Book Section:

Beresnevich, Victor orcid.org/0000-0002-1811-9697, Ramirez, Felipe Alberto and Velani, Sanju orcid.org/0000-0002-4442-6316 (2017) Metric Diophantine Approximation : aspects of recent work. In: Badziahin, D., Gorodnik, A. and Peyerimhoff, N., (eds.) Dynamics and Analytic Number Theory. London Mathematical Society Lecture Note Series . Cambridge University Press , pp. 1-95.

Metric Diophantine Approximation: aspects of recent work

Victor Beresnevich, Felipe Ramírez and Sanju Velani

University of York

Abstract

In these notes, we begin by recalling aspects of the classical theory of metric Diophantine approximation; such as theorems of Khintchine, Jarník, Duffin-Schaeffer and Gallagher. We then describe recent strengthening of various classical statements as well as recent developments in the area of Diophantine approximation on manifolds. The latter includes the well approximable, the badly approximable and the inhomogeneous aspects.

This Chapter is to be published by Cambridge University Press as part of a multi-volume work edited by Badziahin, D., Gorodnik, A., Peyerimhoff, N.

FR is supported by EPSRC Programme Grant: EP/J018260/1.
VB and SV are supported in part by EPSRC Programme Grant: EP/J018260/1.

Contents

1 Background: Dirichlet and Bad

1.1 Dirichlet's Theorem and two important consequences

Diophantine approximation is a branch of number theory that can loosely be described as a quantitative analysis of the density of the rationals $\mathbb{Q}$ in the reals $\mathbb{R}$. Recall that to say that $\mathbb{Q}$ is dense in $\mathbb{R}$ is to say that

> for any real number x and $\epsilon > 0$ there exists a rational number p/q $(q > 0)$ such that $|x - p/q| < \epsilon$.

In other words, any real number can be approximated by a rational number with any assigned degree of accuracy. But how "rapidly" can we approximate a given $x \in \mathbb{R}$?

> Given $x \in \mathbb{R}$ and $q \in \mathbb{N}$, how small can we make ϵ? Trivially we can take any $\epsilon > 1/2q$. Can we do better than $1/2q$?

The following rational numbers all lie within $1/(\text{denominator})^2$ of the circle constant $\pi = 3.141\ldots$:

$$\frac{3}{1}, \frac{22}{7}, \frac{333}{106}, \frac{355}{113}, \frac{103993}{33102}. \tag{1.1}$$

This shows that, at least sometimes, the answer to the last question is "yes." A more complete answer is given by *Dirichlet's theorem*, which is itself a simple consequence of the following powerful fact.

Pigeonhole Principle. *If n objects are placed in m boxes and $n > m$, then some box will contain at least two objects.*

Theorem 1.1 (Dirichlet, 1842)**.** *For any $x \in \mathbb{R}$ and $N \in \mathbb{N}$, there exist $p, q \in \mathbb{Z}$ such that*

$$\left| x - \frac{p}{q} \right| < \frac{1}{qN} \qquad \textit{and} \qquad 1 \leq q \leq N\,. \tag{1.2}$$

The proof can be found in most elementary number theory books. However, given the important consequences of the theorem and its various hybrids, we have decided to include the proof.

Proof. As usual, let $[x] := \max\{n \in \mathbb{Z} : n \leq x\}$ denote the integer part of the real number x and let $\{x\} = x - [x]$ denote the fractional part of x. Note that for any $x \in \mathbb{R}$ we have that $0 \leq \{x\} < 1$.

Consider the $N + 1$ numbers

$$\{0x\}, \{x\}, \{2x\}, \ldots, \{Nx\} \tag{1.3}$$

in the unit interval $[0,1)$. Divide $[0,1)$ into N equal semi-open subintervals as follows:

$$[0,1)=\bigcup_{u=0}^{N-1} I_u \quad \text{where} \quad I_u:=\left[\frac{u}{N},\frac{u+1}{N}\right), \qquad u=0,1,\ldots,N-1. \tag{1.4}$$

Since the $N+1$ points (1.3) are situated in the N subintervals (1.4), the Pigeonhole principle guarantees that some subinterval contains at least two points, say $\{q_2x\},\{q_1x\}\in I_u$, where $0\le u\le N-1$ and $q_1,q_2\in\mathbb{Z}$ with $0\le q_1<q_2\le N$. Since the length of I_u is N^{-1} and I_u is semi-open we have that

$$|\{q_2x\}-\{q_1x\}|<\frac{1}{N}. \tag{1.5}$$

We have that $q_ix=p_i+\{q_ix\}$ where $p_i=[q_ix]\in\mathbb{Z}$ for $i=1,2$. Returning to (1.5) we get

$$|\{q_2x\}-\{q_1x\}|=|q_2x-p_2-(q_1x-p_1)|=|(q_2-q_1)x-(p_2-p_1)|. \tag{1.6}$$

Now define $q=q_2-q_1\in\mathbb{Z}$ and $p=p_2-p_1\in\mathbb{Z}$. Since $0\le q_1,q_2\le N$ and $q_1<q_2$ we have that $1\le q\le N$. By (1.5) and (1.6), we get

$$|qx-p|<\frac{1}{N}$$

whence (1.2) readily follows. $\square$

The following statement is an important consequence of Dirichlet's Theorem.

Theorem 1.2 (Dirichlet, 1842)**.** *Let $x\in\mathbb{R}\setminus\mathbb{Q}$. Then there exist infinitely many integers q,p such that* $\gcd(p,q)=1$, $q>0$ *and*

$$\left|x-\frac{p}{q}\right|<\frac{1}{q^2}. \tag{1.7}$$

Remark 1.1. Theorem 1.2 is true for all $x\in\mathbb{R}$ if we remove the condition that p and q are coprime, that is, if we allow approximations by non-reduced rational fractions.

Proof. Observe that Theorem 1.1 is valid with $\gcd(p,q)=1$. Otherwise $p/q=p'/q'$ with $\gcd(p',q')=1$ and $0<q'<q\le N$ and $|x-p/q|=|x-p'/q'|<1/(qN)<1/(q'N)$.

Suppose x is irrational and that there are only finitely many rationals

$$\frac{p_1}{q_1},\frac{p_2}{q_2},\ldots,\frac{p_n}{q_n},$$

where $\gcd(p_i,q_i)=1$, $q_i>0$ and

$$\left|x-\frac{p_i}{q_i}\right|<\frac{1}{q_i^2}$$

for all $i=1,2,\ldots,n$. Since x is irrational, $x-\frac{p_i}{q_i}\ne 0$ for $i=1,\ldots,n$. Then there exists $N\in\mathbb{N}$ such that

$$\left|x-\frac{p_i}{q_i}\right|>\frac{1}{N} \qquad \text{for all } 1\le i\le n.$$

By Theorem 1.1, there exists a reduced fraction $\dfrac{p}{q}$ such that

$$\left|x - \frac{p}{q}\right| < \frac{1}{qN} \leq \frac{1}{N} \qquad (1 \leq q \leq N).$$

Therefore, $\frac{p}{q} \neq \frac{p_i}{q_i}$ for any i but satisfies (1.7). A contradiction. $\square$

Theorem 1.2 tells us in particular that the list (1.1) of good rational approximations to π is not just a fluke. This list can be extended to an infinite sequence, and furthermore, such a sequence of good approximations exists for *every* irrational number. (See §1.2.)

Another important consequence of Theorem 1.1 is Theorem 1.3, below. Unlike Theorem 1.2, the significance of it is not so immediately clear. However, it will become apparent during the course of these notes that it is the key to the two fundamental theorems of classical metric Diophantine approximation; namely the theorems of Khintchine and Jarník.

First, some notational matters. Unless stated otherwise, given a set $X \subset \mathbb{R}$, we will denote by $m(X)$ the 1-dimensional Lebesgue measure of X. And we will use $B(x,r)$ to denote $(x-r, x+r) \subset \mathbb{R}$, the ball around $x \in \mathbb{R}$ of radius $r > 0$.

Theorem 1.3. *Let $[a,b] \subset \mathbb{R}$ be an interval and $k \geq 6$ be an integer. Then*

$$m\left([a,b] \cap \bigcup_{k^{n-1}<q\leq k^n} \bigcup_{p\in\mathbb{Z}} B\left(\tfrac{p}{q}, \tfrac{k}{k^{2n}}\right)\right) \;\geq\; \tfrac{1}{2}\,(b-a).$$

for all sufficiently large $n \in \mathbb{N}$.

Proof. By Dirichlet's theorem, for any $x \in I := [a,b]$ there are coprime integers p, q with $1 \leq q \leq k^n$ satisfying $|x - p/q| < (qk^n)^{-1}$. We therefore have that

$$\begin{aligned} m(I) &= m\left(I \cap \bigcup_{q\leq k^n} \bigcup_{p\in\mathbb{Z}} B\Big(\frac{p}{q}, \frac{1}{qk^n}\Big)\right) \\ &\leq m\left(I \cap \bigcup_{q\leq k^{n-1}} \bigcup_{p\in\mathbb{Z}} B\Big(\frac{p}{q}, \frac{1}{qk^n}\Big)\right) + m\left(I \cap \bigcup_{k^{n-1}<q\leq k^n} \bigcup_{p\in\mathbb{Z}} B\Big(\frac{p}{q}, \frac{k}{k^{2n}}\Big)\right). \end{aligned}$$

Also, notice that

$$\begin{aligned} m\left(I \cap \bigcup_{q\leq k^{n-1}} \bigcup_{p\in\mathbb{Z}} B\Big(\frac{p}{q}, \frac{1}{qk^n}\Big)\right) &= m\left(I \cap \bigcup_{q\leq k^{n-1}} \bigcup_{p=aq-1}^{bq+1} B\Big(\frac{p}{q}, \frac{1}{qk^n}\Big)\right) \\ &\leq 2 \sum_{q\leq k^{n-1}} \frac{1}{qk^n}\Big(m(I)q + 3\Big) \leq \frac{3}{k} m(I) \end{aligned}$$

for large n. It follows that for $k \geq 6$,

$$m\left(I \cap \bigcup_{k^{n-1}<q\leq k^n} \bigcup_{p\in\mathbb{Z}} B\left(\tfrac{p}{q}, \tfrac{k}{k^{2n}}\right)\right) \;\geq\; m(I) - \tfrac{3}{k}\, m(I) \;\geq\; \tfrac{1}{2}\, m(I)$$

for large n. $\square$

1.2 Basics of continued fractions

From Dirichlet's theorem we know that for any real number x there are infinitely many 'good' rational approximates p/q, but how can we find these? The theory of continued fraction provides a simple mechanism for generating them. We collect some basic facts about continued fractions in this section. For proofs and a more comprehensive account see for example [57, 66, 80].

Let x be an irrational number and let $[a_0; a_1, a_2, a_3, \ldots]$ denote its continued fraction expansion. Denote its n-th convergent by

$$\frac{p_n}{q_n} := [a_0; a_1, a_2, a_3, \ldots, a_n].$$

Recall that the convergents can be obtained by the following recursion

$$\begin{aligned} p_0 &= a_0, & q_0 &= 1, \\ p_1 &= a_1 a_0 + 1, & q_1 &= a_1, \\ p_k &= a_k p_{k-1} + p_{k-2}, & q_k &= a_k q_{k-1} + q_{k-2} \qquad \text{for } k \geq 2, \end{aligned}$$

and that they satisfy the inequalities

$$\frac{1}{q_n(q_{n+1} + q_n)} \leq \left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}}. \tag{1.8}$$

From this it is clear that the convergents provide explicit solutions to the inequality in Theorem 1.2 (Dirichlet); that is,

$$\left| x - \frac{p_n}{q_n} \right| \leq \frac{1}{q_n^2} \qquad \forall n \in \mathbb{N}.$$

In fact, it turns out that for irrational x the convergents are *best approximates* in the sense that if $1 \leq q < q_n$ then any rational $\frac{p}{q}$ satisfies

$$\left| x - \frac{p_n}{q_n} \right| < \left| x - \frac{p}{q} \right|.$$

Regarding $\pi = 3.141\ldots$, the rationals (1.1) are the first 5 convergents.

1.3 Competing with Dirichlet and losing badly

We have presented Dirichlet's theorem as an answer to whether the trivial inequality $|x - p/q| \leq 1/2q$ can be beaten. Naturally, one may also ask if we can do any better than Dirichlet's theorem. Let us formulate this a little more precisely. For $x \in \mathbb{R}$, let

$$\|x\| := \min\{|x - m| : m \in \mathbb{Z}\}$$

denote the distance from x to the nearest integer. Dirichlet's theorem (Theorem 1.2) can be restated as follows: *for any $x \in \mathbb{R}$, there exist infinitely many integers $q > 0$ such that*

$$q \, \|qx\| \leq 1. \tag{1.9}$$

Can we replace right-hand side of (1.9) by arbitrary $\epsilon > 0$? In other words, is it true that $\liminf_{q\to\infty} q\|qx\| = 0$ for every x? One might notice that (1.8) implies that there certainly do exist x for which this is true. (One can write down a continued fraction whose partial quotients grow as fast as one pleases.) Still, the answer to the question is **No.** It was proved by Hurwitz (1891) that for every $x \in \mathbb{R}$, we have $q\,\|qx\| < \epsilon = 1/\sqrt{5}$ for infinitely many $q > 0$, and that this is best possible in the sense that the statement becomes false if $\epsilon < 1/\sqrt{5}$.

The fact that $1/\sqrt{5}$ is best possible is relatively easy to see. Assume that it can be replaced by

$$\frac{1}{\sqrt{5}+\epsilon} \qquad (\epsilon > 0,\ \text{arbitrary}).$$

Consider the Golden Ratio $x_1 = \frac{\sqrt{5}+1}{2}$, root of the polynomial

$$f(t) = t^2 - t - 1 = (t - x_1)(t - x_2)$$

where $x_2 = \frac{1-\sqrt{5}}{2}$. Assume there exists a sequence of rationals $\frac{p_i}{q_i}$ satisfying

$$\left| x_1 - \frac{p_i}{q_i} \right| < \frac{1}{(\sqrt{5}+\epsilon)q_i^2}.$$

Then, for i sufficiently large, the right-hand side of the above inequality is less than ϵ and so

$$\left| x_2 - \frac{p_i}{q_i} \right| \le |x_2 - x_1| + \left| x_1 - \frac{p_i}{q_i} \right| < \sqrt{5} + \epsilon\,.$$

It follows that

$$0 \ne \left| f\left(\frac{p_i}{q_i}\right) \right| < \frac{1}{(\sqrt{5}+\epsilon)q_i^2} \cdot (\sqrt{5}+\epsilon)$$

$$\implies \quad \left| q_i^2 f\left(\frac{p_i}{q_i}\right) \right| < 1.$$

However the left-hand side is a strictly positive integer. This is a contradiction, for there are no integers in $(0,1)$—an extremely useful fact.

The above argument shows that if $x = \frac{\sqrt{5}+1}{2}$ then there are at most finitely many rationals p/q such that

$$\left| x - \frac{p}{q} \right| < \frac{1}{(\sqrt{5}+\epsilon)q^2}.$$

Therefore, there exists a constant $c(x) > 0$ such that

$$\left| x - \frac{p}{q} \right| > \frac{c(x)}{q^2} \qquad \forall\ p/q \in \mathbb{Q}\,.$$

All of this shows that there exist numbers for which we can not improve Dirichlet's theorem arbitrarily. These are called *badly approximable numbers* and are defined by

$$\begin{aligned}\mathbf{Bad} &:= \{x \in \mathbb{R} : \inf_{q\in\mathbb{N}} q\|qx\| > 0\} \\ &= \{x \in \mathbb{R} : c(x) := \liminf_{q\to\infty} q\|qx\| > 0\}.\end{aligned}$$

Note that if x is badly approximable then for the associated badly approximable constant $c(x)$ we have that

$$0 < c(x) \leq \frac{1}{\sqrt{5}}.$$

Clearly, **Bad** $\neq \varnothing$ since the golden ratio is badly approximable. Indeed, if $x \in$ **Bad** then $tx \in$ **Bad** for any $t \in \mathbb{Z} \setminus \{0\}$ and so **Bad** is at least countable.

