been accomplised to some extent:

Theorem 13. *For all $n \in \mathbb{N}$,*

$$Fav(\mathcal{J}_n) \leq e^{-\epsilon_0 \sqrt{\log n}}.$$

Theorem 14. *There is a $p_0 > 0$ such that for all $p < p_0$, there exists $C_p > 0$ such that for all $n \in \mathbb{N}$,*

$$Fav(\mathcal{G}_n) \leq C_p n^{-p}.$$

Further, one may take $p_0 = \frac{1}{[2\log_3(169)]^{-1}+1} \approx \frac{1}{10.262}$, so $p = \frac{1}{11}$ is sufficient.[1]

A method for controlling $|E|$ originates with [18]. One takes the Fourier transform of $f_{n,\theta}$ in the length variable and takes a sample integral of $|\hat{f}_{n,\theta}(x)|^2$ over a chosen small interval I where $\int_{E\times I} |\hat{f}_{n,\theta}(x)|^2 d\theta dx$ is small. One then shows that there is a $\theta \in E$ such that $\int_I |\hat{f}_{n,\theta}(x)|^2 dx$ is large relative to $|E|$, and so $|E|$ must be small.

[1] It is not suspected that this value of p_0 is sharp; on the other hand, $p = 1$ is impossible because of the argument of [3], $Fav(\mathcal{G}_n) \gtrsim \frac{\log n}{n}$.

$\hat{f}_{n,\theta}$ is a decay factor times a finite self-similar product $\prod_k \varphi_\theta(L^{-k}y)$ of trigonometric polynomials φ_θ. The most direct methods don't accomplish the estimate all at once; the high-frequency terms form a product $P_{1,\theta}$ such that $\int_I |P_{1,\theta}|^2 dx$ is large, and the danger is that perhaps the zeroes of the low-frequency terms $P_{2,\theta}$ might be located such that $\int_I |P_1 P_2|^2 dx$ is small. In [18], the four frequencies of φ_θ were symmetric around 0, allowing the terms to simplify to two cosines, and trigonometric identities allowed the whole product to be estimated by a single sine term. In [13], an analogous role was played by tilings of the line on the non-Fourier side by $proj_{\theta_0}(\mathcal{J}_n)$ in the special direction θ_0, and the product structure of $\mathcal{J}_n$ allowed for a change and separation of variables. Separating variables is more difficult when there is no product structure. The simplest case without the product structure is the Sierpinski gasket $\mathcal{G}$ considered in this chapter. We give a sketch of the power estimate (proven in detail in [5]), which is based on the fact that zeroes of $\varphi(3^k\cdot)$ are separated away from each other for different values of k. This special structure of zeros (we call it "analytic tiling" after [13]) is not always available for all angles. We have not yet found an adequate substitute for it in the general case, and this is why the for the general case we still only have $Fav(\mathcal{J}_n) \leq e^{-\epsilon_0\sqrt{\log n}}$. Rather strangely, a claim in the spirit of the Carleson Embedding Theorem, in the form of Lemma 40, plays an important part in our reasoning in the general case of Chapter 5. Because the Fourier transform turns stacks of discs (i.e., sums of overlapping characteristic functions) into clusters of frequencies, this lemma provides important upper bounds when θ belongs to E.

4.1 Reductions and main Fourier-analytic argument

$B(z_0, r) := \{z \in \mathbb{C} : |z - z_0| < r\}$. For $\alpha \in \{-1, 0, 1\}^n$ let

$$z_\alpha := \sum_{k=1}^{n} (\frac{1}{3})^k e^{i\pi[\frac{1}{2}+\frac{2}{3}\alpha_k]}, \ \mathcal{G}_n := \bigcup_{\alpha \in \{-1,0,1\}^n} B(z_\alpha, 3^{-n}).$$

This set is our approximation of a $\mathcal{G}_n$; recall Remark 1. We may still speak of the discs $B(z_\alpha, 3^{-n})$ as "Sierpinski triangles." The result for the Sierpinski gasket is the following:

Theorem 15. *For some $p > 0$, $Fav(\mathcal{G}_n) \lesssim \frac{1}{n^p}$.*

We will simplify the proof by picking specific values for constants; at the end of this paper, a short remark shows how to recover the full range $p < p_0$ as in Theorem 14. As in Chapter 1, let

$$f_{n,\theta} := \sum_{\text{Discs D of } \mathcal{G}_n} \chi_{proj_\theta(D)}.$$

Self-similarity allows us to write $f_{n,\theta}$ in a form well-suited to Fourier analysis:

$$f_{n,\theta} = \frac{1}{2}\nu_n * 3^n \chi_{[-3^{-n},3^{-n}]},$$

where

$$\nu_n := *_{k=1}^{n} \widetilde{\nu}_k$$

$$\widetilde{\nu}_k := \frac{1}{3}\left[\delta_{3^{-k}cos(\pi/2-\theta)} + \delta_{3^{-k}cos(-\pi/6-\theta)} + \delta_{3^{-k}cos(7\pi/6-\theta)}\right]$$

For $K > 0$, let $A_K := A_{K,n,\theta} := \{x : f_{n,\theta} \geq K\}$. $\mathcal{L}_{\theta,n} := \text{proj}_\theta(\mathcal{J}_n) = A_{1,n,\theta}$.

For our result, some maximal versions of these are needed.[2]:

$$f^*_{N,\theta} := \max_{n\le N} f_{n,\theta},\;\; A^*_K := A^*_{K,N,\theta} := \{x : f^*_{N,\theta} \ge K\} = \bigcup_{n=1}^{N} A_{K,n,\theta}.$$

Also, let $E := E_N := \{\theta : |A^*_K| \le K^{-3}\}$ for $K = N^{\epsilon_0}$, where $\epsilon_0 > 0$ is a small enough absolute constant.[3]

Later, we will jump to the Fourier side, where the function

$$\varphi_\theta(x) := \frac{1}{3}\left[e^{-i\cos(\pi/2-\theta)} + e^{-i\cos(-\pi/6-\theta)} + e^{-i\cos(7\pi/6-\theta)}\right]$$

plays the central role: $\widehat{\nu_n}(x) = \prod_{k=1}^{n} \varphi_\theta(3^{-k}x)$.

Let $\mathcal{L}_{n,\theta} := proj_\theta(\mathcal{G}_n)$. The following constitutes the content of **Theorem 27**: If $\theta \notin E_N$, then $|\mathcal{L}_{NK^3,\theta}| \le \frac{C}{K}$. (The same is true of $\mathcal{J}_n$ when everything is again defined analogously)

Now Theorem 15 follows from the following:

Theorem 16. *Let $\epsilon_0 < 1/\log_3(169)$, sufficiently, $\epsilon_0 \le 1/9.262$. Then for $N >> 1$, $|E| < N^{-\epsilon_0} = \frac{1}{K}$.*

This is better than what has currently been done for $\mathcal{J}_n$, Theorem 21. It turns out that L^2 theory on the Fourier side is of great use here. The following is later proved as **Theorem 31** in Section 5.2.2: For all $\theta \in E_N$ and for all $n \le N$, $||f_{n,\theta}||_2^2 \lesssim K$. (The implied constant depends only on the set of self-similarities)

[2]See the **micro-theorem converse** of 1.5 for the rough idea why this is useful, and then Theorem 27 for the formal statement of what one can then say

[3]To get the sharpest exponent in Theorem 14, K^{-3} should be replaced by $K^{-\tau}$ for $\tau > 2$ arbitrary.

One can then take small sample integrals on the Fourier side and look for lower bounds as well. Let $K = N^{\epsilon_0}$, and let $m = 2\epsilon_0 \log_3 N$. Theorem 31 easily implies the existence of $\tilde{E} \subset E$ such that $|\tilde{E}| > |E|/2$ and number n, $N/4 < n < N/2$, such that for all $\theta \in \tilde{E}$,

$$\int_{3^{n-m}}^{3^n} \prod_{k=0}^{n} |\varphi_\theta(3^{-k}x)|^2 dx \lesssim \frac{Km}{N} \lesssim N^{\epsilon_0 - 1} \log N.$$

The number n does not depend on θ; n can be chosen to satisfy the estimate in the average over $\theta \in E$, and then one chooses $\tilde{E}$. Let $I := [3^{n-m}, 3^n]$.

Now the main result amounts to this (with absolute constant α large enough):

Theorem 17.

$$\exists \theta \in \tilde{E} : \int_I \prod_{k=0}^{n} |\varphi_\theta(3^{-k}x)|^2 dx \gtrsim 3^{m-2\cdot\alpha m} = N^{-2\epsilon_0(2\alpha-1)}.$$

The result: $\log N \gtrsim N^{1-\epsilon_0(4\alpha-1)} = N^\delta$, where $\delta > 0$. Then it follows that $N \leq N^*$. Now we sketch the proof of Theorem 17. We split up the product into two parts: high and low-frequency:

$$P_{1,\theta}(z) = \prod_{k=0}^{n-m-1} \varphi_\theta(3^{-k}z),$$

$$P_{2,\theta}(z) = \prod_{k=n-m}^{n} \varphi_\theta(3^{-k}z).$$

The following is **Proposition 23**:

Proposition 18. *For all $\theta \in E$, $\int_I |P_{1,\theta}|^2 \, dx \geq C\, 3^m$.*

Low frequency terms do not have as much regularity, so we must control the damage caused by the **set of small values**, $SSV(\theta) := \{x \in I : |P_2(x)| \leq 3^{-\ell}\}$, $\ell = \alpha\, m$. In the next result we claim the existence of $\mathcal{E} \subset \tilde{E}$, $|\mathcal{E}| > |\tilde{E}|/2$ with the following property:

The next proposition is like **Proposition 24**, except that the following proposition holds for a larger set $SSV(\theta)$ than the corresponding set $SSV(t)$ defined there (t is a reparameterization of θ):

Proposition 19.

$$\int_{\tilde{E}} \int_{SSV(\theta)} |P_{1,\theta}(x)|^2 dx\, d\theta \leq 3^{2m-\ell/2}$$

Therefore, $\exists \mathcal{E} \subset \tilde{E}$ *such that:*

$$\forall \theta \in \mathcal{E} \ \int_{SSV(\theta)} |P_{1,\theta}(x)|^2 dx \lesssim K\, 3^{2m-\ell/2}\,.$$

Then Proposition 23 and 19 give Theorem 17; since $\ell = \alpha m$ and $K^2 = 3^m$, we see that any $\alpha > 2$ may be used for this estimate; however, we will need α to be larger soon.