Bad has a beautiful characterisation via continued fractions.

Theorem 1.4. *Let $x = [a_0; a_1, a_2, a_3, \ldots]$ be irrational. Then*

$$x \in \mathbf{Bad} \iff \exists\ M = M(x) \geq 1 \text{ such that } a_i \leq M \quad \forall i\ .$$

That is, **Bad** *consists exactly of the real numbers whose continued fractions have bounded partial quotients.*

Proof. It follows from (1.8) that

$$\frac{1}{q_n^2(a_{n+1}+2)} \ \leq\ \left|x - \frac{p_n}{q_n}\right| \ <\ \frac{1}{a_{n+1}q_n^2}, \tag{1.10}$$

and from this it immediately follows that if $x \in$ **Bad**, then $a_n \leq \max\{|a_o|, 1/c(x)\}$.

Conversely, suppose the partial quotients of x are bounded, and take any $q \in \mathbb{N}$. Then there is $n \geq 1$ such that $q_{n-1} \leq q < q_n$. On using the fact that convergents are best approximates, it follows that

$$\left|x - \frac{p}{q}\right| \ \geq\ \left|x - \frac{p_n}{q_n}\right| \ \geq\ \frac{1}{q_n^2(M+2)} \ =\ \frac{1}{q^2(M+2)}\,\frac{q^2}{q_n^2}.$$

It is easily seen that

$$\frac{q}{q_n} \geq \frac{q_{n-1}}{q_n} \geq \frac{1}{M+1},$$

which proves that

$$c(x) \geq \frac{1}{(M+2)(M+1)^2} > 0,$$

hence $x \in$ **Bad**. □

Recall that a continued fraction of the form $x = [a_0; \ldots, a_n, \overline{a_{n+1}, \ldots, a_{n+m}}]$ is said to be *periodic*. Also, recall that an irrational number α is called a *quadratic irrational* if α is a solution to a quadratic equation with integer coefficients:

$$ax^2 + bx + c = 0 \qquad (a, b, c \in \mathbb{Z},\ a \neq 0).$$

It is a well-known fact that an irrational number x has periodic continued fraction expansion if and only if x is a quadratic irrational. This and Theorem 1.4 imply the following corollary.

Corollary 1.1. *Every quadratic irrational is badly approximable.*

The simplest instance of this is the golden ratio, a root of $x^2 - x - 1$, whose continued fraction is

$$\frac{\sqrt{5}+1}{2} = [1; 1, 1, 1, \ldots] := [\,\overline{1}\,],$$

with partial quotients clearly bounded.

Indeed, much is known about the badly approximable numbers, yet several simple questions remain unanswered. For example:

Folklore Conjecture. *The only algebraic irrationals that are in* **Bad** *are the quadratic irrationals.*

Remark 1.2. Though this conjecture is widely believed to be true, there is no direct evidence for it. That is, there is no single algebraic irrational of degree greater than two whose membership (or non-membership) in **Bad** has been verified.

A particular goal of these notes is to investigate the 'size' of **Bad**. We will show:

(a) $m(\mathbf{Bad}) = 0$

(b) $\dim \mathbf{Bad} = 1$,

where dim refers to the Hausdorff dimension (see §3.1). In other words, we will see that **Bad** is a small set in that it has measure zero in $\mathbb{R}$, but it is a large set in that it has the same (Hausdorff) dimension as $\mathbb{R}$.

Let us now return to Dirichlet's theorem (Theorem 1.2). Every $x \in \mathbb{R}$ can be approximated by rationals p/q with 'rate of approximation' given by q^{-2}—the right-hand side of inequality (1.7) determines the 'rate' or 'error' of approximation by rationals. The above discussion shows that this rate of approximation cannot be improved by an arbitrary constant for every real number—**Bad** is non-empty. On the other hand, we have stated above that **Bad** is a 0-measure set, meaning that the set of points for which we *can* improve Dirichlet's theorem by an arbitrary constant is full. In fact, we will see that if we exclude a set of real numbers of measure zero, then from a measure theoretic point of view the rate of approximation can be improved not just by an arbitrary constant but by a logarithm (see Remark 2.3).

2 Metric Diophantine approximation: the classical Lebesgue theory

In the previous section, we have been dealing with variations of Dirichlet's theorem in which the right-hand side or rate of approximation is of the form ϵq^{-2}. It is natural to broaden the discussion to general approximating functions. More precisely, for a function $\psi : \mathbb{N} \to \mathbb{R}^+ = [0, \infty)$, a real number x is said to be *ψ–approximable* if there are infinitely many $q \in \mathbb{N}$ such that

$$\|qx\| < \psi(q) \ . \tag{2.1}$$

The function ψ governs the 'rate' at which the rationals approximate the reals and will be referred to as an *approximating function*.

One can readily verify that the set of ψ-approximable numbers is invariant under translations by integer vectors. Therefore without any loss of generality, and to ease the 'metrical' discussion which follows, we shall restrict our attention to ψ–approximable numbers in the unit interval $\mathrm{I} := [0, 1)$. The set of such numbers is clearly a subset of I and will be denoted by $W(\psi)$; i.e.

$$W(\psi) := \{x \in \mathrm{I} \colon \|qx\| < \psi(q) \text{ for infinitely many } q \in \mathbb{N}\} .$$

Notice that in this notation we have that

$$\text{Dirichlet's Theorem (Theorem 1.2)} \quad \Longrightarrow \quad W(\psi) = \mathrm{I} \ \text{ if } \ \psi(q) = q^{-1}.$$

Yet, the existence of badly approximable numbers implies that there exist approximating functions ψ for which $W(\psi) \neq I$. Furthermore, the fact that $m(\mathbf{Bad}) = 0$ implies that we can have $W(\psi) \neq I$ while $m(W(\psi)) = 1$.

A key aspect of the classical theory of Diophantine approximation is to determine the 'size' of $W(\psi)$ in terms of

(a) Lebesgue measure,

(b) Hausdorff dimension, and

(c) Hausdorff measure.

From a measure theoretic point of view, as we move from (a) to (c) in the above list, the notion of size becomes subtler. In this section we investigate the 'size' of $W(\psi)$ in terms of 1- dimensional Lebesgue measure m.

We start with the important observation that $W(\psi)$ is a lim sup set of balls. For a fixed $q \in \mathbb{N}$, let

$$\begin{aligned} A_q(\psi) \ &:= \ \{x \in \mathrm{I} : \|qx\| \ < \ \psi(q)\} \\ &:= \ \bigcup_{p=0}^{q} B\Big(\frac{p}{q}, \frac{\psi(q)}{q}\Big) \ \cap \mathrm{I}\,. \end{aligned} \tag{2.2}$$

Note that

$$m\big(A_q(\psi)\big) \leqslant 2\psi(q) \tag{2.3}$$

with equality when $\psi(q) < 1/2$ since then the intervals in (2.2) are disjoint.

The set $W(\psi)$ is simply the set of real numbers in I which lie in infinitely many sets $A_q(\psi)$ with $q = 1, 2, \ldots$ i.e.

$$W(\psi) = \limsup_{q \to \infty} A_q(\psi) := \bigcap_{t=1}^{\infty} \bigcup_{q=t}^{\infty} A_q(\psi)$$

is a lim sup set. Now notice that for each $t \in \mathbb{N}$

$$W(\psi) \subset \bigcup_{q=t}^{\infty} A_q(\psi)$$

i.e. for each t, the collection of balls $B(p/q, \psi(q)/q)$ associated with the sets $A_q(\psi) : q = t, t+1, \ldots$ form a cover for $W(\psi)$. Thus, it follows via (2.3) that

$$\begin{aligned} m\big(W(\psi)\big) &\leq m\left(\bigcup_{q=t}^{\infty} A_q(\psi)\right) \\ &\leq \sum_{q=t}^{\infty} m\big(A_q(\psi)\big) \\ &\leq 2\sum_{q=t}^{\infty} \psi(q)\,. \end{aligned} \tag{2.4}$$

Now suppose

$$\sum_{q=1}^{\infty} \psi(q) < \infty.$$

Then given any $\epsilon > 0$, there exists t_0 such that for all $t \geq t_0$

$$\sum_{q=t}^{\infty} \psi(q) < \frac{\epsilon}{2}.$$

It follows from (2.4), that

$$m\big(W(\psi)\big) < \epsilon.$$

But $\epsilon > 0$ is arbitrary, whence

$$m\big(W(\psi)\big) = 0$$

and we have established the following statement.

Theorem 2.1. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a function such that*

$$\sum_{q=1}^{\infty} \psi(q) < \infty.$$

Then

$$m(W(\psi)) = 0.$$

This theorem is in fact a simple consequence of a general result in probability theory.

2.1 The Borel-Cantelli Lemma

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space with $\mu(\Omega) < \infty$ and let E_q $(q \in \mathbb{N})$ be a family of measurable sets in Ω. Also, let

$$E_\infty := \limsup_{q\to\infty} E_q := \bigcap_{t=1}^{\infty} \bigcup_{q=t}^{\infty} E_q \ ;$$

i.e. E_∞ is the set of $x \in \Omega$ such that $x \in E_i$ for infinitely many $i \in \mathbb{N}$.

The proof of the Theorem 2.1 mimics the proof of the following fundamental statement from probability theory.

Lemma 2.1 (Convergence Borel-Cantelli)**.** *Suppose that* $\sum_{q=1}^{\infty} \mu(E_q) < \infty$. *Then,*

$$\mu(E_\infty) = 0 \, .$$

Proof. Exercise. □

To see that Theorem 2.1 is a trivial consequence of the above lemma, simply put $\Omega = \mathrm{I} = [0,1]$, $\mu = m$ and $E_q = A_q(\psi)$ and use (2.3).

Now suppose we are in a situation where the sum of the measures diverges. Unfortunately, as the following example demonstrates, it is not the case that if $\sum \mu(E_q) = \infty$ then $\mu(E_\infty) = \mu(\Omega)$ or indeed that $\mu(E_\infty) > 0$.

Example: Let $E_q = (0, \frac{1}{q})$. Then $\sum_{q=1}^{\infty} m(E_q) = \sum_{q=1}^{\infty} \frac{1}{q} = \infty$. However, for any $t \in \mathbb{N}$ we have that

$$\bigcup_{q=t}^{\infty} E_q = E_t \, ,$$

and thus

$$E_\infty = \bigcap_{t=1}^{\infty} E_t = \bigcap_{t=1}^{\infty} (0, \tfrac{1}{t}) = \varnothing$$

implying that $m(E_\infty) = 0$.

The problem in the above example is that the sets E_q overlap 'too much'—in fact they are nested. The upshot is that in order to have $\mu(E_\infty) > 0$, we not only need the sum of the measures to diverge but also that the sets E_q $(q \in \mathbb{N})$ are in some sense independent. Indeed, it is well-known that if we had pairwise independence in the standard sense; *i.e.* if

$$\mu(E_s \cap E_t) = \mu(E_s)\mu(E_t) \qquad \forall s \neq t,$$

then we would have $\mu(E_\infty) = \mu(\Omega)$. However, we very rarely have this strong form of independence in our applications. What is much more useful to us is the following statement, whose proof can be found in [58, 90].

Lemma 2.2 (Divergence Borel-Cantelli)**.** *Suppose that* $\sum_{q=1}^{\infty} \mu(E_q) = \infty$ *and that there exists a constant* $C > 0$ *such that*

$$\sum_{s,t=1}^{Q} \mu(E_s \cap E_t) \leq C \left(\sum_{s=1}^{Q} \mu(E_s) \right)^2 \tag{2.5}$$

holds for infinitely many $Q \in \mathbb{N}$*. Then*

$$\mu(E_\infty) \geq 1/C \, .$$

The independence condition (2.5) is often referred to as *quasi-independence on average*, and, together with the divergent sum condition, it guarantees that the associated lim sup set has positive measure. It does not guarantee *full* measure (*i.e.* that $\mu(E_\infty) = \mu(\Omega)$), which is what we are trying to prove, for example, in Khintchine's Theorem. But this is not an issue if we already know (by some other means) that E_∞ satisfies a zero-full law (which is also often called a zero-one law) with respect to the measure μ, meaning a statement guaranteeing that

$$\mu(E_\infty) = \ 0 \quad \text{or} \quad \mu(\Omega).$$

Happily, this is the case with the lim sup set $W(\psi)$ of ψ-well approximable numbers [38, 37, 58].

Alternatively, assuming Ω is equipped with a metric such that μ becomes a doubling Borel measure, we can guarantee that $\mu(E_\infty) = \mu(\Omega)$ if we can establish *local quasi-independence on average* [14, §8]; *i.e.* we replace (2.5) in the above lemma by the condition that

$$\sum_{s,t=1}^{Q} \mu\big((B \cap E_s) \cap (B \cap E_t)\big) \leq \frac{C}{\mu(B)} \left(\sum_{s=1}^{Q} \mu(B \cap E_s) \right)^2 \tag{2.6}$$

for any sufficiently small ball B with center in Ω and $\mu(B) > 0$. The constant C is independent of the ball B. Recall that μ is doubling if $\mu(2B) \ll \mu(B)$ for balls B centred in Ω. In some literature such measures are also referred to as Federer measures.

The Divergence Borel-Cantelli Lemma is key to determining $m(W(\psi))$ in the case where $\sum_{q=1}^{\infty} \psi(q)$ diverges—the subject of the next section and the main substance of Khintchine's Theorem. Before turning to this, let us ask ourselves one final question regarding quasi-independence on average and positive measure of lim sup sets.

Question. Is the converse to Divergence Borel-Cantelli true? More precisely, if $\mu(E_\infty) > 0$ then is it true that the sets E_t are quasi-independent on average?

The following theorem is a consequence of a more general result established in [29].

Theorem 2.2. *Let* (Ω, d) *be a compact metric space equipped with a Borel probability measure* μ*. Let* E_q $(q \in \mathbb{N})$ *be a sequence of balls in* Ω *such that* $\mu(E_\infty) > 0$*. Then, there exists a strictly increasing sequence of integers* $(q_k)_{k \in \mathbb{N}}$ *such that* $\sum_{k=1}^{\infty} \mu(E_{q_k}) = \infty$ *and the balls* E_{q_k} $(k \in \mathbb{N})$ *are quasi-independent on average.*

2.2 Khintchine's Theorem

The following fundamental statement in metric Diophantine approximation (of which Theorem 2.1 is the "easy case") provides an elegant criterion for the 'size' of the set $W(\psi)$ expressed in terms of Lebesgue measure.

Theorem 2.3 (Khintchine, 1924)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function. Then*

$$m(W(\psi)) = \begin{cases} 0 & \text{if} \quad \sum_{q=1}^{\infty} \psi(q) < \infty \,, \\ \\ 1 & \text{if} \quad \sum_{q=1}^{\infty} \psi(q) = \infty \,. \end{cases}$$

Remark 2.1. It is worth mentioning that Khintchine's original statement [64] made the stronger assumption that $q\psi(q)$ is monotonic.

Remark 2.2. The assumption that ψ is monotonic is only required in the divergent case. It cannot in general be removed—see §2.2.1 below.

Remark 2.3. Khintchine's Theorem implies that

$$m(W(\psi)) = 1 \quad \text{if} \quad \psi(q) = \frac{1}{q \log q} \,.$$

Thus, from a measure theoretic point of view the 'rate' of approximation given by Dirichlet's theorem can be improved by a logarithm.

Remark 2.4. As mentioned in the previous section, in view of Cassels' zero-full law [38] (also known as zero-one) we know that $m(W(\psi)) = 0$ or 1 regardless of whether or not ψ is monotonic.

Remark 2.5. A key ingredient to directly establishing the divergent part is to show that the sets

$$A_s^* = A_s^*(\psi) := \bigcup_{2^{s-1} \le q < 2^s} \ \bigcup_{p=0}^{q} B\Big(\frac{p}{q}, \frac{\psi(2^s)}{2^s}\Big) \ \cap \mathrm{I} \,.$$

are quasi-independent on average. Notice that

- For ψ monotonic, $W(\psi) \supset W^*(\psi) := \limsup_{s\to\infty} A_s^*(\psi)$.
- If $\psi(q) < q^{-1}$, the balls in $A_s^*(\psi)$ are disjoint and so

$$m(A_s^*(\psi)) \ \asymp \ 2^s \psi(2^s) \,.$$

- For ψ monotonic, $\sum \psi(q) \asymp \sum 2^s \psi(2^s)$.

Notation. Throughout, the Vinogradov symbols $\ll$ and $\gg$ will be used to indicate an inequality with an unspecified positive multiplicative constant. If $a \ll b$ and $a \gg b$, we write $a \asymp b$ and say that the two quantities a and b are comparable.

The following is a simple consequence of Khintchine's Theorem.

Corollary 2.1. *Let* **Bad** *be the set of badly approximable numbers. Then*

$$m(\mathbf{Bad}) = 0\,.$$

Proof. Consider the function $\psi(q) = 1/(q \log q)$ and observe that

$$\mathbf{Bad} \cap \mathrm{I} \subset \mathbf{Bad}(\psi) := \mathrm{I} \setminus W(\psi)\,.$$

By Khintchine's Theorem, $m(W(\psi)) = 1$. Thus $m(\mathbf{Bad}(\psi)) = 0$ and so $m(\mathbf{Bad} \cap \mathrm{I}) = 0$. $\square$

2.2.1 The Duffin-Schaeffer Conjecture

The main substance of Khintchine's Theorem is the divergent case and it is where the assumption that ψ is monotonic is necessary. In 1941, Duffin & Schaeffer [48] constructed a non-monotonic approximating function ϑ for which the sum $\sum_q \vartheta(q)$ diverges but $m(W(\vartheta)) = 0$. We now discuss the construction. We start by recalling two well-known facts: for any $N \in \mathbb{N}$, p prime, and $s > 0$,

Fact 1. $\sum_{q|N} q = \prod_{p|N}(1+p)$

Fact 2. $\prod_p (1+p^{-s}) = \zeta(s)/\zeta(2s)$.

In view of Fact 2, we have that

$$\prod_p (1+p^{-1}) = \infty.$$

Thus, we can find a sequence of square free positive integers N_i $(i = 1, 2, \ldots)$ such that $(N_i, N_j) = 1$ $(i \neq j)$ and

$$\prod_{p|N_i} (1+p^{-1}) > 2^i + 1\,. \tag{2.7}$$

Now let

$$\vartheta(q) = \begin{cases} 2^{-i-1} q/N_i & \text{if} \quad q > 1 \quad \text{and} \quad q|N_i \quad \text{for some } i\,, \\ \\ 0 & \text{otherwise}\,. \end{cases} \tag{2.8}$$

As usual let

$$A_q := A_q(\vartheta) = \bigcup_{p=0}^{q} B\Big(\frac{p}{q}, \frac{\vartheta(q)}{q}\Big) \cap \mathrm{I}$$

and observe that if $q|N_i$ $(q > 1)$ then $A_q \subseteq A_{N_i}$ and so

$$\bigcup_{q|N_i} A_q = A_{N_i}\,.$$

In particular

$$m\Big(\bigcup_{q|N_i} A_q\Big) = m(A_{N_i}) = 2\vartheta(N_i) = 2^{-i}\,.$$

By definition
$$W(\vartheta) = \limsup_{q\to\infty} A_q = \limsup_{i\to\infty} A_{N_i} \,.$$
Now
$$\sum_{i=1}^{\infty} m(A_{N_i}) = 1$$
and so the convergence Borel-Cantelli Lemma implies that
$$m(W(\vartheta)) = 0\,.$$
However, it can be verified (*exercise*) on using Fact 1 together with (2.7) that
$$\sum_{q=1}^{\infty} \vartheta(q) = \sum_{i=1}^{\infty} 2^{-i-1} \frac{1}{N_i} \sum_{q>1\,:\,q|N_i} q \;\; = \infty\,.$$

In the same paper [48], Duffin and Schaeffer provided an appropriate statement for arbitrary ψ that we now discuss. The now famous Duffin-Schaeffer Conjecture represents a key open problem in number theory. The integer p implicit in the inequality (2.1) satisfies
$$\left| x - \frac{p}{q} \right| < \frac{\psi(q)}{q}\,. \tag{2.9}$$
To relate the rational p/q with the error of approximation $\psi(q)/q$ uniquely, we impose the coprimeness condition $(p,q) = 1$. In this case, let $W'(\psi)$ denote the set of x in I for which the inequality (2.9) holds for infinitely many $(p,q) \in \mathbb{Z} \times \mathbb{N}$ with $(p,q) = 1$. Clearly, $W'(\psi) \subset W(\psi)$. For any approximating function $\psi : \mathbb{N} \to \mathbb{R}^+$ one easily deduces that
$$m(W'(\psi)) = 0 \quad \text{if} \quad \sum_{q=1}^{\infty} \varphi(q)\, \frac{\psi(q)}{q} \; < \; \infty\,.$$
Here, and throughout, φ is the Euler function.

Conjecture 2.1 (Duffin-Schaeffer, 1941)**.** *For any function* $\psi\colon \mathbb{N} \to \mathbb{R}^+$
$$m(W'(\psi)) = 1 \quad \text{if} \quad \sum_{q=1}^{\infty} \varphi(q)\, \frac{\psi(q)}{q} \; = \; \infty\,.$$

Remark 2.6. Let ϑ be given by (2.8). On using the fact that $\sum_{d|n} \varphi(d) = n$, it is relatively easy to show (*exercise*) that
$$\sum_{q=1}^{\infty} \varphi(q)\, \frac{\vartheta(q)}{q} \; < \; \infty\,.$$
Thus, although ϑ provides a counterexample to Khintchine's Theorem without monotonicity, it is not a counterexample to the Duffin-Schaeffer Conjecture.

Remark 2.7. It is known that $m(W'(\psi)) = 0$ or 1. This is Gallagher's zero-full law [52] and is the natural analogue of Cassels' zero-full law for $W(\psi)$.

Although various partial results have been established (see [58, 90]), the full conjecture is one of the most difficult and profound unsolved problems in metric number theory. In the case where ψ is monotonic it is relatively straightforward to show that Khintchine's Theorem and the Duffin-Schaeffer Conjecture are equivalent statements (*exercise*).

2.3 A limitation of the Lebesgue theory

Let $\tau > 0$ and write $W(\tau)$ for $W(\psi : q \to q^{-\tau})$. The set $W(\tau)$ is usually referred to as the set of *τ-well approximable numbers.* Note that in view of Dirichlet (Theorem 1.2) we have that $W(\tau) = \mathrm{I}$ if $\tau \leq 1$ and so trivially $m(W(\tau)) = 1$ if $\tau \leq 1$. On the other hand, if $\tau > 1$

$$\textstyle\sum_{q=1}^{\infty} q^{-\tau} < \infty$$

and Khintchine's Theorem implies that $m(W(\tau)) = 0$. So for any $\tau > 1$, the set of τ-well approximable numbers is of measure zero. We cannot obtain any further information regarding the 'size' of $W(\tau)$ in terms of Lebesgue measure — it is always zero. Intuitively, the 'size' of $W(\tau)$ should decrease as rate of approximation governed by τ increases. For example we would expect that $W(2015)$ is "smaller" than $W(2)$ – clearly $W(2015) \subset W(2)$ but Lebesgue measure is unable to distinguish between them. In short, we require a more delicate notion of 'size' than simply Lebesgue measure. The appropriate notion of 'size' best suited for describing the finer measure theoretic structures of $W(\tau)$ and indeed $W(\psi)$ is that of Hausdorff measures.

3 Metric Diophantine approximation: the classical Hausdorff theory

3.1 Hausdorff measure and dimension

In what follows, a *dimension function* $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a left continuous, monotonic function such that $f(0) = 0$. Suppose F is a subset of $\mathbb{R}^n$. Given a ball B in $\mathbb{R}^n$, let $r(B)$ denote the radius of B. For $\rho > 0$, a countable collection $\{B_i\}$ of balls in $\mathbb{R}^n$ with $r(B_i) \leq \rho$ for each i such that $F \subset \bigcup_i B_i$ is called a *ρ-cover for F*. Define

$$\mathcal{H}^f_\rho(F) := \inf \sum_i f(r(B_i)),$$

where the infimum is taken over all ρ-covers of F. Observe that as ρ decreases the class of allowed ρ-covers of F is reduced and so $\mathcal{H}^f_\rho(F)$ increases. Therefore, the following (finite or infinite) limit exists

$$\mathcal{H}^f(F) := \lim_{\rho \to 0+} \mathcal{H}^f_\rho(F) \ = \ \sup_{\rho > 0} \mathcal{H}^f_\rho(F) \ ,$$

and is referred to as the *Hausdorff f–measure of F*. In the case that

$$f(r) = r^s \ (s \geq 0),$$

the measure $\mathcal{H}^f$ is the more common *s-dimensional Hausdorff measure* $\mathcal{H}^s$, the measure $\mathcal{H}^0$ being the cardinality of F. Note that when s is a positive integer, $\mathcal{H}^s$ is a constant multiple of Lebesgue measure in $\mathbb{R}^s$. (The constant is explicitly known!) Thus if the s-dimensional Hausdorff measure of a set is known for each $s > 0$, then so is its n-dimensional Lebesgue measure for each $n \geq 1$. The following easy property

$$\mathcal{H}^s(F) < \infty \quad \Longrightarrow \quad \mathcal{H}^{s'}(F) = 0 \qquad \text{if } s' > s$$

implies that there is a unique real point s at which the Hausdorff s-measure drops from infinity to zero (unless the set F is finite so that $\mathcal{H}^s(F)$ is never infinite). This point is called the *Hausdorff dimension* of F and is formally defined as

$$\dim F := \inf \{s > 0 : \mathcal{H}^s(F) = 0\} \ .$$

- By the definition of $\dim F$ we have that

$$\mathcal{H}^s(F) = \begin{cases} 0 & \text{if } s > \dim F \\ \\ \infty & \text{if } s < \dim F. \end{cases}$$

- If $s = \dim F$, then $\mathcal{H}^s(F)$ may be zero or infinite or may satisfy

$$0 < \mathcal{H}^s(F) < \infty;$$

  in this case F is said to be an s-set.

- Let $\mathrm{I} = [0, 1]$. Then $\dim \mathrm{I} = 1$ and

$$2\mathcal{H}^s(\mathrm{I}) = \begin{cases} 0 & \text{if } s > 1 \\ 1 & \text{if } s = 1 \\ \infty & \text{if } s < 1. \end{cases}$$

  Thus, $2\mathcal{H}^1(\mathrm{I}) = m(\mathrm{I})$ and I is an example of a s-set with $s = 1$. Note that the present of the factor '2' here is because in the definition of Hausdorff measure we have used the radii of balls rather than their diameters.

The Hausdorff dimension has been established for many number theoretic sets, e.g. $W(\tau)$ (this is the Jarník-Besicovitch Theorem discussed below), and is easier than determining the Hausdorff measure. Further details regarding Hausdorff measure and dimension can be found in [50, 72].

To calculate $\dim F$ (say $\dim F = \alpha$), it is usually the case that we establish the upper bound $\dim F \leq \alpha$ and lower bound $\dim F \geq \alpha$ separately. If we can exploit a 'natural' cover of F, then upper bounds are usually easier.

Example 3.1. Consider the middle third Cantor set K defined as follows: starting with $I_0 = [0, 1]$ remove the open middle thirds part of the interval. This gives the union of two intervals $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$. Then repeat the procedure of removing the middle third part from each of the intervals in your given collection. Thus, at 'level' n of the construction we will have the union E_n of 2^n closed intervals, each of length 3^{-n}. The middle third Cantor set is defined by

$$K = \bigcap_{n=0}^{\infty} E_n \, .$$

This set consists exactly of all real numbers such that their expansion to the base 3 does not contain the 'digit' 1.

Let $\{I_{n,j}\}$ be the collection of intervals in E_n. This is a collection of 2^n intervals, each of length 3^{-n}. Naturally, $\{I_{n,j}\}$ is a cover of K. Furthermore, for any $\rho > 0$ there is a sufficiently large n such that $\{I_{n,j}\}$ is a ρ-cover of K. It follows that,

$$\mathcal{H}^s_\rho(K) \leq \sum_j r(I_{n,j})^s \asymp 2^n 2^{-s} 3^{-ns} \ll \left(\frac{2}{3^s}\right)^n \to 0$$

as $n \to \infty$ (i.e. $\rho \to 0$) if

$$\frac{2}{3^s} < 1 \Rightarrow s > \frac{\log 2}{\log 3}.$$

In other words

$$\mathcal{H}^s(K) = 0 \text{ if } s > \frac{\log 2}{\log 3}.$$

It follows from the definition of Hausdorff dimension

$$\dim K = \inf\{s : \mathcal{H}^s(K) = 0\}$$

that $\dim K \leqslant \frac{\log 2}{\log 3}$.

In fact, $\dim K = \frac{\log 2}{\log 3}$. To prove that

$$\dim K \geqslant \frac{\log 2}{\log 3}$$

we need to work with arbitrary covers of K and this is much harder. Let $\{B_i\}$ be an arbitrary ρ-cover with $\rho < 1$. K is bounded and closed (intersection of closed intervals), i.e. K is compact. Hence without loss of generality we can assume that $\{B_i\}$ is finite. For each B_i, let r_i and d_i denote its radius and diameter respectively, and let k be the unique integer such that

$$3^{-(k+1)} \leqslant d_i < 3^{-k}. \tag{3.1}$$

Then B_i intersects at most one interval of E_k as the intervals in E_k are separated by at least 3^{-k}.

If $j \geqslant k$, then B_i intersects at most

$$2^{j-k} = 2^j 3^{-sk} \leqslant 2^j 3^s d_i^s \tag{3.2}$$

intervals of E_j, where $s := \frac{\log 2}{\log 3}$ and the final inequality makes use of (3.1). These are the intervals that are contained in the unique interval of E_k that intersects B_i.