4.2 Controlling $SSV(t)$

Up until now, the proof has not differed from the general case other than some choices of m, K, $|E|$ etc., for reasons soon to be established. In this section, we depart dramatically from the general case considered in Chapter 5. Remark 6, as we will see, is indispensible for the proof we consider for the gasket and unavailable in the general case. In particular, a large set of angles lacking properties like those in Remark 6 sometimes impies that $SSV(t)$ is large for a set of angles having size $\gtrsim L^{-\sqrt{m}}$, invalidating the approach we will use here, or at the very least contributing another type of case we don't yet know how to deal with. The general case is handled by much less elementary methods in Section 5.3, which must take into account the possibility of "repeated zeroes".

Remark 6. *Consider $\Phi(x,y) = 1 + e^{ix} + e^{iy}$; note that $\varphi_\theta(z) = \Phi(x_\theta(z), y_\theta(z))$. To understand the small values of Φ, the key observation is the fact that if $\Phi(x,y) = 0$ and $x, y \in \mathbb{R}$, then $\Phi(3x, 3y) = 3$, and further, that $x = \pm 2\pi/3$ mod 2π and $y = \mp 2\pi/3$ mod 2π. See also the Section 5.5.*

These lead to the following estimates:

$$|\Phi(x,y)|^2 \geq a(|4\cos^2 x - 1|^2 + |4\cos^2 y - 1|^2) \tag{4.1}$$

$$\frac{\sin 3x}{\sin x} = 4\cos^2 x - 1\,. \tag{4.2}$$

Actually, we will set $\alpha = a^{-1}$ in the end. Changing variable we can replace $3\varphi_\theta(x)$ by $\phi_t(x) = \Phi(x, tx)$.

Consider $P_{2,t}(x) := \prod_{k=n-m}^{n} \frac{1}{3}\phi_t(3^{-k}x)$, $P_{1,t}(x) := \prod_{k=0}^{n-m} \frac{1}{3}\phi_t(3^{-k}x)$.

We need control over the set $SSV(t) := \{x \in I : |P_{2,t}(x)| \leq 3^{-\ell}\}$. One can easily imagine $SSV(t)$ if one considers $\Omega := \{(x,y) \in [0,2\pi]^2 : |\mathcal{P}(x,y)| := |\prod_{k=0}^{m} \Phi(3^k x, 3^k y)| \leq 3^{m-\ell}\}$. Moreover, (using that if $x \in SSV(t)$ then $3^{-n}x \geq 3^{-m}$, and using $xdxdt = dxdy$) we change variable in the next integral:

$$\int_{\tilde{E}} \int_{SSV(t)} |P_{1,t}(x)|^2 \, dxdt = 3^{-2n+2m} \cdot 3^n \int_{\tilde{E}} \int_{3^{-n}SSV(t)} |\prod_{k=m}^{n} \Phi(3^k x, 3^k tx)|^2 \, dxdt$$

$$\leq 3^{-n+3m} \int_{\Omega} |\prod_{k=m}^{n} \Phi(3^k x, 3^k y)|^2 \, dxdy\,.$$

Now notice that by our key observations

$$\Omega \subset \{(x,y) \in [0,2\pi]^2 : |\sin 3^{m+1}x|^2 + |\sin 3^{m+1}y|^2 \leq a^{-m}3^{2m-2\ell} \leq 3^{-\ell}\}\,. \tag{4.3}$$

The latter set $\mathcal{Q}$ is the union of $4 \cdot 3^{2m+2}$ squares Q of size $3^{-m-\ell/2} \times 3^{-m-\ell/2}$. Fix such a Q and estimate

$$\int_Q |\prod_{k=m}^{n} \Phi(3^k x, 3^k y)|^2 \, dxdy \leq 3^\ell \int_Q |\prod_{k=m+\ell/2}^{n} \Phi(3^k x, 3^k y)|^2 \, dxdy$$

$$\leq 3^\ell \cdot (3^{-m-\ell/2})^2 \int_{[0,2\pi]^2} |\prod_{k=0}^{n-m-\ell/2} \Phi(3^k x, 3^k y)|^2 \, dxdy$$

$$\leq 3^\ell \cdot (3^{-m-\ell/2})^2 \cdot 3^{n-m-\ell/2} = 3^{-2m} \cdot 3^{n-m-\ell/2} .$$

Therefore, taking into account the number of squares Q in $\mathcal{Q}$ and the previous estimates we get

$$\int_E \int_{SSV(t)} |P_{1,t}(x)|^2 \, dxdt \leq 3^{2m-\ell/2} .$$

Proposition 19 is proved.

Remark 7. *It is true that α depends on the constant a in* (4.1)*, since it appears in* (4.3)*. One can use $a = \frac{1}{18}$, attained at $(x, y) = (0, \pi)$. Then from* (4.3)*, we get $\alpha = m/\ell \geq log_3(162) \approx 4.631$ as our last condition on α. We need this to compute the best exponent p.*

Note that in our argument, we cut a couple corners. To get the best exponent currently available, let $\gamma > 1$. Let $m = \gamma \epsilon_0 \log_3 N$. Then the argument works as long as $\epsilon_0 < [2\gamma\alpha + 1 - \gamma]^{-1}$, i.e., $\epsilon_0 < \frac{1}{2 \log_3(169)}$. Using the sharper exponent $\beta > 1$ in Theorem 27, one can get any $p = \frac{1}{\epsilon_0^{-1} + \beta} < \frac{1}{[2 \log_3(169)]^{-1} + 1}$ in the estimate $Fav(\mathcal{G}_n) \leq C_p n^{-p}$. In particular, $p = \frac{1}{10.262}$ is small enough.

This argument can be improved, but not so much that one should expect to get the sharp *exponent without significant, totally new ideas.*

Chapter 5

The upper bound in Buffon's needle problem - general case

See the beginning of the previous chapter for a summary of the main ideas.

5.1 The Fourier-analytic part

5.1.1 The setup

The goal of this section is to prove Theorem 21, which shows that for most directions θ, a considerable amount of stacking occurs orthogonal to θ. The constants c and C will vary from line to line, but will be absolute constants not depending on anything except perhaps L in some cases. The symbols c and C will typically denote constants that are sufficiently small or large, respectively. Everywhere we use the definition $B(z_0, \varepsilon) := \{z \in \mathbb{C} : |z - z_0| < \varepsilon\}$.

Recall Remark 1, which allows us to say

$$\mathcal{J}_1 = \bigcup_{j=1}^{L} B(r_j e^{i\theta_j}, \frac{1}{L}).$$

Also,

$$f_{n,\theta} := \sum_{\text{Discs } D \text{ of } \mathcal{J}_n}^{L} \chi_{proj_\theta(D)}.$$

Observe that $f_{n,\theta} = \nu_n * L^n \chi_{[-L^{-n}, L^{-n}]}$, where $\nu_n := *_{k=1}^{n} \widetilde{\nu}_k$ and

$$\widetilde{\nu}_k = \frac{1}{L}\left[\sum_{l=1}^{L} \delta_{L^{-k} r_l \cos(\theta - \theta_l)}\right].$$

We will now slightly modify f for convenience. Note that

$$\hat{f}_{n,\theta}(x) = L^n \hat{\chi}_{[-L^{-n}, L^{-n}]}(x) \cdot \prod_{k=1}^{n} \phi_\theta(L^{-k} x),$$

where $\phi_\theta(x) = \frac{1}{L}\sum_{l=1}^{L} e^{-ir_l \cos(\theta_l - \theta)x}$. We are interested in L^2 norms, so the argument of ϕ is of no consequence. By factoring out the first term, discarding this factor, and changing the variable, we may instead write in place of ϕ_θ the function

$$\varphi_t(x) = \frac{1}{L}\left[1 + e^{ix} + e^{itx} + \sum_{l=4}^{L} e^{a_l x + b_l t x}\right], \;\; t \in [0,1]. \tag{5.1}$$

We assumed here that $r_1 = 0$, $r_2 = r_3 = 1$, $\theta_2 = 0$, $\theta_3 = \pi/2$. We can do this by affine change of variable.

For numbers $K, N > 0$, define the following[1]:

$$f_N^*(s) := f_{N,t}^* \sup_{n \le N} f_{n,t}(s) \tag{5.2}$$

$$A_K^* := A_{K,N,t}^* := \{s : f_N^*(s) \ge K\} \tag{5.3}$$

$$E := \{t : |A_K^*| \le \frac{1}{K^3}\}\,. \tag{5.4}$$

E is essentially the set of pathological t such that $||f_{n,t}||_{L^2(s)}$ is small for all $n \le N$, as in [18]. In fact, we have this result, proved in Section 5.2.2:

Theorem 20. *Let $t \in E$. Then*

$$\max_{0 \le n \le N} \|f_{n,t}\|^2_{L^2(s)} \le c\,K\,.$$

The aim of Section 5.1 is to prove the following:

Theorem 21. *Let ϵ_0 be a fixed small enough constant. Then for $N >> 1$, $|E| < e^{-\epsilon_0 \sqrt{\log N}}$.*

So let $K \approx e^{\epsilon_0 \sqrt{\log N}}$, and suppose $|E| > \frac{1}{K}$. We will show that $N < N^*$, for some finite constant $N^* >> 1$.

5.1.2 Initial reductions

Because of Theorem 20, we have $\forall t \in E$,

$$K \ge ||f_{N,t}||^2_{L^2(s)} \approx ||\widehat{f_{N,t}}||^2_{L^2(x)} \ge C \int_1^{L^{N/2}} |\nu_{\widehat{N}}(x)|^2 dx \tag{5.5}$$

[1]Note that our result could be sharper if K^3 were replaced by K^τ, $\tau > 2$. The constant ϵ_0 could be computed explicitly, and it depends on τ. We will not do this, though.

Let $m \approx (\frac{\epsilon_0}{2}\log N)^{1/2}$. Split $[1, L^{N/2}]$ into $N/2$ pieces $[L^k, L^{k+1}]$ and take a sample integral of $|\widehat{\nu_N}|^2$ on a small block $I := [L^{n-m}, L^n]$, with $n \in [N/4, N/2]$ chosen so that

$$\frac{1}{|E|}\int_E \int_{L^{n-m}}^{L^n} |\widehat{\nu_N}(x)|^2 dx\, dt \leq CKm/N\,.$$

This choice is possible by (5.5). Define

$$\tilde{E} := \{t \in E : \int_{L^{n-m}}^{L^n} |\widehat{\nu_N}(x)|^2 dx \leq 2CKm/N\}\,.$$

It then follows that $|\tilde{E}| \geq \frac{1}{2K}$.

Note that $\widehat{\nu_N}(x) = \prod_{k=1}^{N} \varphi(L^{-k}x) \approx \prod_{k=1}^{n} \varphi(L^{-k}x)$ for $x \in [L^{n-m}, L^n]$.

So for $t \in E$,

$$\int_{L^{n-m}}^{L^n} \prod_{k=1}^{n} |\varphi_t(L^{-k}x)|^2 dx \leq \frac{CKm}{N} \leq 2\epsilon_0 N^{\epsilon_0 - 1}\log N.$$

Recall that $m \approx (\frac{\epsilon_0}{2}\log N)^{1/2}$. Later, we will show that $\exists t \in E$ and absolute constant α such that

$$\int_{L^{n-m}}^{L^n} \prod_{k=1}^{n} |\varphi_t(L^{-k}x)|^2 dx \geq cL^{m-2\cdot\alpha m^2} \geq cN^{-\alpha\epsilon_0}. \tag{5.6}$$

The result: $2\epsilon_0 \log N \geq N^{1-4\alpha\epsilon_0-\epsilon_0}$, i.e., $N \leq N^*$ if ϵ_0 is small enough. In other words:

Proposition 22. *Inequality* (5.6) *is sufficient to prove Theorem* (21)*. Further, inequality* 5.6 *can be deduced from Propositions* 23 *and* 24*, as will be seen shortly.*

So let us prove inequality (5.6).

First, let us write $\prod_{k=1}^{n}\varphi_t(L^{-k}x) = P_t(x) = P_{1,t}(x)P_{2,t}(y)$, where P_2 is the low frequency part, and P_1 is has medium and high frequencies:

$$P_{1,t}(x) := \prod_{k=1}^{n-m}\varphi_t(L^{-k}x) = \widehat{\nu_{n-m}}(x)$$

$$P_{2,t}(x) = \prod_{k=n-m}^{n}\varphi_t(L^{-k}x) = \widehat{\nu_m}(L^{m-n}x)$$

We want the following:

Proposition 23. *Let $t \in E$ be fixed. Then $\int_{L^{n-m}}^{L^n}|P_{1,t}(x)|^2 dx \geq C\,L^m$.*

Recall that we defined the set $\tilde{E}, |\tilde{E}| > |E|/2$, and we assume that

$$|E| > 1/K\,. \tag{5.7}$$

Recall that we denoted

$$I = [L^{n-m}, L^n]\,.$$

We also want a proportion of the contribution to the integral separated away from the complex zeroes of $P_{2,t}$:

Proposition 24. *Let $SSV(t) := \{x \in I : |P_{2,t}(x)| \leq L^{-\alpha m^2}\}$. Suppose also that E is unable to hide, that is (5.7) is valid. Then there exists a subset $\mathcal{E} \subset \tilde{E}$, $|\mathcal{E}| \geq 1/4K$, such that for every $\theta \in \mathcal{E}$ one has*

$$\int_{SSV(t)}|P_{1,t}(x)|^2 dxdt \leq 2c\,L^m\,,$$

where $2c$ *is less than the* C *from Proposition* 23*. In particular,*

$$\frac{1}{|\tilde{E}|}\int_{\tilde{E}}\int_{SSV(t)}|P_{1,t}(x)|^2dxdt \leq c\,L^m,$$

Remark 8. *The set* $SSV(t)$ *is so named because it is the* ***set of small values*** *of* P_2 *on* I*. Combining this with Proposition* 23*,*

$$\int_{L^{n-m}}^{L^n}|P_{1,t}(x)|^2|P_{2,t}(x)|^2\,dx \geq \int_{I\setminus SSV(t)}|P_{1,t}(x)|^2\cdot L^{-\alpha m^2}\,dx \geq c\,L^{m-2\alpha m^2},$$

which gives (5.6)*–exactly what we promised to obtain from Propositions 24, 23. Thus Propositions* 23 *and* 24 *suffice to prove Theorem* 21*, and Proposition* 22 *has been demonstrated.*

Remark 9. *We want to show that for* $N >> 1$, (5.7) *fails. After showing this, we will have:*

$$|E| \leq 1/K = L^{-\frac{m}{2}} = e^{-C(L)\epsilon_0(\log N)^{1/2}}, \tag{5.8}$$

proving the main result, since the projections decay quickly enough on E^c*.*

First, let us fix $t \in E$ and prove Proposition 23.

Proof. We are using first Salem's trick on

$$\int_0^{L^n}|P_1(x)|^2dx:$$

Let $h(x) := (1-|x|)\chi_{[-1,1]}(x)$, and note that $\hat{h}(\alpha) = C\frac{1-\cos\alpha}{\alpha^2} > 0$. Then if we write

$P_1 = L^{m-n-1} \sum_{j=0}^{L^{n-m}} e^{i\alpha_j x}$, we get

$$\int_0^{L^n} |P_1(x)|^2 dx \geq 2 \int_{-L^n}^{L^n} h(L^{-n}x) |P_1(x)|^2 dx$$

$$\geq C(L^{m-n})^2 [L^n \cdot L^{n-m} + \sum_{j \neq k; j,k=1}^{L^{n-m}} L^n \hat{h}(L^n(\alpha_j - \alpha_k))] \geq CL^m.$$

To show that this is not concentrated on $[0, L^{n-m}]$, we will use Theorem 20 and Lemma 40. We get

$$\int_0^{L^{n-m}} |P_1(x)|^2 dx = \int_0^{L^{n-m}} |\widehat{\nu_{n-m}}(x)|^2 dx = L^{2(m-n)} \int_0^{L^{n-m}} |\sum_{j=0}^{n-m} e^{i\alpha_j x}|^2 dx$$

$$\leq CK \leq CL^{\frac{m}{2}}. \tag{5.9}$$

□

So now we have Proposition 23. The greater challenge will be Proposition 24.

5.1.3 The proof of Proposition 24

Recall that $SSV(t) := \{x \in I = [L^{n-m}, L^n] : |P_{2,t}(x)| \leq L^{-\alpha m^2}\}$.

To get Proposition 24, we will split $P_{1,t}$ into two parts, $P^\sharp_{1,t}(x)$ and $P^\flat_{1,t}(x)$ corresponding to medium and high frequencies.

A straightforward application of Lemma 40 to high frequency part $P^\sharp_{1,t}(x)$ will get us part of the way there (see Proposition 26), and the claim 25 applied to medium frequency term $P^\flat_{1,t}(x)$ will further sharpen the final estimate to what we need. This latter refinement will be a "for most t..." statement about $P^\flat_{1,t}(x)$ that contributes a small amount to the possible size of E.

Naturally, $P^{\flat}_{1,t}(x)$ and $P^{\sharp}_{1,t}(x)$ are defined as the medium and high frequency parts of $P_{1,t}(x)$. Below, $\ell := \alpha m$:

$$P^{\flat}_{1,t}(x) := \prod_{k=n-m-\ell}^{n-m-1} \varphi_t(L^{-k}x) = \hat{\nu}_{\ell-1}(L^{m+\ell-n}x)\,,$$

$$P^{\sharp}_{1,t}(x) := \prod_{k=1}^{n-m-\ell-1} \varphi_t(L^{-k}x) = \hat{\nu}_{n-m-\ell-1}(x).$$

What follows is the first claim of this subsection. The idea is simply that $|\Phi_t| \leq 1$, with equality only when the exponents all belong to $2\pi\mathbb{Z}$. As it is quite difficult for this to happen simultaneously for even just two exponential terms, one gains a lot of information about the decimal expansion of t whenever $|\Phi_t|$ is close to 1.

Proposition 25. *For all sufficiently small positive numbers $\tau \leq \tau_0$ and for all sufficiently large m and $\ell = \alpha\, m$ there exists an exceptional set H of directions t such that*

$$|H| \leq L^{-\ell/2}\,, \tag{5.10}$$

$$\forall t \notin H\, \forall x \in [L^{n-m}, L^n],\ |P^{\flat}_{1,t}(x)| \leq e^{-\tau\,\ell}\,. \tag{5.11}$$

Proof. Notice that

$$\phi_\theta(r) = \Phi(r\cos\theta, r\sin\theta)\,,$$

where for $x = (x_1, x_2)$,

$$\Phi(x) := \Phi(x_1, x_2) = \frac{1}{L}\sum_{l=1}^{L} e^{2\pi i\langle a_l, x\rangle}\,.$$

As some pair of vectors $a_l - a_1$, $l \in [1, L]$ must span a two-dimensional space, we can assume without the loss of generality (make an affine change of variable) that

$$a_1 = (0,0)\,, a_2 = (1,0)\,, a_3 = (0,1)\,.$$

Then

$$\Phi(x_1, x_2) = \frac{1}{L}(1 + e^{2\pi i x_1} + e^{2\pi i x_2} + \sum_{l=4}^{L} e^{2\pi i \langle a_l, x\rangle})\,. \tag{5.12}$$

We make the change of variable $y = (y_1, y_2) = L^{-(n-m)}x$. Let R_t denote the ray $y_2 = t y_1$. Then we need to prove that there exists a small set H of t's such that if $y \in R_t \cap \{y : |y| \in [1, L^m]\}$, $t \notin H$ then

$$|\Phi(y) \cdot \dots \cdot \Phi(L^\ell y)| \le e^{-\tau \ell}\,. \tag{5.13}$$

We consider only the case $t \in [0,1]$, all our y's will be such that $0 < y_2 \le y_1$, and as $|y| \ge 1$ we have $y_1 \ge \frac{1}{\sqrt{2}}$.