Now choose j large enough so that

$$3^{-(j+1)} \leqslant d_i \quad \forall B_i \in \{B_i\}.$$

This is possible because the collection $\{B_i\}$ is finite. Then $j \geqslant k$ for each B_i and (3.2) is valid. Furthermore, since $\{B_i\}$ is a cover of K, it must intersect every interval of E_j. There are 2^j intervals in E_j. Thus

$$\begin{aligned} 2^j &= \#\{I \in E_j : \cup B_i \cap I \neq \varnothing\} \\ &\leq \sum_i \#\{I \in E_j : B_i \cap I \neq \varnothing\} \\ &\leq \sum_i 2^j 3^s d_i^s \,. \end{aligned}$$

The upshot is that for any arbitrary cover $\{B_i\}$, we have that

$$2^s \sum r_i^s \asymp \sum d_i^s \geq 3^{-s} = \frac{1}{2}\,.$$

By definition, this implies that implies $\mathcal{H}^s(K) \geq 2^{-(1+s)}$ and so $\dim K \geq \frac{\log 2}{\log 3}$.

Even for this simple Cantor set example, the lower bound for $\dim K$ is much more involved than the upper bound. This is usually the case and the number theoretic sets $W(\psi)$ and $W(\tau)$ are no exception.

3.2 The Jarník-Besicovitch Theorem

Recall, the limsup nature of $W(\psi)$; namely that

$$W(\psi) = \limsup_{q\to\infty} A_q(\psi) := \bigcap_{t=1}^{\infty} \bigcup_{q=t}^{\infty} A_q(\psi)$$

where

$$A_q(\psi) = \bigcup_{p=0}^{q} B\Big(\frac{p}{q}, \frac{\psi(q)}{q}\Big) \cap \mathrm{I}\,.$$

By definition, for each t, the collection of balls $B(p/q, \psi(q)/q)$ associated with the sets $A_q(\psi)$: $q = t, t+1, \ldots$ form a cover for $W(\psi)$. Suppose for the moment that ψ *is monotonic* and $\psi(q) < 1$ for q large. Now for any $\rho > 0$, choose t large enough so that $\rho > \psi(t)/t$. Then the balls in $\{A_q(\psi)\}_{q\geqslant t}$ form a ρ cover of $W(\psi)$. Thus,

$$\mathcal{H}^s_\rho\big(W(\psi)\big) \leq \sum_{q=t}^{\infty} q\big(\psi(q)/q\big)^s \to 0$$

as $t \to \infty$ (i.e. $\rho \to 0$) if

$$\sum_{q=1}^{\infty} q^{1-s}\psi^s(q) < \infty\,;$$

i.e. $\mathcal{H}^s\big(W(\psi)\big) = 0$ if the above s-volume sum converges. Actually, monotonicity on ψ can be removed (*exercise*) and we have proved the following Hausdorff measure analogue of Theorem 2.1. Recall, that $\mathcal{H}^1$ and one-dimensional Lebesgue measure m are comparable.

Theorem 3.1. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a function and $s \geq 0$ such that*

$$\sum_{q=1}^{\infty} q^{1-s}\psi^s(q) < \infty.$$

Then

$$\mathcal{H}^s\big(W(\psi)\big) = 0\,.$$

Now put $\psi(q) = q^{-\tau}$ $(\tau \geq 1)$ and notice that for $s > \frac{2}{\tau+1}$

$$\sum_{q=1}^{\infty} q^{1-s}\psi^s(q) = \sum_{q=1}^{\infty} q^{-(\tau s+s-1)} < \infty\,.$$

Then the following statement is a simple consequence of the above theorem and the definition of Hausdorff dimension.

Corollary 3.1. *For $\tau \geq 2$, we have that* $\dim W(\tau) \leq \frac{2}{\tau+1}$.

Note that the above convergence result and thus the upper bound dimension result, simply exploit the natural cover associated with the limsup set under consideration. The corollary constitutes the easy part of the famous Jarník-Besicovitch Theorem.

Theorem 3.2 (The Jarník-Besicovitch Theorem)**.** *Let $\tau > 1$. Then*

$$\dim\big(W(\tau)\big) = 2/(\tau+1)\,.$$

Jarník proved the result in 1928. Besicovitch proved the same result in 1932 by completely different methods. The Jarník-Besicovitch Theorem implies that

$$\dim W(2) = 2/3 \quad \text{and} \quad \dim W(2015) = 2/2016$$

and so $W(2015)$ is "smaller" than $W(2)$ as expected. In view of Corollary 3.1, we need to establish the lower bound result $\dim\big(W(\tau)\big) \geq 2/(\tau+1)$ in order to complete the proof of Theorem 3.2. We will see that this is a consequence of Jarník's measure result discussed in the next section.

The dimension theorem is clearly an excellent result but it gives no information regarding $\mathcal{H}^s$ at the critical exponent $d := 2/(\tau+1)$. By definition

$$\mathcal{H}^s(W(\tau)) = \begin{cases} 0 & \text{if } s > d \\ \infty & \text{if } s < d \end{cases}$$

but

$$\mathcal{H}^s(W(\tau)) = \textbf{?} \quad \text{if } s = d\,.$$

In short, it would be highly desirable to have a Hausdorff measure analogue of Khintchine's Theorem.

3.3 Jarník's Theorem

Theorem 3.1 is the easy case of the following fundamental statement in metric Diophantine approximation. It provides an elegant criterion for the 'size' of the set $W(\psi)$ expressed in terms of Hausdorff measure.

Theorem 3.3 (Jarník's Theorem, 1931)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $s \in (0,1)$. Then*

$$\mathcal{H}^s\big(W(\psi)\big) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q^{1-s}\psi^s(q) < \infty \\ \\ \infty & \text{if } \sum_{q=1}^{\infty} q^{1-s}\psi^s(q) = \infty \end{cases}$$

Remark 3.1. With $\psi(q) = q^{-\tau}$ $(\tau > 1)$, not only does the above theorem imply that $\dim W(\tau) = 2/(1+\tau)$ but it tells us that the Hausdorff measure at the critical exponent is infinite; i.e.

$$\mathcal{H}^s\big(W(\tau)\big) = \infty \quad \text{at} \quad s = 2/(1+\tau)\,.$$

Remark 3.2. As in Khintchine's Theorem, the assumption that ψ is monotonic is only required in the divergent case. In Jarník's original statement, apart from assuming stronger monotonicity conditions, various technical conditions on ψ and indirectly s were imposed, which prevented $s = 1$. Note that even as stated, it is natural to exclude the case $s = 1$ since

$$\mathcal{H}^1\big(W(\psi)\big) \asymp m\big(W(\psi)\big) = 1\,.$$

The clear cut statement without the technical conditions was established in [14] and it allows us to combine the theorems of Khintchine and Jarník into a unifying statement.

Theorem 3.4 (Khintchine-Jarník 2006)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $s \in (0,1]$. Then*

$$\mathcal{H}^s\big(W(\psi)\big) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q^{1-s}\psi^s(q) < \infty\,, \\ \\ \mathcal{H}^s(\mathrm{I}) & \text{if } \sum_{q=1}^{\infty} q^{1-s}\psi^s(q) = \infty\,. \end{cases}$$

Obviously, the Khintchine-Jarník Theorem implies Khintchine's Theorem.

In view of the Mass Transference Principle established in [21] one actually has that

$$\text{Khintchine's Theorem} \quad \Longrightarrow \quad \text{Jarník's Theorem.}$$

Thus, the Lebesgue theory of $W(\psi)$ underpins the general Hausdorff theory. At first glance this is rather surprising because the Hausdorff theory had previously been thought to be a subtle refinement of the Lebesgue theory. Nevertheless, the Mass Transference Principle allows us to transfer Lebesgue measure theoretic statements for limsup sets to Hausdorff statements and naturally obtain a complete metric theory.

3.4 The Mass Transference Principle

Let (Ω, d) be a locally compact metric space and suppose there exist constants $\delta > 0$, $0 < c_1 < 1 < c_2 < \infty$ and $r_0 > 0$ such that

$$c_1\, r^\delta \leq \mathcal{H}^\delta(B) \leq c_2\, r^\delta \;, \tag{3.3}$$

for any ball $B = B(x, r)$ with $x \in \Omega$ and radius $r \leq r_0$. For the sake of simplicity, the definition of Hausdorff measure and dimension given in §3.1 is restricted to $\mathbb{R}^n$. Clearly, it can easily be adapted to the setting of arbitrary metric spaces – see [50, 72]. A consequence of (3.3) is that

$$0 < \mathcal{H}^\delta(\Omega) < \infty \quad \text{and} \quad \dim \Omega = \delta \,.$$

Next, given a dimension function f and a ball $B = B(x, r)$ we define the scaled ball

$$B^f := B\big(x, f(r)^{\frac{1}{\delta}}\big) \,.$$

When $f(r) = r^s$ for some $s > 0$, we adopt the notation B^s, i.e.

$$B^s := B\big(x, r^{\frac{s}{\delta}}\big)$$

and so by definition $B^\delta = B$.

The Mass Transference Principle [21] allows us to transfer $\mathcal{H}^\delta$-measure theoretic statements for limsup subsets of Ω to general $\mathcal{H}^f$-measure theoretic statements. Note that in the case $\delta = k \in \mathbb{N}$, the measure $\mathcal{H}^\delta$ coincides with k-dimensional Lebesgue measure and the Mass Transference Principle allows us to transfer Lebesgue measure theoretic statements for limsup subsets of $\mathbb{R}^k$ to Hausdorff measure theoretic statements.

Theorem 3.5. *Let $\{B_i\}_{i\in\mathbb{N}}$ be a sequence of balls in Ω with $r(B_i) \to 0$ as $i \to \infty$. Let f be a dimension function such that $x^{-\delta} f(x)$ is monotonic. For any ball $B \in \Omega$ with $\mathcal{H}^\delta(B) > 0$, if*

$$\mathcal{H}^\delta\big(B \cap \limsup_{i\to\infty} B_i^f\,\big) = \mathcal{H}^\delta(B)$$

then

$$\mathcal{H}^f\big(B \cap \limsup_{i\to\infty} B_i^\delta\,\big) = \mathcal{H}^f(B) \;.$$

Remark 3.3. There is one point that is well worth making. The Mass Transference Principle is purely a statement concerning limsup sets arising from a sequence of balls. There is absolutely no monotonicity assumption on the radii of the balls. Even the imposed condition that $r(B_i) \to 0$ as $i \to \infty$ is redundant but is included to avoid unnecessary tedious discussion.

3.4.1 Khintchine's Theorem implies Jarník's Theorem

First of all let us dispose of the case that $\psi(r)/r \nrightarrow 0$ as $r \to \infty$. Then trivially, $W(\psi) = \mathrm{I}$ and the result is obvious. Without loss of generality, assume that $\psi(r)/r \to 0$ as $r \to \infty$. With respect to the Mass Transference Principle, let $\Omega = \mathrm{I}$, d be the supremum norm, $\delta = 1$ and $f(r) = r^s$ with $s \in (0, 1)$. We are given that $\sum q^{1-s}\psi(q)^s = \infty$. Let $\theta(r) := q^{1-s}\psi(q)^s$. Then θ is an approximating function and $\sum \theta(q) = \infty$. Thus, Khintchine's Theorem implies that $\mathcal{H}^1(B \cap W(\theta)) = \mathcal{H}^1(B \cap \mathrm{I})$ for any ball B in $\mathbb{R}$. It now follows via the Mass Transference Principle that $\mathcal{H}^s(W(\psi)) = \mathcal{H}^s(\mathrm{I}) = \infty$ and this completes the proof of the divergence part of Jarník's Theorem. As we have already seen, the convergence part is straightforward.

3.4.2 Dirichlet's Theorem implies the Jarník-Besicovitch Theorem

Dirichlet's theorem (Theorem 1.2) states that for any irrational $x \in \mathbb{R}$, there exist infinitely many reduced rationals p/q ($q > 0$) such that $|x - p/q| \le q^{-2}$; i.e. $W(1) = \mathrm{I}$. Thus, with $f(r) := r^d$ ($d := 2/(1+\tau)$) the Mass Transference Principle implies that $\mathcal{H}^d(W(\tau)) = \infty$. Hence $\dim W(\tau) \ge d$. The upper bound is trivial. Note that we have actually proved a lot more than the Jarník-Besicovitch theorem. We have proved that the s–dimensional Hausdorff measure $\mathcal{H}^s$ of $W(\tau)$ at the critical exponent $s = d$ is infinite.

3.5 The Generalised Duffin-Schaeffer Conjecture

As with Khintchine's Theorem, it is natural to seek an appropriate statement in which one removes the monotonicity condition in Jarník's Theorem. In the case of Khintchine's Theorem, the appropriate statement is the Duffin-Schaeffer Conjecture – see §2.2.1. With this in mind, we work with the set $W'(\psi)$ in which the coprimeness condition $(p,q) = 1$ is imposed on the rational approximates p/q. For any function $\psi : \mathbb{N} \to \mathbb{R}^+$ and $s \in (0,1]$ it is easily verified that

$$\mathcal{H}^s\big(W(\psi)\big) = 0 \quad \text{if} \quad \sum_{q=1}^{\infty} \varphi(q) \left(\frac{\psi(q)}{q}\right)^s \; < \; \infty \,.$$

In the case the above s-volume sum diverges it is reasonable to believe in the truth of the following Hausdorff measure version of the Duffin-Schaeffer Conjecture [21].

Conjecture 3.1 (Generalised Duffin-Schaeffer Conjecture, 2006)**.** *For any function $\psi\colon \mathbb{N} \to \mathbb{R}^+$ and $s \in (0,1]$*

$$\mathcal{H}^s\big(W'(\psi)\big) = \mathcal{H}^s\big(\mathrm{I}\big) \quad \text{if} \quad \sum_{q=1}^{\infty} \varphi(q) \left(\frac{\psi(q)}{q}\right)^s \; = \; \infty \,.$$

Remark 3.4. If $s = 1$, then $\mathcal{H}^1(\mathrm{I}) = m(\mathrm{I})$ and Conjecture 3.1 reduces to the Lebesgue measure conjecture of Duffin & Schaeffer (Conjecture 2.1).

Remark 3.5. In view of the Mass Transference Principle, it follows that

$$\text{Conjecture 2.1} \quad \Longrightarrow \quad \text{Conjecture 3.1}$$

Exercise: Prove the above implication.

4 The higher dimensional theory

We start with a generalisation of Theorem 1.1 to simultaneous approximation in $\mathbb{R}^n$.

Theorem 4.1 (Dirichlet in $\mathbb{R}^n$)**.** *Let $(i_1, \dots, i_n)$ be any n-tuple of numbers satisfying*

$$0 < i_1, \dots, i_n < 1 \quad \text{and} \quad \sum_{t=1}^{n} i_t = 1\,. \tag{4.1}$$

Then, for any $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ *and* $N \in \mathbb{N}$, *there exists* $q \in \mathbb{Z}$ *such that*

$$\max\{\|qx_1\|^{1/i_1}, \ldots, \ \|qx_n\|^{1/i_n}\} < N^{-1} \qquad \textit{and} \qquad 1 \le q \le N\,. \tag{4.2}$$

Remark 4.1. The symmetric case corresponding to $i_1 = \ldots = i_n = 1/n$ is the more familiar form of the theorem. In this symmetric case, when N is an n'th power, the one-dimensional proof using the pigeon-hole principle can be easily adapted to prove the associated statement (*exercise*). The above general form is a neat consequence of a fundamental theorem in the geometry of numbers; namely Minkowski's theorem for systems of linear forms – see §4.1 below. At this point simply observe that for a fixed q the first inequality in (4.2) corresponds to considering rectangles centered at rational points

$$\Big(\frac{p_1}{q}, \ldots, \frac{p_n}{q}\Big) \quad \text{of sidelength} \quad \frac{2}{qN^{i_1}}, \ldots, \frac{2}{qN^{i_n}} \quad \text{respectively}\,.$$

Now the shape of the rectangles are clearly governed by $(i_1, \ldots, i_n)$. However the volume is not. Indeed, for any $(i_1, \ldots, i_n)$ satisfying (4.1), the n-dimensional Lebesgue measure m_n of any rectangle centered at a rational point with denominator q is $2^n q^{-n} N^{-1}$.