It is very difficult if at all possible for function Φ to satisfy $|\Phi(y)| = 1$. In fact, looking at (5.12) we can see that

$$|\Phi(y)| \le 1 - b\mathrm{dist}(y, \mathbb{Z}^2) \le e^{-b\mathrm{dist}(y,\mathbb{Z}^2)}\,. \tag{5.14}$$

Therefore, we are left to understand that there are few t's such that

$$\exists y \in R_t, : y_1 \in [\frac{1}{\sqrt{2}}, L^m] : \; b \cdot \sum_{k=0}^{\ell} \mathrm{dist}(L^k\, y, \mathbb{Z}^2) \le \tau\, \ell\,. \tag{5.15}$$

Now may be a good time to consult Figure 5.1.

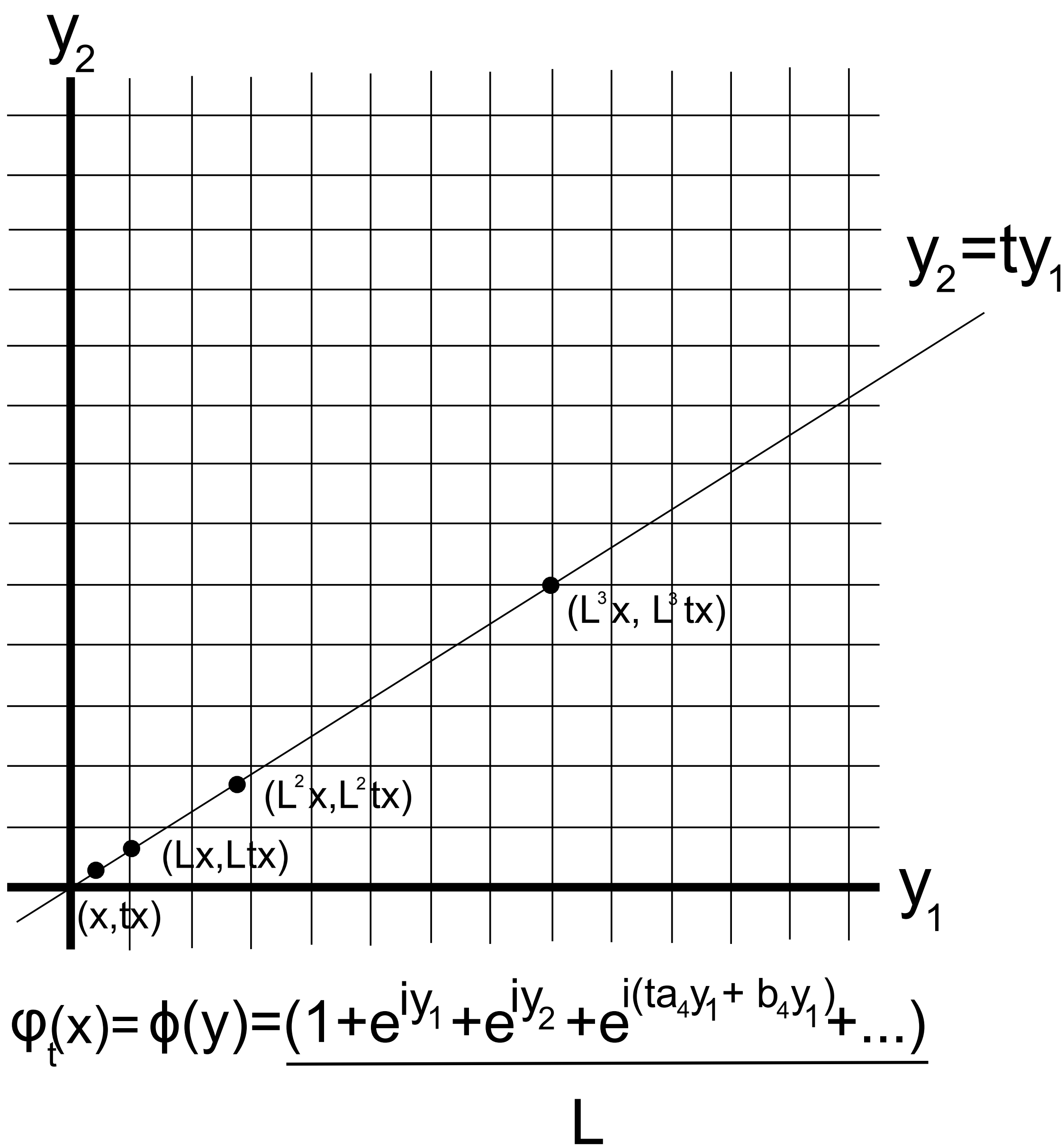

Figure 5.1: It is quite difficult for a large number of factors of $P^{\flat}_{1,t}(x)$ to be close to 1 simultaneously. In particular, $L^k x$ and $L^k tx$ must be close to $\mathbb{Z}$ for many values of k.

Fix $y \in R_t$ as above. If (5.15) holds then for 90 per cent of $k's$ one has

$$\mathrm{dist}(L^k y, \mathbb{Z}^2) \leq 10\tau \ell\,. \tag{5.16}$$

Denote $Z_y := \{k \in [0, \ell] : \mathrm{dist}(L^k y, \mathbb{Z}^2) \leq 10\tau \ell\}$. We know that

$$|Z_y| \geq 0.9\ell\,.$$

Let us call **scenario** the collection $s := \{m_1; k_1, ..., k_{0.1\ell}\}$, where $m_1 = 0, .., m; 0 \leq k_1 < ... < k_{0.1\ell}$.

Every t such that there exists y such that (5.15) holds generates several scenarios according to

$$y_1 \in [L^{m_1 - 1}, L^{m_1})$$

and according to what is the set $[0, \ell] \setminus Z_y$—this is the set $k_1, ..., k_{0.1\ell}$ of the scenario.

We will calculate the number of scenarios later. Now let us fix a scenario $s = \{m_1; k_1, ..., k_{0.1\ell}\}$, and let us estimate the measure of the set $T(s)$, $T(s) := \{t \in (0, 1) : \exists y, y_2 = ty_1, y_1 \in [L^{m_1 - 1}, L^{m_1})$ such that $[0, \ell] \setminus Z_y = \{k_1, \ldots, k_{0.1\ell}\}$. To do that for this fixed scenario we fix a **net**. To explain what is a net we fix

$$a := \left[\frac{\log \frac{100}{\eta}}{\log L}\right] + 1\,,$$

where $\eta = C\,\tau$ and C is an absolute constant to be chosen soon.

A net is a collection $N(s) := \{n_1, \ldots, n_j\}$, $n_1 < n_2 < \ldots$, where every n_i is not among

k_j included in the scenario, $j \geq \frac{3\ell}{4a} + 1$, and

$$n_{i+1} - n_i \geq 2a\,.$$

Given a scenario it is always possible to built a net. In fact we just delete from $[0, \ell]$ the numbers $k_1, ..., k_{0.1\ell}$ belonging to the scenario, we are left with at least 0.9ℓ numbers. We choose an arithmetic progression with step a (enumerating them anew first). This arithmetic progression will be long enough, its length $j \geq \frac{3\ell}{4a}$ because after eliminating $k_1, ..., k_{0.1\ell}$ we still have at least 0.9ℓ numbers left. We mark the numbers of this progression. Then we put back $k_1, ..., k_{0.1\ell}$. The marked numbers will form our net.

If $t \in T(s)$ then there exists $y = (y_1, ty_1)$ as above, in particular,

$$\mathrm{dist}(L^{n_i}\, y, \mathbb{Z}^2) \leq 10\tau\, \ell\,,\ \forall n_i \in N(s)\,.$$

Let us write that then there exist integers $p_1 \leq q_1$: $|L^{n_1} y_1 - q_1| < 10\tau$, $|L^{n_1} y_2 - p_1| < 10\tau$, so

$$\left| t - \frac{p_1}{q_1} \right| = \left| \frac{L^{n_1} y_2}{L^{n_1} y_1} - \frac{p_1}{q_1} \right| = \left| \frac{L^{n_1} y_2 - p_1 + p_1}{L^{n_1} y_1 - q_1 + q_1} - \frac{p_1}{q_1} \right|$$

$$\left| \frac{(L^{n_1} y_2 - p_1 + p_1)q_1 - (L^{n_1} y_1 - q_1 + q_1)p_1}{(L^{n_1} y_1 - q_1 + q_1)q_1} \right| \leq \frac{|L^{n_1} y_2 - p_1||q_1| + |L^{n_1} y_1 - q_1||p_1|}{(q_1 - 10\tau)q_1}$$

$$\leq 40\tau \frac{1}{q_1}\,.$$

As promised we choose C: $C = 40$, $\eta := 40\tau$ and we get

$$\exists p_1 \leq q_1 \ :\ \left| t - \frac{p_1}{q_1} \right| \leq \eta \frac{1}{q_1}\,. \tag{5.17}$$

Next we choose integers $p_2 \le q_2$: $|L^{n_2} y_1 - q_2| < 10\tau$, $|L^{n_2} y_2 - p_2| < 10\tau$ and obtain

$$\exists p_2 \le q_2 \;:\; \left| t - \frac{p_2}{q_2} \right| \le \eta \frac{1}{q_2}\,. \tag{5.18}$$

Notice also that because of $|L^{n_1} y_1 - q_1| < 10\eta$, $|L^{n_2} y_1 - q_2| < 10\eta$, $y_1 \ge 1/\sqrt{2}$, and smallness of τ, and the fact that $n_2 - n_1 \ge 2a$, we get

$$\frac{q_2}{q_1} \ge L^a \ge \frac{100}{\eta}\,. \tag{5.19}$$

We continue in the same vein, $i = 2, \dots, j-1 \ge \frac{3\ell}{4a}$:

$$\exists p_i \le q_i \;:\; \left| t - \frac{p_i}{q_i} \right| \le \eta \frac{1}{q_i}\,. \tag{5.20}$$

Notice also that because of $|L^{n_1} y_1 - q_1| < 10\eta$, $|L^{n_2} y_1 - q_2| < 10\eta$, $y_1 \ge 1/\sqrt{2}$, and smallness of τ, and the fact that $n_2 - n_1 \ge 2a$, we get

$$\frac{q_{i+1}}{q_i} \ge L^a \ge \frac{100}{\eta}\,. \tag{5.21}$$

Inequality (5.17) gives that $|T(s)| \le \eta$, inequalities (5.17) and (5.18) in conjunction with (5.19) give $|T(s)| \le \left(1 + \frac{1}{100}\right)\eta^2$, similarly all inequalities (5.20), (5.21) together give

$$|T(s)| \le (1.01\eta)^{\frac{3\ell}{4a}} \ge e^{0.1\,\ell} L^{-\frac{3}{4}\ell(1-\epsilon(\eta))}\,.$$

Here we used of course that $a := \left[\frac{\log \frac{100}{\eta}}{\log L}\right] + 1$. Finally, if η is sufficiently small we have

$$|T(s)| \leq L^{-\frac{2}{3}\ell}. \tag{5.22}$$

Let $\mathcal{S}$ denote the set of all scenarios. Now we want to calculate the number of scenarios. This is easy:

$$\#\mathcal{S} \leq m \cdot \binom{\ell}{0.1\ell} \leq \ell \cdot \left(\frac{10}{9}\right)^{0.9\ell} \cdot 10^{0.1\ell}.$$

We just proved that the measure of the set of all $t \in (0,1)$ such that one has (5.15)

$$\exists y \in R_t, : y_1 \in [\frac{1}{\sqrt{2}}, L^m] \; : \; \sum_{k=0}^{\ell} \mathrm{dist}(L^k y, \mathbb{Z}^2) \leq \tau\,\ell$$

can be estimated as

$$\leq \ell \cdot \left(\frac{10}{9}\right)^{0.9\ell} \cdot 10^{0.1\ell} \cdot L^{-\frac{2}{3}\ell} \leq L^{-\ell/2}.$$

Proposition 25 is proved. Except for a small set of exceptional directions, the uniform bound $|P^{\flat}_{1,t}(x)| < e^{-\tau\ell}$ holds.

□

Here is the second claim of the subsection:

Proposition 26.

$$t \in E \Rightarrow \int_{SSV(t)} |P^{\sharp}_{1,t}(x)|^2 dx \leq C'' K\, L^m.$$

We will see in Section 5.3 that for each t, $SSV(t)$ is contained in $C \cdot L^m$ neighborhoods of size $L^{n-m-\ell}$ around the complex zeroes λ_j of P_2.

Fix t. Let

$$I_j = [\lambda_j - L^{n-m-\ell}, \lambda_j + L^{n-m-\ell}], \tag{5.23}$$

$$\text{where } SSV(t) \subseteq \bigcup_j I_j \tag{5.24}$$

Choose j for which $\int_{I_j} |P^{\sharp}_{1,t}(x)|^2 dx$ is maximized. Then

$$\int_{SSV(t)} |P^{\sharp}_{1,t}(x)|^2 dx \leq CL^m \int_{I_j} |P^{\sharp}_{1,t}(x)|^2 dx \leq CL^m (L^{\ell+m-n})^2 \int_{I_j} |\sum_{k=0}^{n-m-\ell} e^{i\alpha_j x}|^2.$$

As $|I_j| \leq 2 \cdot L^{n-m-\ell}$, so Lemma 40 and the definition of E give us Proposition 26.

The estimate for $t \in \tilde{E} \setminus H$ follows. If $|E| \geq 1/K, K = L^{m/2}$, $|\tilde{E}| > 1/2K$, and we also just proved that $|H| \leq L^{-\ell/2}$, $\ell = \alpha\, m$ with large α, we have a set $\mathcal{E} \subset \tilde{E} \setminus H$, $\mathcal{E} > 1/4K$, such that for every $t \in \mathcal{E}$

$$\int_{SSV(t)} |P_1(r)|^2 \, dr \leq L^{-\ell} \int_{SSV(t)} |P^{\sharp}_{1,t}(x)(r)|^2 \, dr \leq C'' K \, L^m \cdot L^{-\alpha m} \, .$$

So we proved

$$\int_{SSV(t)} |P_1(r)|^2 \, dr \leq c \, L^m \tag{5.25}$$

with c as small as we wish. In particular, Proposition 24 is completely proved.

5.2 Two combinatorial lemmas

In this section, we will prove two combinatorial lemmas. The objective in each case is to rigorously estimate one quantity by another, clearly related, quantity. The two, taken together, reduce the problem of finding an upper bound in Buffon's needle problem to the problem of finding a bound on $|E|$.

For this section, regard the set E from Section 5.1 as parameterized by θ, and use the variable x instead of s on the non-Fourier side, since we will not work on the Fourier side at all during this section.

5.2.1 $|A^*_{K,N,\theta}|$ vs. $|\mathcal{L}_{NK^\beta,\theta}|$

In this section, we show how Theorem 13 follows from Theorem 21. The theorem we prove here is the big brother of the **micro-theorem converse**.

First, let us define

$$\mathcal{L}_{N,\theta} := proj_\theta \mathcal{J}_N. \tag{5.26}$$

Theorem 27. *Let $\beta > 1$ (we used $\beta = 3$ in the previous section). Let K and N be large enough, possibly depending on L. If $t \notin E$ (see definition (5.4) and use $\tau > 2$ as suggested), then $|\mathcal{L}_{NK^\beta,\theta}| \le \frac{C}{K}$.*

Proof. Let us use θ instead of t and use x for the space variable on the non-Fourier side, since we do not use Fourier analysis in this proof. Fix θ, and for $j \in \mathbb{N}$, let $F_j := A^*_{K,jN,\theta} = \{x : f^*_{jN}(x) \ge K\}$. Let $F := F_1$. $\theta \notin E$ means $|F| \ge K^{-\tau}$, where $\tau > 2$ is fixed.

Note that this theorem is the sophisticated analog of the **micro-theorem converse** of Chapter 1.

Consider the discs of $\mathcal{J}_N$. All discs are white initially. Now each disc lying above any $x \in F$ green. We will now consider the sets $\mathcal{J}_{jN}$, for $j = 1, 2, ...$ and label these discs as green or white according to these rules:

1) If a disc in $\mathcal{J}_{jN}$ is green, its offspring in $\mathcal{J}_{(j+1)N}$ are all green. 2) If a disc in $\mathcal{J}_{jN}$ is

white, its offspring in $\mathcal{J}_{(j+1)N}$ are white except for those discs which are self-similar copies of the discs which were green in $\mathcal{J}_N$.

Let G_j denote the set of green discs in $\mathcal{J}_{jN}$. Note that $\theta \notin E$ tells us that $|G_1|$ is fairly large - let us prove a statement to this effect. Consider $\phi_j(x) := \sum_{D \in G_j} \chi_{proj_\theta(D)}$, and let $\phi(x) := \phi_1(x)$.

Proposition 28.

$$\bigcup_{D \in G_j} D \subseteq \{x : \mathcal{M}\phi_j > K/4\}$$

Proof. When a disc D in G_j with projected center at x_0 has white ancestor in G_{j-1} – that is, it is "green for the first time" – it is clear that $\mathcal{M}\phi_j > K/4$ by taking the average of ϕ_j on $[x_0 - 2L^{-jN}, x_0 + 2L^{-jN}]$. In fact, the L^1 mass of the green discs above such an interval cannot decrease below this bound, simply because the offspring have L^1 mass summing to that of its parent, and the interval contains all of these K discs entirely. □

Proposition 29. *$F \subseteq \{x : \mathcal{M}\phi \geq K/2\}$, where $\mathcal{M}$ is the (uncentered or centered; we will take it to be centered) Hardy-Littlewood maximal operator.*

Proof. Fix $x \in F$. By definition, $\exists n < N$ such that $f_n(x) \geq K$. Thus the interval $[x - 2L^{-n}, x + 2L^{-n}]$ contains the projections of K green discs of $\mathcal{J}_n$, i.e., $\phi(x) \geq K$. In fact, the total L^1 mass of the sum of characteristic functions of the children of these projected green discs remains constant as n increases. So clearly

$$\mathcal{M}\phi(x) \geq \frac{L^n}{4} \int_{x-2L^{-n}}^{x+2L^{-n}} f_{N,\theta}(x)dx \geq K2L^{-n}\frac{L^n}{4} \geq K/2.$$

□

Of course one sees where this is headed:

$$|F| \leq |\{x : \mathcal{M}\phi(x) > K/2\}| \lesssim \frac{1}{K}||\phi||_1 = \frac{2}{K}L^{-N}|G_1|. \tag{5.27}$$

Since $\theta \notin E$, this immediately proves:

Proposition 30.

$$|G_1| \gtrsim K^{1-\tau}L^N.$$

□

Let P_j denote $|G_j| \cdot L^{-jN}$, that is, the proportion of discs of $\mathcal{J}_{jN}$ which are green. Note that $Q_j := 1 - P_j = \left(1 - \frac{|G_1|}{L^N}\right)^j \leq \left(1 - cK^{1-\tau}\right)^j = \left(1 - \frac{cjK^{1-\tau}}{j}\right)^j \approx e^{-cjK^{1-\tau}}$. Note that

$$\left|\bigcup_{W \text{ a white disc of } \mathcal{J}_{jN}} proj_\theta(W)\right| \leq 2Q_j \lesssim e^{-cjK^{1-\tau}}.$$

Also, we saw already that the remaining discs of D of $\mathcal{J}_{jN}$ are exactly the green discs, i.e., $D \in G_j$. Using Proposition 28, we see that

$$\left|\bigcup_{D \in G_j} proj_\theta(D)\right| \leq |\{x : \mathcal{M}\phi_j(x) > K/2\}| \lesssim \frac{1}{K}||\phi_j||_1 \leq \frac{2}{K}.$$

In particular, if $j > K^{\tau-1+\varepsilon} = K^\beta$, there are few enough white discs, and all is well. This completes the proof of Theorem 27.

□

5.2.2 $\sup_{n\le N} ||f_{n,\theta}||_2^2$ vs. $|A^*_{K,N,\theta}|$

Theorem 31. *Let $\theta \in E$. Then*

$$\max_{n:0\le n\le N} \|f_{n,\theta}\|^2_{L^2(\mathbb{R})} \le C\,K\,.$$

To prove this we first need the following claim, which is the main combinatorial assertion of this subsection. It repeats the one in [18] but we give a slightly different proof.