4.1 Minkowski's Linear Forms Theorem

We begin by introducing various terminology and establishing Minkowski's theorem for convex bodies.

Definition 4.1. A subset B of $\mathbb{R}^n$ is said to be *convex* if for any two points $\mathbf{x}, \mathbf{y} \in B$

$$\big\{\lambda\mathbf{x} + (1-\lambda)\mathbf{y} : 0 \le \lambda \le 1\big\} \subset B\,,$$

that is the line segment joining $\mathbf{x}$ and $\mathbf{y}$ is contained in B. A *convex body* in $\mathbb{R}^n$ is a bounded convex set.

Definition 4.2. A subset B in $\mathbb{R}^n$ is said to be *symmetric about the origin* if for every $\mathbf{x} \in B$ we have that $-\mathbf{x} \in B$.

The following is a simple but nevertheless powerful observation concerning symmetric convex bodies.

Theorem 4.2 (Minkowski's Convex Body Theorem)**.** *Let B be a convex body in $\mathbb{R}^n$ symmetric about the origin. If* $\operatorname{vol}(B) > 2^n$ *then B contains a non-zero integer point.*

Proof. The following proof is attributed to Mordell. For $m \in \mathbb{N}$ let $A(m, B) = \{\mathbf{a} \in \mathbb{Z}^m : \mathbf{a}/m \in B\}$. Then we have that

$$\lim_{m\to\infty} m^{-n} \# A(m, B) = \operatorname{vol}(B)\,.$$

Since $\operatorname{vol}(B) > 2^n$, there is a sufficiently large m such that $m^{-n}\#A(m, B) > 2^n$, that is $\#A(m, B) > (2m)^n$. Since there are $2m$ different residue classes modulo $2m$ and each point in $A(Q,m)$ has n coordinates, there are two distinct points in $A(Q,m)$, say $\mathbf{a}=(a_1,\dots,a_n)$ and $\mathbf{b}=(b_1,\dots,b_n)$ such that

$$a_i \equiv b_i \pmod{2m} \quad \text{for each } i=1,\dots,n\,.$$

Hence

$$\mathbf{z}=\frac{1}{2}\frac{\mathbf{a}}{m}+\frac{1}{2}\left(-\frac{\mathbf{b}}{m}\right)=\frac{\mathbf{a}-\mathbf{b}}{2m}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}\,.$$

Since B is symmetric about the origin, $-\mathbf{b}/m\in B$ and since B is convex $\mathbf{z}\in B$. The proof is complete. $\square$

The above convex body result enables us to prove the following extremely useful statement.

Theorem 4.3 (Minkowski's theorem for systems of linear forms)**.** *Let $\beta_{i,j}\in\mathbb{R}$, where $1\le i,j\le k$, and let $C_1,\dots,C_k>0$. If*

$$|\det(\beta_{i,j})_{1\le i,j\le k}|\le\prod_{i=1}^{k}C_i, \tag{4.3}$$

then there exists a non-zero integer point $\mathbf{x}=(x_1,\dots,x_k)$ such that

$$\begin{cases} |x_1\beta_{i,1}+\dots+x_k\beta_{i,k}|<C_i & (1\le i\le k-1)\\ |x_1\beta_{k,1}+\dots+x_n\beta_{k,k}|\le C_k \end{cases} \tag{4.4}$$

Proof. The set of $(x_1,\dots,x_k)\in\mathbb{R}^k$ satisfying (4.4) is a convex body symmetric about the origin. First consider the case when $\det(\beta_{i,j})_{1\le i,j\le k}\ne 0$ and (4.3) is strict. Then

$$\operatorname{vol}(B)=\frac{\prod_{i=1}^{k}(2C_i)}{|\det(\beta_{i,j})_{1\le i,j\le k}|}>2^n\,.$$

Then, by Theorem 4.2, the body contains a non-zero integer point $(x_1,\dots,x_k)$ as required.

If $\det(\beta_{i,j})_{1\le i,j\le k}=0$ then B is unbounded and has infinite volume. Then there exists a sufficiently large $m\in\mathbb{N}$ such that $B_m=B\cap[-m,m]$ has volume $\operatorname{vol}(B_m)>2^n$. Next, B_m is convex and symmetric about the origin, since it is the intersection of 2 sets with these properties. Again, by Theorem 4.2, B_m contains a non-zero integer point $(x_1,\dots,x_k)$. Since $B_m\subset B$ we again get the required statement.

Finally, consider the situation when (4.3) is an equation. In this case $\det(\beta_{i,j})_{1\le i,j\le k}\ne 0$. Define $C_k^\varepsilon=C_k+\varepsilon$ for some $\varepsilon>0$. Then

$$|\det(\beta_{i,j})_{1\le i,j\le k}|<\prod_{i=1}^{k-1}C_i\times C_k^\varepsilon \tag{4.5}$$

and by what we have already shown there exists a non-zero integer solution $\mathbf{x}_\varepsilon=(x_1,\dots,x_k)$ to the system

$$\begin{cases} |x_1\beta_{i,1}+\dots+x_k\beta_{i,k}|<C_i & (1\le i\le k-1)\\ |x_1\beta_{k,1}+\dots+x_n\beta_{k,k}|\le C_k^\varepsilon\,. \end{cases} \tag{4.6}$$

For $\varepsilon \le 1$ all the points $\mathbf{x}_\varepsilon$ satisfy (4.6) with $\varepsilon = 1$. That is they lie in a bounded body. Hence, there are only finitely many of them. Therefore there is a sequence ε_i tending to 0 such that $\mathbf{x}_{\varepsilon_i}$ are all the same, say $\mathbf{x}_0$. On letting $i \to \infty$ within (4.6) we get that (4.4) holds with $\mathbf{x} = \mathbf{x}_0$. $\square$

It is easily verified that Theorem 4.1 (Dirichlet in $\mathbb{R}^n$) is an immediate consequence of Theorem 4.3 with $k = n + 1$ and

$$C_t = N^{-i_t} \quad (1 \le t \le k-1) \quad \text{and} \qquad C_k = N$$

and

$$(\beta_{i,j}) = \begin{pmatrix} -1 & 0 & 0 & \dots & \alpha_1 \\ 0 & -1 & 0 & \dots & \alpha_2 \\ 0 & 0 & -1 & \dots & \\ \vdots & & & \ddots & \alpha_n \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Another elegant application of Theorem 4.3 is the following statement.

Corollary 4.1. *For any $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ and any real $N > 1$, there exist $q_1, \dots, q_n, p \in \mathbb{Z}$ such that*

$$|q_1\alpha_1 + \dots + q_n\alpha - p| < N^{-n} \qquad \text{and} \qquad 1 \le \max_{1 \le i \le n} |q_i| \le N\,.$$

In particular, there exist infinitely many $((q_1, \dots, q_n), p) \in \mathbb{Z}^n \setminus \{0\} \times \mathbb{Z}$ such that

$$|q_1\alpha_1 + \dots + q_n\alpha - p| < \Big(\max_{1 \le i \le n} |q_i|\Big)^{-n}.$$

Proof. Exercise $\square$

4.2 Bad in $\mathbb{R}^n$

An important consequence of Dirichlet's theorem (Theorem 4.1) is the following higher dimensional analogue of Theorem 1.2.

Theorem 4.4. *Let $(i_1, \dots, i_n)$ be any n-tuple of real numbers satisfying (4.1). Let $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$. Then there exist infinitely many integers $q > 0$ such that*

$$\max\{\|qx_1\|^{1/i_1}, \dots, \ \|qx_n\|^{1/i_n}\} < q^{-1}\,. \tag{4.7}$$

Now just as in the one-dimensional setup we can ask the following natural question.

Question. Can we replace the right-hand side of (4.7) by ϵq^{-1} where $\epsilon > 0$ is arbitrary?

> **No.** For any $(i_1, \dots, i_n)$ satisfying (4.1), there exists $(i_1, \dots, i_n)$-*badly approximable points.*

Denote by $\mathbf{Bad}(i_1,\ldots,i_n)$ the set of $(i_1,\ldots,i_n)$-badly approximable points; that is the set of $(x_1,\ldots,x_n)\in\mathbb{R}^n$ such that there exists a positive constant $c(x_1,\ldots,x_n)>0$ so that

$$\max\{\|qx_1\|^{1/i_1},\ldots,\ \|qx_n\|^{1/i_n}\}>c(x_1,\ldots,x_n)\ q^{-1}\quad\forall q\in\mathbb{N}\ .$$

Remark 4.2. Let $n=2$ and note that if $(x,y)\in\mathbf{Bad}(i,j)$ for some pair (i,j), then it would imply that

$$\liminf_{q\to\infty} q\|qx\|\|qy\|=0.$$

Hence $\cap_{i+j=1}\mathbf{Bad}(i,j)=\varnothing$ would imply that Littlewood's Conjecture is true. We will return to this famous conjecture in §4.4.

Remark 4.3. Geometrically speaking, $\mathbf{Bad}(i_1,\ldots,i_n)$ consists of points $\mathbf{x}\in\mathbb{R}^n$ that avoid all rectangles of size $c^{i_1}q^{-(1+i_1)}\times\ldots\times c^{i_n}q^{-(1+i_n)}$ centred at rational points $(p_1/q,\ldots,p_n/q)$ with $c=c(\mathbf{x})$ sufficiently small. Note that in the symmetric case $i_1=\ldots=i_n=1/n$, the rectangles are squares (or essentially balls) and this makes a profound difference when investigating the 'size' of $\mathbf{Bad}(i_1,\ldots,i_n)$ – it makes life significantly easier!

Perron [74] in 1921 observed that $(x,y)\in\mathbf{Bad}(\frac{1}{2},\frac{1}{2})$ whenever x and y are linearly independent numbers in a cubic field; e.g $x=\cos\frac{2\pi}{7}, y=\cos\frac{4\pi}{7}$. Thus, certainly $\mathbf{Bad}(\frac{1}{2},\frac{1}{2})$ is not the empty set. It was shown by Davenport in 1954 that $\mathbf{Bad}(\frac{1}{2},\frac{1}{2})$ is uncountable and later in [42] he gave a simple and more illuminating proof of this fact. Furthermore, the ideas in his 1964 paper show that $\mathbf{Bad}(i_1,\ldots,i_n)$ is uncountable. In 1966, Schmidt [85] showed that in the symmetric case the corresponding set $\mathbf{Bad}(\frac{1}{n},\ldots,\frac{1}{n})$ is of full Hausdorff dimension. In fact, Schmidt proved the significantly stronger statement that the symmetric set is winning in the sense of his now famous (α,β)-games (see §7.2 below). Almost forty years later it was proved in [77] that

$$\dim\mathbf{Bad}(i_1,\ldots,i_n)=n\,.$$

Now let us return to the symmetric case of Theorem 4.4. It implies that every point $\mathbf{x}=(x_1,\ldots,x_n)\in\mathbb{R}^n$ can be approximated by rational points $(p_1/q,\ldots p_n/q)$ with rate of approximation given by $q^{-(1+\frac{1}{n})}$. The above discussion shows that this rate of approximation cannot in general be improved by an arbitrary constant—$\mathbf{Bad}(\frac{1}{n},\ldots,\frac{1}{n})$ is non-empty. However, if we exclude a set of real numbers of measure zero, then from a measure theoretic point of view the rate of approximation can be improved, just as in the one-dimensional setup.

4.3 Higher dimensional Khintchine

Let $\mathrm{I}^n:=[0,1)^n$ denote the unit cube in $\mathbb{R}^n$ and for $\mathbf{x}=(x_1,\ldots,x_n)\in\mathbb{R}^n$ let

$$\|q\mathbf{x}\|:=\max_{1\le i\le n}\|qx_i\|\,.$$

Given $\psi:\mathbb{N}\to\mathbb{R}^+$, let

$$W(n,\psi):=\{\mathbf{x}\in\mathrm{I}^n\colon\|q\mathbf{x}\|<\psi(q)\text{ for infinitely many }q\in\mathbb{N}\}$$

denote the set of *simultaneously ψ-well approximable* points $\mathbf{x}\in\mathrm{I}^n$. Thus, a point $\mathbf{x}\in\mathrm{I}^n$ is ψ-well approximable if there exist infinitely many rational points

$$\Big(\frac{p_1}{q},\ldots,\frac{p_n}{q}\Big)$$

with $q > 0$, such that the inequalities

$$\left|x_i - \frac{p_i}{q}\right| < \frac{\psi(q)}{q}$$

are simultaneously satisfied for $1 \leq i \leq n$. For the same reason as in the $n = 1$ case there is no loss of generality in restricting our attention to the unit cube. In the case $\psi : q \to q^{-\tau}$ with $\tau > 0$, we write $W(n, \tau)$ for $W(n, \psi)$. The set $W(n, \tau)$ is the set of *simultaneously τ-well approximable numbers.* Note that in view of Theorem 4.4 we have that

$$W(n, \tau) = \mathrm{I}^n \quad \text{if} \quad \tau \leq \frac{1}{n}. \tag{4.8}$$

The following is the higher dimensional generalisation of Theorem 2.3 to simultaneous approximation. Throughout, m_n will denote n-dimensional Lebesgue measure.

Theorem 4.5 (Khintchine's Theorem in $\mathbb{R}^n$)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function. Then*

$$m_n(W(n, \psi)) = \begin{cases} 0 & \text{if} \quad \sum_{q=1}^{\infty} \psi^n(q) < \infty \,, \\ \\ 1 & \text{if} \quad \sum_{q=1}^{\infty} \psi^n(q) = \infty \,. \end{cases}$$

Remark 4.4. The convergent case is a straightforward consequence of the Convergence Borel-Cantelli Lemma and does not require monotonicity.

Remark 4.5. The divergent case is the main substance of the theorem. When $n \geq 2$, a consequence of a theorem of Gallagher [54] is that the monotonicity condition can be dropped. Recall, that in view of the Duffin-Schaeffer counterexample (see §2.2.1) the monotonicity condition is crucial when $n = 1$.

Remark 4.6. Theorem 4.5 implies that

$$m_n(W(n, \psi)) = 1 \quad \text{if} \quad \psi(q) = 1/(q \log q)^{\frac{1}{n}} \,.$$

Thus, from a measure theoretic point of view the 'rate' of approximation given by Theorem 4.4 can be improved by $(\text{logarithm})^{\frac{1}{n}}$.

Remark 4.7. Theorem 4.5 implies that $m_n(\mathbf{Bad}(\frac{1}{n}, \ldots, \frac{1}{n})) = 0$.

Remark 4.8. For a generalisation of Theorem 4.5 to Hausdorff measures—that is, the higher dimension analogue of Theorem 3.4 (Khintchine-Jarník Theorem))—see Theorem 4.12 with $m = 1$ in §4.6. Also, see §5.3.1.

In view of Remark 4.5, one may think that there is nothing more to say regarding the Lebesgue theory of ψ-well approximable points in $\mathbb{R}^n$. After all, for $n \geq 2$ we do not even require monotonicity in Theorem 4.5. For ease of discussion let us restrict our attention to the plane $\mathbb{R}^2$ and assume that the n-volume sum in Theorem 4.5 diverges. So we know that almost all points (x_1, x_2) are ψ-well approximable but it tells us nothing for a given fixed x_1. For example, are there any points $(\sqrt{2}, x_2) \in \mathbb{R}^2$ that are ψ-well approximable? This will be discussed in §4.5 and the more general question of approximating points on a manifold will be the subject of §6.