We fix a direction θ, we think that the line ℓ_θ on which we project is $\mathbb{R}$. If $x \in \mathbb{R}$ then by N_x we denote the line orthogonal to $\mathbb{R}$ and passing through point x, we call N_x a needle.

Recall that $A^*_{K,N,\theta} := \{x \in \mathbb{R} : f^*_{N,\theta}(x) > K\}$. When N and θ are understood from context, we can write $F_K := A^*_{K,N,\theta}$.

Theorem 32. *There exists an absolute constant C such that for any large enough K, M, and N,*

$$|F_{2LKM}| \le CLK\,|F_K| \cdot |F_M|\,. \tag{5.28}$$

Proof. One can see this by considering maximal discs above F_{2LK}. Suppose $x \in F_{2LK}$. Then there are at least $2LK$ "light green" (relative to x and n) discs of some generation $n \le N$ above x; call these $L_{x,n}$. In generation $n-1$, there are still at least $2K$ discs above x - namely, the fathers of the $2LK$ discs of generation n. Keep going back one generation until you reach $j_0 = j_0(x)$, the largest $j < n$ such that the generation j ancestors of the light green discs of $L_{x,n}$ are fewer than $2LK$ in number. Call these discs of generation $j_0(x)$ the green discs (relative to x and n), or $G_{x,n}$. Then $|G_{x,n}| \ge 2K$. Form the union $G = \cup_{x,n} G_{x,n}$ of green discs. The discs of this union are just called green.

Each green disc is maximal for some (x,n), but it may be the case that a green disc above (x_1, n_1) is properly contained in a green disc above (x_2, n_2). We want our maximal discs to be truly maximal, so mark as dark green all green discs which are not sub-discs of a larger green disc. Call the family of dark green discs D.

The largest dark green disc has some radius L^{-n_0}. Call one such dark green disc Q_0. $Q_0 \in G_{x,n}$ for some (x, n), so it belongs to a stack of K or more green discs. In fact, they are all dark green by the maximality of Q_0.

Let $I_0 = 20proj_\theta(Q_0)$, where the rescaling is concentric. Consider all $Q \in D$ whose projection intersects I_0. Call this set of such Q by the name $\mathcal{F}(Q_0)$. For all $x \in \mathbb{R}$, the needle at x intersects fewer than $2LK$ discs from the set $\mathcal{F}(Q_0)$ (Otherwise, larger green discs could be found by taking ancestors, contradiction). Since $\mathcal{F}(Q_0)$ lives above $I_0 + [-2L^{-n_0}, 2L^{n_0}]$, $|\mathcal{F}(Q_0)| < 100LK$.

Let x_0 be the projected center of Q_0. Let $J_0 := [x_0, x_0 + L^{-n_0}]$ or $J_0 := [x_0 - L^{-n_0}, x_0]$, whichever contains at least K projected centers of dark green discs. Thus $J_0 \subseteq F_K$ and $|J_0| \geq L^{-n_0} \gtrsim |I_0|$.

Lemma 33. $|F_{2LKM} \cap I_0| \lesssim KL|I_0||F_M| \lesssim KL|J_0||F_M|$

Proof. Let $x \in F_{2LKM} \cap I_0$. Note that $F_{2LKM} \cap I_0 \subseteq F_{2LK} \cap I_0 \subseteq \mathcal{F}(Q_0)$. So in generation n_0, x has fewer than $2LK$ discs above it, whose projected lengths sum to at most $cKL|I_0|$. For some $n \leq N$, the stack must reach height $2LKM$, which means that one of the discs of $\mathcal{F}(Q_0)$ must give birth to a stack of M discs. That is, x must belong to one of $\leq 2KL$ self-similar copies of F_M living inside of $\mathcal{F}(Q_0)$. The lemma follows. □

To finish the proof, one needs to induct. That is, one needs intervals $I_1, I_2, ...$ covering F_{2LKM} such that comparable subintervals $J_1, J_2, ...$ can be substituted for I_0 and J_0 in the statement of this last lemma. This, in fact, can be done; one deletes $\bigcup_{r=1}^{s} I_r$ from

F_{2KL} and starts the maximality argument over again to get I_{s+1} and J_{s+1}. Note that by maximality, it is impossible for the sets I_r to overlap too much; each is centered outside of the previous, and they only shrink. The problem is finite, so in fact all of F_{2LKM} is exhausted in this way.

This completes the proof of Theorem 32.

□

Now we can prove Theorem 31.

Proof. Let $E_j := \{x : f_{n,\theta}(x) > (2LK)^{j+1}\}$, $j = 0, 1, ...$ We know by Theorem 32 that

$$|E_j| \leq (CLK)^j |E_0|^{j+1} .$$

Hence,

$$\int f_{n,\theta}(x)^2 \, dx \leq 2LK \int f_{n,\theta}(x) \, dx + \sum_{j=0}^{\infty} \int_{E_j \setminus E_{j+1}} f_{n,\theta}(x)^2 \, dx$$

$$\leq 2LK \int f_{n,\theta}(x) \, dx + \sum_{j=0}^{\infty} (2LK)^{j+2} \int_{E_j \setminus E_{j+1}} f_{n,\theta}(x) \, dx$$

$$\leq 2CLK + \sum_{j=0}^{\infty} (2LK)^{j+2} \, (CLK)^j |E_0|^{j+1} .$$

If $|\{x : f_N^*(x) > K\}| \leq 1/K^\tau$, $\tau > 2$, then for all $n \leq N$ we can immediately read the previous inequality as

$$\int f_{n,\theta}(x)^2 \, dx \leq C(\tau) \, K .$$

□

5.3 Controlling $SSV(t)$

Now we have to consider $P_{2,t}(r) = \phi_t(r)\phi_t(L^{-1}r) \cdot \dots \cdot \phi_t(L^{-m}r)$. We are interested in the set

$$SSV(t) := \{r \in [1, L^m] : |P_{2,t}(r)| \le L^{-Am^2}\} .$$

We will be using so-called Turan's lemma:

Lemma 34. *Let $f(x) = \sum_{l=1}^{L} c_l e^{\lambda_l x}$, let $E \subset I$, I being any interval. Then*

$$\sup_I |f(x)| \le e^{\max |\Re\lambda_n|\,|I|} \left(\frac{A|I|}{|E|}\right)^L \sup_E |f(x)| .$$

Here A is an absolute constant.

In this form it is proved by F. Nazarov [21].

Now let us consider any square $Q = [x' - 1, x' + 1] \times [-1, 1]$. We call $\frac{1}{2}Q$ the concentric square of half the size.

Lemma 35. *With uniform constant C depending only on L one has*

$$\sup_Q |\phi_t(z)| \le C \sup_{\frac{1}{2}Q} |\phi_t(z)| .$$

Proof. Let $z_0 = x_0 + iy_0$ is a point of maximum in the closure of Q. We first want to compare $|f(z_0)|$ and $|f(x_0)|$. Consider $f_{x_0}(y) := \phi_t(x_0 + iy)$. Notice that uniformly in Q and x_0

$$|f'_{x_0}(y)| \le C(L) .$$

This means that $|f_{x_0}(y)| \ge \frac{1}{2}|f_{x_0}(0)|$ on an interval of uniform length $c(L)$.

Notice also that the exponents $\lambda_l(t), l = 1, \dots, L$, encountered in ϕ_t are all uniformly bounded. Then applying Lemma 34 we get

$$|\phi_t(z_0)| = |f x_0(y_0)| \leq C'(L)|f x_0(0)|\,.$$

Now consider $F(x) = \phi_t(x)$. We want to compare $F(x_0) = f x_0(0) = \phi_t(x_0)$ with

$$\max_{[x'-\frac{1}{2}, x'+\frac{1}{2}]} |F(x)|.$$

By Lemma 34 we get again

$$|f x_0(0)| = |F(x_0)| \leq \sup_{[x'-1, x'+1]} |F(x)| \leq C''(L) \sup_{[x'-1/2, x'+1/2]} |F(x)| \leq C''(L) \sup_{\frac{1}{2}Q} |\phi_t(z)|$$

Combining the last two display inequalities we get Lemma 35 completely proved.

□

Lemma 36. *With uniform constant C depending only on L (and not on m) one has*

$$\sup_{Q} |\phi_t(L^{-k} z)| \leq C \sup_{\frac{1}{2}Q} |\phi_t(L^{-k} z)|\,,\ k = 0, \dots, m\,.$$

The proof is exactly the same. We just use $L^{-k}\lambda_l(t), l = 1, \dots, L$, encountered in $\phi_t(L^{-k}\cdot)$ are all uniformly bounded.

By complex analysis lemmas from Section 5.4 we know that Lemma 36 implies that every $\frac{1}{2}Q$ has at most M (depending only on L) zeros of $\phi_t(z)$. And if we denote them by

$\mu_1, \ldots, \mu_M$ then

$$\{x \in \frac{1}{2}Q \cap \mathbb{R} : |\phi_t(x)| \le L^{-M\ell}\} \subseteq \bigcup_{i=1}^{M} B(\mu_i, L^{-\ell})\,. \tag{5.29}$$

Consider $\mu_1, \ldots, \mu_S$ being all zeros of $P_{2,t}$ in $[1/2, L^m + 1] \times [1/2, 1/2]$. By abovementioned lemmas from Section 5.4 and by Lemma 36 we get that

$$S \le M(L)\, L^m\,.$$

From (5.29) it is immediate that

$$\{x \in [1, L^m] :\, |P_{2,t}(L^{-(n-m)}x)| \le L^{-M\ell m}\} \subset \bigcup_{i=1}^{M\,L^m} B(\mu_i, L^{-\ell})\,. \tag{5.30}$$

Changing the variable $y = L^{n-m}x$ we get the structure of the set of small values used above during the proof of Proposition 26:

$$SSV(t) \subset \cup_{i=1}^{C\,L^m} I_i\,, \tag{5.31}$$

where each interval I_i has the length $2\,L^{n-m-\ell}$.

In this section, we also include Lemmas 37 and 38. Given a bounded holomorphic function on the disc, its supremum, and an interior non-zero value, these lemmas bound the number of zeroes and contain the set of small values within certain neighborhoods of these zeroes. They are somewhat standard, but are included for completeness.