4.4 Multiplicative approximation: Littlewood's Conjecture

For any pair of real numbers $(\alpha, \beta) \in \mathrm{I}^2$, there exist infinitely many $q \in \mathbb{N}$ such that

$$\|q\alpha\|\,\|q\beta\| \leq q^{-1}\,.$$

This is a simple consequence of Theorem 4.4 or indeed the one-dimensional Dirichlet theorem and the trivial fact that $\|x\| < 1$ for any x. For any arbitrary $\epsilon > 0$, the problem of whether or not the statement remains true by replacing the right-hand side of the inequality by $\epsilon\, q^{-1}$ now arises. This is precisely the content of Littlewood's conjecture.

Littlewood's Conjecture . *For any pair* $(\alpha, \beta) \in \mathrm{I}^2$*,*

$$\liminf_{q\to\infty} q\,||q\alpha||\,||q\beta|| = 0\,.$$

Equivalently, for any pair $(\alpha, \beta) \in \mathrm{I}^2$ there exist infinitely many rational points $(p_1/q, p_2/q)$ such that

$$\left|\alpha - \frac{p_1}{q}\right| \left|\beta - \frac{p_2}{q}\right| \;<\; \frac{\epsilon}{q^3} \quad (\epsilon > 0 \;\text{ arbitrary})\,.$$

Thus geometrically, the conjecture states that every point in the (x, y)-plane lies in infinitely many hyperbolic regions given by $|x| \cdot |y| < \epsilon/q^3$ centred at rational points.

The analogous conjecture in the one-dimensional setting is false—Hurwitz's theorem tells us that the set **Bad** is nonempty. However, in the multiplicative situation the problem is still open.

We make various simple observations:

(i) *The conjecture is true for pairs* (α, β) *when either* α *or* β *are not in* **Bad**. Suppose $\beta \notin$ **Bad** and consider its convergents p_n/q_n. It follows from the right-hand side of inequality (1.8) that $q_n||q_n\alpha||\,||q_n\beta|| \leq 1/a_{n+1}$ for all n. Since β is not badly approximable the partial quotients a_i are unbounded and the conjecture follows. Alternatively, by definition if $\beta \notin$ **Bad**, then $\liminf_{q\to\infty} q\,||q\beta|| = 0$ and we are done. See also Remark 4.2.

(ii) *The conjecture is true for pairs* (α, β) *when either* α *or* β *lie in a set of full Lebesgue measure.* This follows at once from Khintchine's theorem. In fact, one has that for all α and almost all $\beta \in \mathrm{I}$,

$$q\,\log q\,\|q\alpha\|\,\|q\beta\| \leq 1 \quad \text{for infinitely many } q \in \mathbb{N} \tag{4.9}$$

or even

$$\liminf_{q\to\infty} q\,\log q\,||q\alpha||\,||q\beta|| = 0\,.$$

We now turn our attention to 'deeper' results regarding Littlewood.

Theorem (Cassels & Swinnerton-Dyer, 1955). *If* α, β *are both cubic irrationals in the same cubic field then Littlewood's Conjecture is true.*

This was subsequently strengthened by Peck [73].

Theorem (Peck, 1961). *If α, β are both cubic irrationals in the same cubic field then (α, β) satisfy (4.9) with the constant 1 on the right hand side replaced by a positive constant dependent on α and β.*

In view of (ii) above, when dealing with Littlewood we can assume without loss of generality that both α and β are in **Bad**. As mentioned in Chapter 1, it is conjectured (the Folklore Conjecture) that the only algebraic irrationals which are badly approximable are the quadratic irrationals. Of course, if this conjecture is true then the Cassels & Swinnerton–Dyer result follows immediately. On restricting our attention to just badly approximable pairs we have the following statement [76].

Theorem PV (2000). *Given $\alpha \in$ **Bad** we have that*

$$\dim\big(\{\beta \in \mathbf{Bad} : (\alpha, \beta) \text{ satisfy (4.9)}\}\big) = 1 .$$

Regarding, potential counterexamples to Littlewood we have the following elegant statement [49].

Theorem EKL (2006). $\dim\big(\{(\alpha, \beta) \in \mathrm{I}^2 : \liminf_{q\to\infty} q\,||q\alpha||\,||q\beta|| > 0\}\big) = 0.$

Now let us turn our attention to non-trivial, purely metrical statements regarding Littlewood. The following result due to Gallagher [53] is the analogue of Khintchine's simultaneous approximation theorem (Theorem 4.5) within the multiplicative setup. Given $\psi : \mathbb{N} \to \mathbb{R}^+$ let

$$W^{\times}(n, \psi) := \{\mathbf{x} \in \mathrm{I}^n \colon \|qx_1\| \, \ldots \, \|qx_n\| < \psi(q) \text{ for infinitely many } q \in \mathbb{N}\} \tag{4.10}$$

denote the set of multiplicative ψ-well approximable points $\mathbf{x} \in \mathrm{I}^n$.

Theorem 4.6 (Gallagher, 1962)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function. Then*

$$m_n(W^{\times}(n, \psi)) = \begin{cases} 0 & \text{if} \quad \sum_{q=1}^{\infty} \psi(q) \log^{n-1} q < \infty \, , \\ \\ 1 & \text{if} \quad \sum_{q=1}^{\infty} \psi(q) \log^{n-1} q = \infty \, . \end{cases}$$

Remark 4.9. In the case of convergence, we can remove the condition that ψ is monotonic if we replace the above convergence condition by $\sum \psi(q) \, | \log \psi(q)|^{n-1} < \infty$; see [16] for more details.

An immediate consequence of Gallagher's Theorem is that almost all (α, β) beat Littlewood's Conjecture by 'log squared'; equivalently, almost surely Littlewood's Conjecture is true with a 'log squared' factor to spare.

Corollary 4.2. *For almost all $(\alpha, \beta) \in \mathbb{R}^2$*

$$\liminf_{q\to\infty} q \, \log^2 q \, ||q\alpha|| \, ||q\beta|| = 0 \, . \tag{4.11}$$

Recall, that this is beyond the scope of what Khintchine's theorem can tell us; namely that

$$\liminf_{q\to\infty} q\,\log q\,||q\alpha||\,||q\beta|| = 0 \quad \forall\,\alpha\in\mathbb{R} \quad \text{and} \quad \text{for almost all } \beta\in\mathbb{R}\,. \tag{4.12}$$

However the extra log factor in the corollary comes at a cost of having to sacrifice a set of measure zero on the α side. As a consequence, unlike with (4.12) which is valid for any α, we are unable to claim that the stronger 'log squared' statement (4.11) is true for say when $\alpha = \sqrt{2}$. Obviously, the role of α and β in (4.12) can be reversed. This raises the natural question of whether (4.11) holds for every α. If true, it would mean that for any α we still beat Littlewood's Conjecture by 'log squared' for almost all β.

4.4.1 Gallagher on fibers

The following result is established in [17].

Theorem 4.7. *Let $\alpha\in\mathrm{I}$ and $\psi:\mathbb{N}\to\mathbb{R}^+$ be a monotonic function such that*

$$\sum_{q=1}^{\infty} \psi(q)\,\log q \;=\; \infty \tag{4.13}$$

and such that

$$\exists\,\delta>0 \qquad \liminf_{n\to\infty} q_n^{3-\delta}\psi(q_n) \ge 1\,, \tag{4.14}$$

where q_n denotes the denominators of the convergents of α. Then for almost every $\beta\in\mathrm{I}$, there exists infinitely many $q\in\mathbb{N}$ such that

$$\|q\alpha\|\,\|q\beta\| < \psi(q)\,. \tag{4.15}$$

Remark 4.10. Condition (4.14) is not particularly restrictive. It holds for all α with Diophantine exponent $\tau(\alpha)<3$. By definition,

$$\tau(x) = \sup\{\tau>0 : \|q\alpha\| < q^{-\tau} \quad \text{for infinitely many } q\in\mathbb{N}\}\,.$$

Recall that by the Jarník-Besicovitch theorem (Theorem 3.2), the complement is of relatively small dimension; namely $\dim\{\alpha\in\mathbb{R}:\tau(\alpha)\ge 3\} = \frac{1}{2}$.

The theorem can be equivalently formulated as follows. Working within the (x,y)-plane, let L_x denote the line parallel to the y-axis passing through the point $(x,0)$. Then, given $\alpha\in\mathrm{I}$, Theorem 4.7 simply states that

$$m_1(W^\times(2,\psi)\,\cap\,\mathrm{L}_\alpha) = 1 \quad \text{if} \quad \psi \text{ statisfies (4.13) and (4.14).}$$

An immediate consequence of the theorem is that (4.11) holds for every α as desired.

Corollary 4.3. *For every $\alpha\in\mathbb{R}$ one has that*

$$\liminf_{q\to\infty} q\,\log^2 q\,||q\alpha||\,||q\beta|| = 0 \quad \textit{for almost all } \beta\in\mathbb{R}\,.$$

Pseudo sketch proof of Theorem 4.7. Given α and ψ, rewrite (4.15) as follows:

$$\|q\beta\| < \Psi_\alpha(q) \quad \textit{where} \quad \Psi_\alpha(q) := \frac{\psi(q)}{\|q\alpha\|} \,. \tag{4.16}$$

We are given (4.13) rather than the above divergent sum condition. So we need to show that

$$\sum_{q=1}^{\infty} \psi(q) \, \log q \; = \; \infty \qquad \Longrightarrow \qquad \sum_{q=1}^{\infty} \Psi_\alpha(q) \; = \; \infty \; . \tag{4.17}$$

This follows (*exercise*) on using partial summation together with the following fact established in [17]. For any irrational α and $Q \geq 2$

$$\sum_{q=1}^{Q} \frac{1}{\|q\alpha\|} \; \geq \; 2\, Q \log Q \; . \tag{4.18}$$

This lower bound estimate strengthens a result of Schmidt [84] – his result is for almost all α rather than all irrationals. Now, if $\Psi_\alpha(q)$ was a monotonic function of q we could have used Khintchine's Theorem, which would then imply that

$$m_1(W(\Psi_\alpha)) = 1 \quad \text{if} \quad \sum_{q=1}^{\infty} \Psi_\alpha(q) \; = \; \infty \; . \tag{4.19}$$

Unfortunately, Ψ_α is not monotonic. Nevertheless, the argument given in [17] overcomes this difficulty.

$\boxtimes$

It is worth mentioning that Corollary 4.3 together with Peck's theorem and Theorem PV adds weight to the argument made in [8] for the following strengthening of Littlewood's Conjecture.

Conjecture 4.1. *For any pair $(\alpha, \beta) \in \mathrm{I}^2$,*

$$\liminf_{q \to \infty} q \; \log q \, ||q\alpha|| \, ||q\beta|| < +\infty \; .$$

Furthermore, it is argued in [8] that the natural analogue of **Bad** within the multiplicative setup is the set:

$$\mathbf{Mad} := \{(\alpha, \beta) \in \mathbb{R}^2 \; : \; \liminf_{q \to \infty} q \cdot \log q \cdot ||q\alpha|| \cdot ||q\beta|| > 0\}.$$

Note that Badziahin [4] has proven that there is a set of (α, β) of full Hausdorff dimension such that

$$\liminf_{q \to \infty} q \cdot \log q \cdot \log \log q \cdot ||q\alpha|| \cdot ||q\beta|| > 0 \, .$$

Regarding the convergence counterpart to Theorem 4.7, the following statement is established in [17].

Theorem 4.8. *Let $\alpha \in \mathbb{R}$ be any irrational real number and let $\psi : \mathbb{N} \to \mathbb{R}^+$ be such that*

$$\sum_{q=1}^{\infty} \psi(q)\, \log q \, < \, \infty$$

Furthermore, assume either of the following two conditions:

(i) *$n \mapsto n\psi(n)$ is decreasing and*

$$\sum_{n=1}^{N} \frac{1}{n\|n\alpha\|} \ll (\log N)^2 \qquad \textit{for all } N \geq 2\,; \tag{4.20}$$

(ii) *$n \mapsto \psi(n)$ is decreasing and*

$$\sum_{n=1}^{N} \frac{1}{\|n\alpha\|} \ll N \log N \qquad \textit{for all } N \geq 2\,. \tag{4.21}$$

Then for almost all $\beta \in \mathbb{R}$, there exist only finitely many $n \in \mathbb{N}$ such that

$$\|n\alpha\|\,\|n\beta\| < \psi(n)\,. \tag{4.22}$$

The behaviour of the sums (4.20) and (4.21) is explicitly studied in term of the continued fraction expansion of α. In particular, it is shown in [17] that (4.20) holds for almost all real numbers α while (4.21) fails for almost all real numbers α. An intriguing question formulated in [17] concerns the behaviour of the above sums for algebraic α of degree ≥ 3. In particular, it is conjectured that (4.20) is true for any real algebraic number α of degree ≥ 3. As is shown in [17], this is equivalent to the following statement.

Conjecture 4.2. *For any algebraic $\alpha = [a_0; a_1, a_2, \dots] \in \mathbb{R} \setminus \mathbb{Q}$, we have that*

$$\sum_{k=1}^{n} a_k \ll n^2\,.$$

Remark 4.11. Computational evidence for specific algebraic numbers does support this conjecture [34].

4.5 Khintchine on fibers

In this section we look for a strengthening of Khintchine simultaneous theorem (Theorem 4.5) akin to the strengthening of Gallagher's multiplicative theorem described above in §4.4.1. For ease of discussion, we begin with the case that $n = 2$ and whether or not Theorem 4.5 remains true if we fix $\alpha \in \mathrm{I}$. In other words, if L_α is the line parallel to the y-axis passing through the point $(\alpha, 0)$ and ψ is monotonic, then is it true that

$$m_1(W(2,\psi) \cap \mathrm{L}_\alpha) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} \psi^2(q) < \infty \\ & \\ 1 & \text{if } \sum_{q=1}^{\infty} \psi^2(q) = \infty \end{cases} \qquad \textbf{????}$$

The question marks are deliberate. They emphasize that the above statement is a question and not a fact or a claim. Indeed, it is easy to see that the convergent statement is false. Simply take α to be rational, say, $\alpha = \frac{a}{b}$. Then, by Dirichlet's theorem, for any β there exist infinitely many $q \in \mathbb{N}$ such that $\|q\beta\| < q^{-1}$ and so it follows that

$$\|bq\beta\| < \frac{b}{q} = \frac{b^2}{bq} \quad \text{and} \quad \|bq\alpha\| = 0 < \frac{b^2}{bq} .$$

This shows that every point on the rational vertical line L_α is $\psi(q) = b^2 q^{-1}$ - approximable and so

$$m_1(W(2,\psi) \cap \mathrm{L}_\alpha) = 1 \quad \text{but} \quad \sum_{q=1}^{\infty} \psi^2(q) = \sum_{q=1}^{\infty} b^4 q^{-2} < \infty .$$

Now, concerning the divergent statement, we claim it is true.