5.3.1 A Blaschke estimate

Lemma 37. *Let D be the closed unit disc in $\mathbb{C}$. Suppose ϕ is holomorphic in an open neighborhood of D, $|\phi(0)| \geq 1$, and the zeroes of ϕ in $\frac{1}{2}D$ are given by $\lambda_1, \lambda_2, ..., \lambda_M$. Let $C = ||\phi||_{L^\infty(D)}$. Then $M \leq log_2(C)$.*

Proof. Let

$$B(z) = \prod_{k=1}^{M} \frac{z - \lambda_k}{1 - \bar{\lambda}_k z}.$$

Then $|B| \leq 1$ on D, with $=$ on the boundary. If we let $g := \frac{\phi}{B}$, then g is holomorphic and nonzero on $\frac{1}{2}D$, and $|g(e^{i\theta})| \leq C \; \forall \theta \in [0, 2\pi]$. Thus $|g(0)| \leq C$ by the maximum modulus principle. So we have

$$C \geq |g(0)| = \frac{|\phi(0)|}{|B(0)|} \geq \prod_{k=1}^{M} \frac{1}{|\lambda_k|} \geq 2^M.$$

□

Lemma 38. *In the same setting as Theorem 37, the following is also true for all $\delta \in (0, 1/3)$: $\{z \in \frac{1}{4}D : |\phi| < \delta\} \subseteq \bigcup_{1 \leq k \leq M} B(\lambda_k, \epsilon)$, where*

$$\epsilon := \frac{9}{16}(3\delta)^{1/M} \leq \frac{9}{16}(3\delta)^{1/log_2(C)}.$$

Proof. Let $\delta \in (0, 1/3)$, and let $z \in \frac{1}{4}D$ such that $|z - \lambda_k| > \epsilon \; \forall k$. Note that g is harmonic and nonzero on $\frac{1}{2}D$ with $|g(0)| \geq 2^M$. Thus Harnack's inequality ensures that $|g| \geq \frac{1}{3}2^M$ on $\frac{1}{4}D$, so there

$$|\phi(z)| \geq |g(z)B(z)| \geq \frac{1}{3}2^M \prod_{k=1}^{M} |\frac{z - \lambda_k}{1 - \bar{\lambda}_k z}| \geq (\frac{16\epsilon}{9})^M \frac{1}{3} = \delta.$$

We can conclude the proof by the contrapositive. $\square$

5.4 A localized upper bound on $||P_1||_2$.

By manipulating some estimates with Poisson kernels, it is possible to localize information about $||f_n||^2$ to say something about $||P_1 \cdot \chi_I||^2$ for an arbitrary interval I. We used this to show that P_1 doesn't "live too much on small intervals," in particular, near the origin, $[0, L^{n-m}]$ - this lemma is used (in the form of Corollary 41) to get (5.9).

The first claim, Lemma 39, uses the Carleson imbedding theorem. It can be skipped, though, as a stronger version, Lemma 40, is proved using general H^2 theory on the upper half-plane $\mathbb{C}_+$. The Carleson imbedding theorem and some H^p theory can be found in [10] and its references.

Lemma 39. *Let $j = 1, 2, ...k$, $c_j \in \mathbb{C}$, $|c_j| = 1$, and $\alpha_j \in \mathbb{R}$. Let $A := \{\alpha_j\}_{j=1}^k$. Then*

$$\int_0^1 |\sum_{j=1}^k c_j e^{i\alpha_j y}|^2 dy \leq C\, k \cdot \sup_{I \text{ a unit interval}} \#\{A \cap I\}\,.$$

Proof. Let $A_1 := \{\mu = \alpha + i : \alpha \in A\}$. Let $\nu := \sum_{\mu \in A_1} \delta_\mu$. This is a measure in $\mathbb{C}_+$. Obviously its Carleson constant

$$\|\nu\|_C := \sup_{J \subset \mathbb{R},\, J \text{ is an interval}} \frac{\nu(J \times [0, |J|])}{|J|}$$

can be estimated as follows

$$\|\nu\|_C \leq 2 \sup_{I \text{ a unit interval}} \#\{A \cap I\}\,. \tag{5.32}$$

Recall that

$$\forall f \in H^2(\mathbb{C}_+) \ \int_{C_+} |f(z)|^2 \, d\nu(z) \le C_0 \, \|\nu\|_C \|f\|_{H^2}^2 \,, \tag{5.33}$$

where C_0 is an absolute constant. Now we compute

$$\int_0^1 |\sum_{j=1}^k c_j e^{i\alpha_j y}|^2 dy \le e^2 \int_0^1 |\sum_{j=1}^k c_j e^{i(\alpha_j + i)y}|^2 dy \le$$

$$e^2 \int_0^\infty |\sum_{j=1}^k c_j e^{i(\alpha_j + i)y}|^2 dy = e^2 \int_{\mathbb{R}} |\sum_{\mu \in A_1} \frac{c_\mu}{x - \mu}|^2 \,,$$

where $c_\mu := c_j$ for $\mu = \alpha_j + i$. The last equality is by Plancherel's theorem.

We continue

$$\int_{\mathbb{R}} |\sum_{\mu \in A_1} \frac{c_\mu}{x - \mu}|^2 = \sup_{f \in H^2(C_+), \, \|f\|_2 \le 1} \left| \langle f, \sum_{\mu \in A_1} \frac{c_\mu}{x - \mu} \rangle \right|^2 =$$

$$4\pi^2 \sup_{f \in H^2(C_+), \, \|f\|_2 \le 1} |\sum_{\mu \in A_1} c_\mu f(\mu)|^2 \le C \, \#\{A_1\} \sup_{f \in H^2(C_+), \, \|f\|_2 \le 1} \sum_{\mu \in A_1} |f(\mu)|^2 \le$$

$$C \, \#\{A\} \sup_{f \in H^2(C_+), \, \|f\|_2 \le 1} \int_{C_+} |f(z)|^2 \, d\nu(z) \le 2 C_0 C \, \#\{A\} \sup_{I \text{ a unit interval}} \#\{A \cap I\} \,.$$

This is by (5.39) and (5.32). The lemma is proved.

□

Now we are going to prove a stronger assertion by a simpler approach. This stronger assertion is what is used in the main part of the article.

Lemma 40. *Let* $j = 1, 2, ...k$, $c_j \in \mathbb{C}$, $|c_j| = 1$, *and* $\alpha_j \in \mathbb{R}$. *Let* $A := \{\alpha_j\}_{j=1}^k$. *Then Suppose*

$$\int_{\mathbb{R}} (\sum_{\alpha\in A} \chi_{[\alpha-1,\alpha+1]}(x))^2\, dx \le S\,, \tag{5.34}$$

Then there exists an absolute constant C *such that*

$$\int_0^1 |\sum_{\alpha\in A} c_\alpha e^{i\alpha y}|^2\, dy \le C\, S\,. \tag{5.35}$$

Of course, one can change variables and get:

Corollary 41. *Let* $j = 1, 2, ...k$, $c_j \in \mathbb{C}$, $|c_j| = 1$, *and* $\alpha_j \in \mathbb{R}$. *Let* $A := \{\alpha_j\}_{j=1}^k$, *and let* $\delta > 0$. *Suppose*

$$\int_{\mathbb{R}} (\sum_{\alpha\in A} \chi_{[\alpha-\delta,\alpha+\delta]}(x))^2\, dx \le S\,, \tag{5.36}$$

Then there exists an absolute constant C *such that for any* $a \in \mathbb{R}$

$$\int_a^{a+\delta^{-1}} |\sum_{\alpha\in A} c_\alpha e^{i\alpha y}|^2\, dy \le C\, S\,/\delta^2. \tag{5.37}$$

Remark. Lemma 40 is obviously stronger than Lemma 39. In fact, let S_0 be the maximal number of points A in any unit interval. Then

$$f(x) := \sum_{\alpha\in A} \chi_{[\alpha-1,\alpha+1]}(x) \le 2S_0.$$

Now $\int_{\mathbb{R}} f^2(x)dx \le 4kS_0$, where k as above is the cardinality of A. We can put now $S := 4kS_0$, apply Lemma 40 and get the conclusion of Lemma 39. The proof of Lemma 40 does not require the Carleson imbedding theorem. Here it is.

Proof. Using Plancherel's theorem we write

$$\int_0^1 |\sum_{\alpha\in A} c_\alpha e^{i\alpha\, y}\, dy|^2 \le e^2 \int_0^1 |\sum_{\alpha\in A} c_\alpha e^{i(\alpha+i)\, y}\, dy|^2 \le e^2 \int_0^\infty |\sum_{\alpha\in A} c_\alpha e^{i(\alpha+i)\, y}\, dy|^2 =$$

$$e^2 \int_{\mathbb{R}} \Big| \sum_{\alpha\in A} \frac{c_\alpha}{\alpha+i-x}\Big|^2 dx\,.$$

Identify $z = x + iy$ in the usual way ($x, y \in \mathbb{R}$). Let $\mathbb{C}_+ = \{x+iy : y > 0\}$. Let H_0^2 be the space of measurable functions $f : \mathbb{C}_+ \to \mathbb{C}$ such that $\sup_{y>0} \int_R |f(x+iy)|dx < \infty$. Let H^2 be the subspace of H_0^2 consisting of analytic functions. H_0^2 is a Hiblert space:

$$< f_1, \overline{f_2} >:= \sup_{y>0} \int_{\mathbb{R}} f_1(x+iy)\overline{f_2(x+iy)}dx. \tag{5.38}$$

It is a standard fact[2] that $H^2(\mathbb{C}_+)$ is orthogonal to $\overline{H^2(\mathbb{C}_+)}$, implying in particular that if f_1, f_2 are analytic in $\mathbb{C}_+$ with $\int_{\mathbb{R}} |f_j(x+iy)|^2 dx < M$ for all $y > 0$ and for $j = 1, 2$, then

$$0 =< f_1, f_2 >= \int_{\mathbb{R}} f_1(x+iy)\overline{f_2(x+iy)}dx \ \ \forall y > 0. \tag{5.39}$$

In our application, we can use

$$f_1(z) = \sum_{\alpha\in A} \frac{c_\alpha}{\alpha-i-z},\ \ f_2(z) = \sum_{\alpha\in A} \frac{\bar{c_\alpha}}{\alpha-i-z}$$

We can just evaluate at $y = 0$ directly. Note that (5.39) says that $< f_1 - \bar{f_2}, f_1 - \bar{f_2} >= ||f_1||^2 + ||f_2||^2 \ge ||f_1||^2$.

[2] $f_1\overline{f_2}$ is analytic in this case, and by conformal identification of $\mathbb{C}_+ \cup \infty$ with the unit disc, one sees that the complex integral along a circle is equal 0, the "value at infinity" of $f_1\overline{f_2}$ on $\mathbb{C}_+$.

Then we get

$$\int_{\mathbb{R}} \Big| \sum_{\alpha\in A} \frac{c_\alpha}{\alpha+i-x}\Big|^2 dx \le \int_{\mathbb{R}} \Big| \sum_{\alpha\in A} \frac{c_\alpha}{\alpha-i-x} - \sum_{\alpha\in A} \frac{c_\alpha}{\alpha+i-x}\Big|^2 dx$$

$$= \int_{\mathbb{R}} \Big| \sum_{\alpha\in A} \frac{-2ic_\alpha}{1+(\alpha-x)^2}\Big|^2 dx = 4\pi^2 \int_{\mathbb{R}} \Big| \sum_{\alpha\in A} c_\alpha \mathcal{P}_1(\alpha-x)\Big|^2 dx\,,$$

where $\mathcal{P}_1$ is the Poisson kernel in the upper half-plane $\mathbb{C}_+$ at height $h=1$:

$$\mathcal{P}_h(x) := \frac{1}{\pi}\frac{h}{h^2+x^2}\,.$$

We continue by noticing that $\mathcal{P}_1 * \chi_{[\lambda-1,\lambda+1]}(x) \ge c\,\mathcal{P}_1(\lambda-x)$ with absolute positive c. This is an elementary calculation, or, if one wishes, Harnack's inequality. Now we can continue

$$\int_0^1 |\sum_{\alpha\in A} c_\alpha e^{i\alpha\, y}\, dy|^2 \lesssim \int_{\mathbb{R}} \Big|(\mathcal{P}_1 * \sum_{\alpha\in A} c_\alpha \chi_{[\alpha-1,\alpha+1]})(x)\Big|^2 dx\,.$$

Now we use the fact that $f \to \mathcal{P}_1 * f$ is a contraction in $L^2(\mathbb{R})$. So

$$\int_0^1 |\sum_{\alpha\in A} c_\alpha e^{i\alpha\, y}\, dy|^2 \lesssim \int_{\mathbb{R}} |\sum_{\alpha\in A} c_\alpha \chi_{[\alpha-1,\alpha+1]}(x)|^2\, dx \lesssim S\,.$$

The lemma is proved.

□

5.5 Discussion

The reason we were able to prove the stronger estimate for the Sierpinski gasket is exactly given by (4.1) and (4.2). They are a quantified version of the fact that the three-term sum

$\varphi(z) = 1 + e^{iz} + e^{itz}$ is zero if and only if the summands are $e^{\frac{2}{3}j\pi i}$, $j = 0, 1, 2$, and that for such z, $\varphi(3^k z) = 3$ for all integers $k \geq 1$. An alternate argument using this fact in this form is employed in [7]. Both versions of this fact we call by the general term "analytic tiling". It is not a tiling of the interval by projected Cantor squares as in [13], but there is a certain tiling pattern to the zeroes of the Fourier transform.

However, there cannot be such a thing in the general case. Suppose we had 5 self-similarities, and that for for some direction θ, we had $\phi_\theta(x_0) = 1+(-i)+i+e^{2\pi i/3}+e^{4\pi i/3} = 0$. Then clearly, taking fifth powers of the summands results in another zero with exactly the same summands, in complete and utter contrast to the three-point case. Similar examples using partitions into relatively prime roots of unity exist for numbers other than 5. In fact, there are examples where $L = 5$ and for θ in a pathological set of size $>> L^{-m}$, $|\{x \in [L^{n-m}, L^n] : \prod_{k=n-\log n}^{n} |\varphi_\theta(x)| < e^{-cm}\}| > L^{n-\sqrt{\log n}}$. That is, $SSV(t)$ takes up a proportion of I much larger than one that is exponentially small in the number of terms in the product. By taking $\frac{1}{m} \log |P_2|$, one gets a certain ergodic sum which one may hope has nice properties, but for some sets $\mathcal{J}$, such nice properties fail for a set of directions far too large to ignore.

It is not yet known whether some separate argument is valid for this new set of "bad directions." One thought is that perhaps there are "structured" and "pseudo-random" directions, and that a separate argument works for each. In the latter case, a pseudo-random analog of the large deviations theory for i.i.d. random variables may hold. But much remains to be seen.

For example, if one considers $\mathcal{K}_n$ as in [18], one gets $\varphi_\theta(z) = 1 + e^{i\pi z} + e^{i\lambda z} + e^{i(\lambda+\pi)z}$, which has the zero $z = 1$. Then $\varphi_\theta(4^k) = 2(1 + \cos(4^k \lambda))$ for $k > 0$. λ depends continuously on θ, and for fixed λ such an ergodic sampling results in a sequence $a_k := \varphi(4^k)$, and either:

1: a_k is eventually periodic and non-zero,

2: a_k takes values other than 4 only finitely often,

or 3 (the case for almost every λ): $4^k\lambda \bmod 2\pi$ evenly samples $[0, 2\pi]$ over the long term, with long-term average $\frac{1}{N}\sum_{k=1}^{N} \log a_k \to \log 2$ as $N \to \infty$.

This regularity agrees with the result [18], which already proved a result without using ergodic theory or large deviation theory. There was a θ and x separation of variables, and the zeroes obeyed an "analtyic tiling" property like the one for the gasket.

Chapter 6

Epilogue

This thesis was written by a minotaur in a manner compliant with federal policies on the use of human and/or animal subjects in research projects. No harm was sustained by the minotaur except for maybe some loss of sleep and the formation of a coffee dependence.

Figure 6.1: Minotaur.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] M. Bateman, N.Katz, *Kakeya sets in Cantor directions*, arXiv:math/0609187v1, 2006, pp. 1–10.

[2] M. Bateman, *Kakeya sets and the directional maximal operators in the plane,* arXiv:math.CA/0703559v1, 2007, pp. 1–20.

[3] M. Bateman, A. Volberg *An estimate from below for the Buffon needle probability of the four-corner Cantor set*, arXiv:0807.2953v1 [math.CA], 2008.

[4] A. S. Besicovitch, *Tangential properties of sets and arcs of infinite linear measure*, *Bull. Amer. Math. Soc.* **66** (1960), 353–359.

[5] M. Bond, A. Volberg *Buffon needle lands in ϵ-neighborhood of a 1-dimensional Sierpinski Gasket with probability at most* $|\log \epsilon|^{-c}$, *Comptes Rendus Mathematique*, Volume 348, Issues 11-12, June 2010, 653-656

[6] M. Bond, A. Volberg: *Circular Favard Length of the Four-Corner Cantor Set*, *J. of Geometric Analysis*, online July 2010, DOI: 10.1007/s12220-010-9141-4.

[7] M. Bond, A. Volberg: *The power law for Buffon's needle landing near the Sierpinski gasket*, arXiv: 0911.0233v2, 2009.

[8] J. Bourgain, *Averages in the plane over convex curves and maximal operators,* *J. Analyse Math.* **47** (1986), 69–85.

[9] G. David, *Analytic capacity, Calderón-Zygmund operators, and rectifiability,* *Publ. Mat.* **43** (1999),3–25.

[10] J. Garnett, *Bounded Analytic Functions*, *Springer Graduate Texts in Mathematics* **236**, 2007.

[11] P. W. Jones and T. Murai, *Positive analytic capacity but zero Buffon needle probability*, *Pacific J. Math.* **133** (1988), 99–114.

[12] R. Kenyon, *Projecting the one-dimensional Sierpinski gasket,* *Israel J. Math.* **97** (1997), 221–238.

[13] I. Łaba, K. Zhai, *The Favard length of product Cantor sets*, *Bulletin of the London Mathematical Society*, doi: 10.1112/blms/bdq059, 2010.

[14] J. C. Lagarias and Y. Wang, *Tiling the line with translates of one tile*, *Invent. Math.* **124** (1996), 341–365.

[15] J. Mateu, X. Tolsa and J. Verdera, *The planar Cantor sets of zero analytic capacity and the local $T(b)$-theorem. J. Amer. Math. Soc.* **16** (2003), 19–28.

[16] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, *Cambridge University Press*, 1995.

[17] P. Mattila, On the analytic capacity and curvature of some Cantor sets with non-σ-finite length, *Publ. Mat.* **40** (1996),no. 1, 195–204.

[18] F. Nazarov, Y. Peres, A. Volberg, *The power law for the Buffon needle probability of the four-corner Cantor set*, arXiv:0801.2942, 2008.

[19] H. Pajot. *Analytic Capacity, Rectifiability, Menger Curvature and the Cauchy Integral*, *Lecture Notes in Mathematics*, vol. 1799, Springer, Berlin, 2002.

[20] Y. Peres and B. Solomyak, *How likely is Buffon's needle to fall near a planar Cantor set? Pacific J. Math. 204*, 2 (2002), 473–496.

[21] F. Nazarov, *Local estimates of exponential polynomials and their applications to inequalities of uncertainty principle type* , St Petersburg Math. J., v. 5 (1994), No. 4, pp. 3–66.

[22] A. Seeger, T. Tao, J. Wright, *Singular Maximal Functions and Radon Transforms near L^1*, *Amer. J. Math*, 126 (2002), 607–647.

[23] E.M. Stein, *Maximal functions: Spherical means*, *Proc. Nat. Acad. Sci. U.S.A.*, 73 (1976), 2174–2175.

[24] T. Tao, *A quantitative version of the Besicovitch projection theorem via multiscale analysis,* pp. 1–28, arXiv:0706.2446v1 [math.CA] 18 Jun 2007.

[25] X. Tolsa, *Analytic capacity, rectifiability, and the Cauchy integral*, Proceedings of the ICM, 2006, Madrid.