Conjecture 4.3. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $\alpha \in \mathrm{I}$. Then*

$$m_1(W(2,\psi) \cap \mathrm{L}_\alpha) = 1 \quad \text{if} \quad \sum_{q=1}^{\infty} \psi^2(q) = \infty . \tag{4.23}$$

In order to state the current results, we need the notion of the Diophantine exponent of a real number. For $\mathbf{x} \in \mathbb{R}^n$, we let

$$\tau(\mathbf{x}) := \sup\{\tau : \mathbf{x} \in W(n,\tau)\} \tag{4.24}$$

denote the *Diophantine exponent of* $\mathbf{x}$. A word of warning, this notion of Diophantine exponent should not be confused with the Diophantine exponents introduced later in §4.6.1. Note that in view of (4.8), we always have that $\tau(\mathbf{x}) \geq 1/n$. In particular, for $\alpha \in \mathbb{R}$ we have that $\tau(\alpha) \geq 1$. The following result is established in [79].

Theorem 4.9 (F. Ramírez, D. Simmons, F. Süess)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $\alpha \in \mathrm{I}$.*

A. *If $\tau(\alpha) < 2$, then* (4.23) *is true.*

B. *If $\tau(\alpha) > 2$ and for $\epsilon > 0$, $\psi(q) > q^{-\frac{1}{2}-\epsilon}$ for q large enough, then $W(2,\psi) \cap \mathrm{L}_\alpha = \mathrm{I}^2 \cap \mathrm{L}_\alpha$. In particular, $m_1(W(2,\psi) \cap \mathrm{L}_\alpha) = 1$.*

Remark 4.12. Though we have only stated it for lines in the plane, Theorem 4.9 is actually true for lines in $\mathbb{R}^n$. There, we fix an $(n-1)$-tuple of coordinates $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{n-1})$, and consider the line $L_{\boldsymbol{\alpha}} \subset \mathbb{R}^n$. We obtain the same result, with a "cut-off" at n in the *dual Diophantine exponent* of $\boldsymbol{\alpha} \in \mathbb{R}^{n-1}$. The dual Diophantine exponent $\tau^*(\mathbf{x})$ of a vector $\mathbf{x} \in \mathbb{R}^n$ is defined similarly to the (simultaneous) Diophantine exponent, defined above by (4.24), and in the case of numbers (*i.e.*, one-dimensional vectors), the two notions coincide – see §4.6.1 for the formal definition of $\tau^*(\mathbf{x})$.

Remark 4.13. This cut-off in Diophantine exponent, which in Theorem 4.9 happens at $\tau(\alpha) = 2$, seems quite unnatural: why should real numbers with Diophantine exponent 2 be special? Still, such points are inaccessible to our methods. We will see the obstacle in the counting estimate (4.26) which is used for the proof of Part A and is unavailable for $\tau(\alpha) = 2$, and in our application of Khintchine's Transference Principle for the proof of Part B.

Remark 4.14. Note that in Part B, the 'in particular' full measure conclusion is immediate and does not even require the divergent sum condition associated with (4.23).

Regarding the natural analogous conjecture for higher-dimensional subspaces, we have the following statement from [79] which provides a complete solution in the case of affine co-ordinate subspaces of dimension at least two.

Theorem 4.10. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and given $\boldsymbol{\alpha} \in \mathrm{I}^{n-d}$ where $2 \leq d \leq n-1$, let $\mathrm{L}_{\boldsymbol{\alpha}} := \{\boldsymbol{\alpha}\} \times \mathbb{R}^d$. Then*

$$m_d(W(n,\psi) \cap \mathrm{L}_{\boldsymbol{\alpha}}) = 1 \quad \text{if} \quad \sum_{q=1}^{\infty} \psi^n(q) = \infty\,. \tag{4.25}$$

Remark 4.15. Notice that Theorem 4.10 requires $d \geq 2$, thereby excluding lines in $\mathbb{R}^n$. In this case, the obstacle is easy to describe: the proof of Theorem 4.10 relies on Gallagher's extension of Khintchine's theorem, telling us that the monotonicity assumption can be dropped in higher dimensions (see Remark 4.5). In the proof of Theorem 4.10 we find a natural way to apply this directly to the fibers, therefore, we must require $d \geq 2$.

But this is again only a consequence of the chosen method of proof, and not necessarily a reflection of reality. Indeed, Theorem 4.9 (and its more general version for lines in $\mathbb{R}^n$) suggests that we should be able to relax Theorem 4.10 to include the case where $d = 1$.

Remark 4.16. The case when $d = n-1$ was first treated in [78]. There, a number of results are proved in the direction of Theorem 4.10, but with various restrictions on Diophantine exponent, or on the approximating function.

Regarding the proof of Theorem 4.9, Part B makes use of Khintchine's Transference Principle (see §4.6.1 below) while the key to establishing Part A is the following measure theoretic statement (cf. Theorem 1.3) and ubiquity (see §5 below).

Proposition 4.1. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function such that for all $\epsilon > 0$ we have $\psi(q) > q^{-\frac{1}{2}-\epsilon}$ for all q large enough. Let $\alpha \in \mathbb{R}$ be a number with Diophantine exponent $\tau(\alpha) < 2$. Then for any $0 < \epsilon < 1$ and integer $k \geq k_0(\epsilon)$, we have that*

$$m_1\left(\bigcup_{\substack{k^{n-1} < q \leq k^n: \\ \|q\alpha\| \leq \psi(k^n)}} \bigcup_{p=0}^{q} B\left(\tfrac{p}{q}, \tfrac{k}{k^{2n}\psi(k^n)}\right) \right) \geq 1 - \epsilon\,.$$

Remark 4.17. Note that within the context of Theorem 4.9, since α is fixed it is natural to consider only those $q \in \mathbb{N}$ for which $\|q\alpha\| \leq \psi(q)$ when considering solutions to the inequality $\|q\beta\| \leq \psi(q)$. In other words, if we let

$$\mathcal{A}_\alpha(\psi) := \{q \in \mathbb{N} : \|q\alpha\| \leq \psi(q)\}$$

then by definition

$$W(2,\psi)\cap \mathrm{L}_\alpha=\{(\alpha,\beta)\in \mathrm{L}_\alpha\cap \mathrm{I}^2:\|q\beta\|\le\psi(q) \text{ for infinitely many } q\in\mathcal{A}_\alpha(\psi)\}\,.$$

It is clear that the one-dimensional Lebesgue measure m_1 of this set is the same as that of

$$\{\beta\in \mathrm{I}:\|q\beta\|\le\psi(q) \text{ for infinitely many } q\in\mathcal{A}_\alpha(\psi)\}\,.$$

Sketch proof of Proposition 4.1. In view of Minkowski's theorem for systems of linear forms, for any $(\alpha,\beta)\in\mathbb{R}^2$ and integer $N\ge 1$, there exists an integer $q\ge 1$ such that

$$\begin{aligned}\|q\alpha\| &\le \psi(N)\\ \|q\beta\| &\le \frac{1}{N\,\psi(N)}\\ q &\le N\,.\end{aligned}$$

The desired statement follows on exploiting this with $N=k^n$ together with the following result which is a consequence of a general counting result established in [17]: given ψ and α satisfying the conditions imposed in Proposition 4.1, then for n sufficiently large

$$\#\{q\le k^{n-1}:\|q\alpha\|\le\psi(k^n)\}\ \le\ 31\,\psi(k^n)\,k^{n-1}\,. \tag{4.26}$$

(An analogous count is established in [79] for vectors $\boldsymbol{\alpha}\in\mathbb{R}^{n-1}$.) *Exercise:* Fill in the details of the above sketch. ⊠

4.6 Dual approximation and Khintchine's Transference

Instead of simultaneous approximation by rational points as considered in the previous section, one can consider the closeness of the point $\mathbf{x}=(x_1,\dots,x_m)\in\mathbb{R}^m$ to rational hyperplanes given by the equations $\mathbf{q}\cdot\mathbf{x}=p$ with $p\in\mathbb{Z}$ and $\mathbf{q}\in\mathbb{Z}^m$. The point $\mathbf{x}\in\mathbb{R}^n$ will be called *dually ψ-well approximable* if the inequality

$$|\mathbf{q}\cdot\mathbf{x}-p|<\psi(|\mathbf{q}|)$$

holds for infinitely many $(p,\mathbf{q})\in\mathbb{Z}\times\mathbb{Z}^m$ with $|\mathbf{q}|:=|\mathbf{q}|_\infty=\max\{|q_1|,\dots,|q_m|\}>0$. The set of dually ψ-approximable points in I^m will be denoted by $W^*(m,\psi)$. In the case $\psi:q\to q^{-\tau}$ with $\tau>0$, we write $W^*(m,\tau)$ for $W^*(m,\psi)$. The set $W^*(n,\tau)$ is the set of *dually τ-well approximable numbers.* Note that in view of Corollary 4.1 we have that

$$W^*(m,\tau)=\mathrm{I}^m \quad\text{if}\quad \tau\le m. \tag{4.27}$$

The simultaneous and dual forms of approximation are special cases of a system of linear forms, covered by a general extension due to A. V. Groshev (see [90]). This treats real $m\times n$ matrices X, regarded as points in $\mathbb{R}^{mn}$, which are ψ-approximable. More precisely, $X=(x_{ij})\in\mathbb{R}^{mn}$ is said to be ψ-approximable if the inequality

$$\|\mathbf{q}X\|<\psi(|\mathbf{q}|)$$

is satisfied for infinitely many $\mathbf{q} \in \mathbb{Z}^m$. Here $\mathbf{q}X$ is the system

$$q_1 x_{1j} + \cdots + q_m x_{m,j} \qquad (1 \leq j \leq n)$$

of n real linear forms in m variables and $\|\mathbf{q}X\| := \max_{1 \leq j \leq n} \|\mathbf{q} \cdot X^{(j)}\|$, where $X^{(j)}$ is the j'th column vector of X. As the set of ψ-approximable points is translation invariant under integer vectors, we can restrict attention to the mn-dimensional unit cube I^{mn}. The set of ψ-approximable points in I^{mn} will be denoted by

$$W(m,n,\psi) := \{X \in \mathrm{I}^{mn} : \|\mathbf{q}X\| < \psi(|\mathbf{q}|) \text{ for infinitely many } \mathbf{q} \in \mathbb{Z}^m\}.$$

Thus, $W(n,\psi) = W(1,n,\psi)$ and $W^*(m,\psi) = W(m,1,\psi)$. The following result naturally extends Khintchine's simultaneous theorem to the linear forms setup. For obvious reasons, we write $|X|_{mn}$ rather than $m_{mn}(X)$ for mn-dimensional Lebesgue measure of a set $X \subset \mathbb{R}^{mn}$.

Theorem 4.11 (Khintchine-Groshev, 1938)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$. Then*

$$|W(m,n,\psi)|_{mn} = \begin{cases} 0 & \text{if } \displaystyle\sum_{r=1}^{\infty} r^{m-1}\psi(r)^n < \infty, \\[2ex] 1 & \text{if } \displaystyle\sum_{r=1}^{\infty} r^{m-1}\psi(r)^n = \infty \quad \text{and } \psi \text{ is monotonic.} \end{cases}$$

The counterexample due to Duffin and Schaeffer mentioned in §2.2.1 means that the monotonicity condition cannot be dropped from Groshev's theorem when $m = n = 1$. To avoid this situation, let $mn > 1$. Then for $m = 1$, we have already mentioned (Remark 4.5) that the monotonicity condition can be removed. Furthermore, the monotonicity condition can also be removed for $m > 2$ – see [13, Theorem 8] and [90, Theorem 14]. The $m = 2$ situation was resolved only recently in [27], where it was shown that the monotonicity condition can be safely removed. The upshot of this discussion is that we only require the monotonicity condition in the Khintchine-Groshev theorem in the case when $mn = 1$.

Naturally, one can ask for a Hausdorff measure generalisation of the Khintchine-Groshev theorem. The following is such a statement and as one should expect it coincides with Theorem 3.4 when $m = n = 1$. In the simultaneous case ($m = 1$), the result was alluded to within Remark 4.8 following the simultaneous statement of Khintchine's theorem.

Theorem 4.12. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$. Then*

$$\mathcal{H}^s(W(m,n,\psi)) = \begin{cases} 0 & \text{if } \displaystyle\sum_{r=1}^{\infty} r^{m(n+1)-1-s}\psi(r)^{s-n(m-1)} < \infty\,, \\[3ex] \mathcal{H}^s(\mathrm{I}^{mn}) & \text{if } \displaystyle\sum_{r=1}^{\infty} r^{m(n+1)-1-s}\psi(r)^{s-n(m-1)} = \infty \\ & \qquad\qquad\qquad \text{and } \psi \text{ is monotonic}\,. \end{cases}$$

This Hausdorff theorem follows from the corresponding Lebesgue statement in the same way that Khintchine's theorem implies Jarník's theorem via the Mass Transference Principle—see §3.4.1. The Mass Transference Principle introduced in §3.4 deals with lim sup sets which are defined by a sequence of balls. However, the 'slicing' technique introduced in [22] extends the Mass Transference Principle to deal with lim sup sets defined by a sequence of neighborhoods of 'approximating' planes. This naturally enables us to generalise the Lebesgue measure statements for systems of linear forms to Hausdorff measure statements. The last sentence should come with a warning. It gives the impression that in view of the discussion preceding Theorem 4.11, one should be able to establish Theorem 4.12 directly, without the monotonicity assumption except when $m = n = 1$. However, as things currently stand we also need to assume monotonicity when $m = 2$. For further details see [13, §8].

Returning to Diophantine approximation in $\mathbb{R}^n$, we consider the following natural question.

Question. Is there a connection between the simultaneous ($m = 1$) and dual ($n = 1$) forms of approximating points in $\mathbb{R}^n$?

4.6.1 Khintchine's Transference

The simultaneous and dual forms of Diophantine approximation are related by a 'transference' principle in which a solution of one form is related to a solution of the other. In order to state the relationship we introduce the quantities ω^* and ω. For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, let

$$\omega^*(\mathbf{x}) \; := \; \sup \{\omega \in \mathbb{R} : \mathbf{x} \in W^*(n, n+\omega)\}$$

and

$$\omega(\mathbf{x}) \; := \; \sup \left\{\omega \in \mathbb{R} : \mathbf{x} \in W(n, \tfrac{1+\omega}{n})\right\} .$$

Note that

$$\tau(\mathbf{x}) = \frac{1+\omega(\mathbf{x})}{n}$$

where $\tau(\mathbf{x})$ is the Diophantine exponent of $\mathbf{x}$ as defined by (4.24). For the sake of completeness we mention that the quantity

$$\tau^*(\mathbf{x}) = n + \omega^*(\mathbf{x})$$

is called the dual Diophantine exponent. The following statement provides a relationship between the dual and simultaneous Diophantine exponents.

Theorem 4.13 (Khintchine's Transference Principle)**.** *For $\mathbf{x} \in \mathbb{R}^n$, we have that*

$$\frac{\omega^*(\mathbf{x})}{n^2 + (n-1)\omega^*(\mathbf{x})} \leq \omega(\mathbf{x}) \leq \omega^*(\mathbf{x})$$

with the left hand side being interpreted as $1/(n-1)$ if $\omega^*(\mathbf{x})$ is infinite.

Remark 4.18. The transference principle implies that given any $\epsilon > 0$, if $\mathbf{x} \in W(n, \frac{1+\epsilon}{n})$ then $\mathbf{x} \in W^*(n, n+\epsilon^*)$ for some ϵ^* comparable to ϵ, and vice versa.

Proof of Part B of Theorem 4.9

Part B of Theorem 4.9 follows by plugging $n = 2$ and $d = 1$ into the following proposition, which is in turn a simple consequence of Khintchine's Transference Principle.

Proposition 4.2. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and given $\boldsymbol{\alpha} \in \mathrm{I}^{n-d}$ where $1 \leq d \leq n-1$, let $\mathrm{L}_{\boldsymbol{\alpha}} := \{\boldsymbol{\alpha}\} \times \mathbb{R}^d$. Assume that $\tau(\boldsymbol{\alpha}) > \frac{1+d}{n-d}$ and for $\epsilon > 0$, $\psi(q) > q^{-\frac{1}{n}-\epsilon}$ for q large enough. Then*

$$W(n,\psi) \cap \mathrm{L}_{\boldsymbol{\alpha}} = \mathrm{I}^n \cap \mathrm{L}_{\boldsymbol{\alpha}} \,.$$

In particular, $m_d(W(n,\psi) \cap \mathrm{L}_{\boldsymbol{\alpha}}) = 1$.

Proof. We are given that $\tau(\boldsymbol{\alpha}) > \frac{1+d}{n-d}$ and so by definition $\omega(\boldsymbol{\alpha}) > d$. Thus, by Khintchine's Transference Principle, it follows that $\omega^*(\boldsymbol{\alpha}) > d$ and so $\omega^*(\mathbf{x}) > 0$ for any point $\mathbf{x} = (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}^n$; i.e. $\boldsymbol{\beta} \in \mathbb{R}^d$ and $\mathbf{x}$ is a point on the d-dimensional plane $\mathrm{L}_{\boldsymbol{\alpha}}$. On applying Khintchine's Transference Principle again, we deduce that $\omega(\mathbf{x}) > 0$ which together with the growth condition imposed on ψ implies the desired conclusion. □

5 Ubiquitous systems of points

In [14], a general framework is developed for establishing divergent results analogous to those of Khintchine and Jarník for a natural class of lim sup sets. The framework is based on the notion of 'ubiquity', which goes back to [10] and [46] and captures the key measure theoretic structure necessary to prove such measure theoretic laws. The 'ubiquity' introduced below is a much simplified version of that in [14]. In particular, we make no attempt to incorporate the linear forms theory of metric Diophantine approximation. However this does have the advantage of making the exposition more transparent and also leads to cleaner statements which are more than adequate for the application we have in mind; namely to systems of points.

5.1 The general framework and fundamental problem

The *general framework of ubiquity* considered within is as follows.

- (Ω, d) is a compact metric space.
- μ is a Borel probability measure supported on Ω.
- There exist positive constants δ and r_o such that for any $x \in \Omega$ and $r \leq r_0$,

$$a\, r^{\delta} \;\leq\; \mu(B(x,r)) \;\leq\; b\, r^{\delta}. \tag{5.1}$$

  The constants a and b are independent of the ball $B(x,r) := \{y \in \Omega : d(x,y) < r\}$.
- $\mathcal{R} = (R_\alpha)_{\alpha \in J}$ a sequence of points R_α in Ω indexed by an infinite countable set J. The points R_α are commonly referred to as *resonant points*.
- $\beta : J \to \mathbb{R}^+ : \alpha \mapsto \beta_\alpha$ is a positive function on J. It attaches a 'weight' β_α to the resonant point R_α.

- To avoid pathological situations:

$$\#\{\alpha \in J : \beta_\alpha \leq x\} < \infty \quad \text{for any } x \in \mathbb{R}. \tag{5.2}$$

Remark 5.1. The measure condition (5.1) on the *ambient measure* μ implies that μ is non-atomic, that is $\mu(\{x\}) = 0$ for any $x \in \Omega$, and that

$$\mu(\Omega) := 1 \asymp \mathcal{H}^\delta(\Omega) \quad \text{and} \quad \dim \Omega = \delta\,.$$

Indeed, μ is comparable to δ–dimensional Hausdorff measure $\mathcal{H}^\delta$.

Given a decreasing function $\Psi : \mathbb{R}^+ \to \mathbb{R}^+$ let

$$\Lambda(\Psi) = \{x \in \Omega : x \in B(R_\alpha, \Psi(\beta_\alpha)) \text{ for infinitely many } \alpha \in J\}\ .$$

The set $\Lambda(\Psi)$ is a 'lim sup' set; it consists of points in Ω which lie in infinitely many of the balls $B(R_\alpha, \Psi(\beta_\alpha))$ centred at resonant points. As in the classical setting introduced in §2, it is natural to refer to the function Ψ as the *approximating function.* It governs the 'rate' at which points in Ω must be approximated by resonant points in order to lie in $\Lambda(\Psi)$. In view of the finiteness condition (5.2), it follows that for any fixed $k > 1$, the number of α in J with $k^{t-1} < \beta_\alpha \leq k^t$ is finite regardless of the value of $t \in \mathbb{N}$. Therefore $\Lambda(\Psi)$ can be rewritten as the limsup set of

$$\Upsilon(\Psi, k, t) := \bigcup_{\alpha \in J\ :\ k^{t-1} < \beta_\alpha \leq k^t} B(R_\alpha, \Psi(\beta_\alpha))\,;$$

that is

$$\Lambda(\Psi) = \limsup_{t \to \infty} \Upsilon(\Psi, k, t) := \bigcap_{m=1}^{\infty} \bigcup_{t=m}^{\infty} \Upsilon(\Psi, k, t)\ .$$

It is reasonably straightforward to determine conditions under which $\mu(\Lambda(\Psi)) = 0$. In fact, this is implied by the convergence part of the Borel–Cantelli lemma from probability theory whenever

$$\textstyle\sum_{t=1}^{\infty} \mu(\Upsilon(\Psi, k, t)) < \infty\ . \tag{5.3}$$

In view of this it is natural to consider the following **fundamental problem**:

Under what conditions is $\mu(\Lambda(\psi)) > 0$ *and more generally* $\mathcal{H}^s(\Lambda(\Psi)) > 0$ *?*

Ideally, we would like to be able to conclude the full measure statement $\mathcal{H}^s(\Lambda(\Psi)) = \mathcal{H}^s(\Omega)$. Recall that when $s = \delta$, the ambient measure μ coincides with $\mathcal{H}^\delta$. Also, if $s < \delta$ then $\mathcal{H}^s(\Omega) = \infty$.

5.1.1 The basic example

In order to illustrate and clarify the above general setup, we show that the set $W(n, \psi)$ of simultaneously ψ-well approximable points $\mathbf{x} \in \mathrm{I}^n := [0,1]^n$ can be expressed in the form of $\Lambda(\Psi)$. With this in mind, let

- $\Omega := \mathrm{I}^n$ and $d(\mathbf{x},\mathbf{y}) := \max_{1\leq i\leq n} |x_i - y_i|$,
- μ be Lebesgue measure restricted to I^n and $\delta := n$,
- $J := \{(\mathbf{p},q) \in \mathbb{Z}^n \times \mathbb{N} : \mathbf{p}/q \in \mathrm{I}^n\}$ and $\alpha := (\mathbf{p},q) \in J$,
- $\mathcal{R} := (\mathbf{p}/q)_{(\mathbf{p},q)\in J}$ and $\beta_{(\mathbf{p},q)} := q$.

Thus, the resonant points R_α are simply rational points $\mathbf{p}/q := (p_1/q, \ldots, p_n/q)$ in the unit cube I^n. It is readily verified that the measure condition (5.1) and the finiteness condition (5.2) are satisfied and moreover that for any decreasing function $\psi : \mathbb{N} \to \mathbb{R}^+$,

$$\Lambda(\Psi) = W(n,\psi) \quad \text{with} \quad \Psi(q) := \psi(q)/q\,.$$

For this basic example, the solution to the fundamental problem is given by the simultaneous Khintchine-Jarník Theorem (see Theorem 4.12 with $m = 1$ in §4.6).

5.2 The notion of ubiquity

The following 'system' contains the key measure theoretic structure necessary for our attack on the fundamental problem.

Let $\rho : \mathbb{R}^+ \to \mathbb{R}^+$ be a function with $\rho(r) \to 0$ as $r \to \infty$ and let

$$\Delta(\rho, k, t) := \bigcup_{\alpha\in J\,:\,\beta_\alpha\leq k^t} B(R_\alpha, \rho(k^t))\ ,$$

where $k > 1$ is a fixed real number. Note that when $\rho = \Psi$ the composition of $\Delta(\rho, k, t)$ is very similar to that of $\Upsilon(\Psi, k, t)$.

Definition (Ubiquitous system) Let $B = B(x, r)$ denote an arbitrary ball with centre x in Ω and radius $r \leq r_0$. Suppose there exists a function ρ and absolute constants $\kappa > 0$ and $k > 1$ such that for any ball B as above

$$\mu\,(B \cap \Delta(\rho, k, t)) \ \geq\ \kappa\ \mu(B) \quad \text{for } t \geq t_0(B). \tag{5.4}$$

Then the pair $(\mathcal{R}, \beta)$ is said to be a *local μ-ubiquitous system relative to* (ρ, k). If (5.4) does not hold for arbitrary balls with centre x in Ω and radius $r \leq r_0$, but does hold with $B = \Omega$, the pair $(\mathcal{R}, \beta)$ is said to be a *global μ-ubiquitous system relative to* (ρ, k).

Loosely speaking, the definition of local ubiquity says that the set $\Delta(\rho, k, t)$ locally 'approximates' the underlying space Ω in terms of the measure μ. By 'locally' we mean balls centred at points in Ω. The function ρ is referred to as the *ubiquitous function*. The actual values of the constants κ and k in the above definition are irrelevant—it is their existence that is important. In practice, the μ-ubiquity of a system can be established using standard arguments concerning the distribution of the resonant points in Ω, from which the function ρ arises naturally. To illustrate this, we return to the basic example of §5.1.1.

Proposition 5.1. *There is a constant $k > 1$ such that the pair $(\mathcal{R}, \beta)$ defined in §5.1.1 is a local μ-ubiquitous system relative to (ρ, k) where $\rho : r \mapsto \text{const} \times r^{-(n+1)/n}$.*

The one-dimensional case of this proposition follows from Theorem 1.3.

Exercise: Prove the above proposition for arbitrary n. Hint: you will need to use the multi-dimensional version of Dirichlet's theorem, or Minkowski's theorem.

5.3 The ubiquity statements

Before stating the main results regarding ubiquity we introduce one last notion. Given a real number $k > 1$, a function $h : \mathbb{R}^+ \to \mathbb{R}^+$ will be said to be *k-regular* if there exists a strictly positive constant $\lambda < 1$ such that for t sufficiently large

$$h(k^{t+1}) \leq \lambda\, h(k^t) \ . \tag{5.5}$$

The constant λ is independent of t but may depend on k. A consequence of local ubiquity is the following result.

Theorem 5.1 (Ubiquity - the Hausdorff measure case)**.** *Let (Ω, d) be a compact metric space equipped with a probability measure μ satisfying condition* (5.1) *and such that any open subset of Ω is μ-measurable. Suppose that $(\mathcal{R}, \beta)$ is a locally μ-ubiquitous system relative to (ρ, k) and that Ψ is an approximating function. Furthermore, suppose that $s \in (0, \delta]$, that ρ is k-regular and that*

$$\sum_{t=1}^{\infty} \frac{\Psi(k^t)^s}{\rho(k^t)^\delta} \ = \ \infty \, . \tag{5.6}$$

Then

$$\mathcal{H}^s\left(\Lambda(\Psi)\right) \ = \ \mathcal{H}^s\left(\Omega\right) \ .$$

As already mentioned, if $s < \delta$ then $\mathcal{H}^s(\Omega) = \infty$. On the other hand, if $s = \delta$, the Hausdorff measure $\mathcal{H}^\delta$ is comparable to the ambient measure μ and the theorem implies that

$$\mu\left(\Lambda(\Psi)\right) = \mu(\Omega) := 1.$$

Actually, the notion of global ubiquity has implications in the ambient measure case.

Theorem 5.2 (Ubiquity - the ambient measure case)**.** *Let (Ω, d) be a compact metric space equipped with a measure μ satisfying condition* (5.1) *and such that any open subset of Ω is μ-measurable. Suppose that $(\mathcal{R}, \beta)$ is a globally μ-ubiquitous system relative to (ρ, k) and that Ψ is an approximating function. Furthermore, suppose that either ρ or Ψ is k-regular and that*

$$\sum_{t=1}^{\infty} \left(\frac{\Psi(k^t)}{\rho(k^t)}\right)^\delta \ = \ \infty \, . \tag{5.7}$$

Then

$$\mu\left(\Lambda(\Psi)\right) > 0.$$

If in addition $(\mathcal{R}, \beta)$ is a locally μ-ubiquitous system relative to (ρ, k), then

$$\mu\left(\Lambda(\Psi)\right) = 1.$$

Remark 5.2. Note that in Theorem 5.2 we can get away with either ρ or Ψ being k-regular. In the ambient measure case, it is also possible to weaken the measure condition (5.1) (see Theorem 1 in [14, §3]).

Remark 5.3. If we know via some other means that $\Lambda(\Psi)$ satisfies a zero-full law (as indeed is the case for the classical set of $W(n,\psi)$ of ψ-well approximable points), then it is enough to show that $\mu(\Lambda(\Psi)) > 0$ in order to conclude full measure.

The above results constitute the main theorems appearing in [14] tailored to the setup considered here. In fact, Theorem 5.1 as stated appears in [25] for the first time. Previously, the Hausdorff and ambient measure cases had been thought of and stated separately.

The concept of ubiquity was originally formulated by Dodson, Rynne & Vickers [46] to obtain lower bounds for the Hausdorff dimension of lim sup sets. Furthermore, the ubiquitous systems of [46] essentially coincide with the regular systems of Baker & Schmidt [10] and both have proved very useful in obtaining lower bounds for the Hausdorff dimension of lim sup sets. However, unlike the framework developed in [14], both [10] and [46] fail to shed any light on establishing the more desirable divergent Khintchine and Jarník type results. The latter clearly implies lower bounds for the Hausdorff dimension. For further details regarding regular systems and the original formulation of ubiquitous systems see [14, 31].

5.3.1 The basic example and the simultaneous Khintchine-Jarník Theorem

Regarding the basic example of §5.1.1, recall that

$$\Lambda(\Psi) = W(n,\psi) \quad \text{with} \quad \Psi(q) := \psi(q)/q$$

and that Proposition 5.1 states that for k large enough, the pair $(\mathcal{R},\beta)$ is a local μ-ubiquitous system relative to (ρ,k) where

$$\rho : r \mapsto \text{const} \times r^{-(n+1)/n}\,.$$

Now, clearly the function ρ is k-regular. Also note that the divergence sum condition (5.6) associated with Theorem 5.1 becomes

$$\sum_{t=1}^{\infty} k^{t(n+1-s)}\psi(k^t)^s = \infty\,.$$

If ψ is monotonic, this is equivalent to

$$\sum_{q=1}^{\infty} q^{n-s}\psi(q)^s = \infty\,,$$

and Theorem 5.1 implies that

$$\mathcal{H}^s(W(n,\psi)) = \mathcal{H}^s(\mathrm{I}^n)\,.$$

The upshot is that Theorem 5.1 implies the divergent case of the simultaneous Khintchine-Jarník Theorem; namely, Theorem 4.12 with $m = 1$ in §4.6.

Remark 5.4. It is worth standing back a little and thinking about what we have actually used in establishing the classical results—namely, local ubiquity. Within the classical setup, local ubiquity is a simple measure theoretic statement concerning the distribution of rational points with respect to Lebesgue measure—the natural measure on the unit interval. From this we are able to obtain the divergent parts of both Khintchine's Theorem (a Lebesgue measure statement) and Jarník's Theorem (a Hausdorff measure statement). In other words, the Lebesgue measure statement of local ubiquity underpins the general Hausdorff measure theory of the lim sup set $W(n,\psi)$. This of course is very much in line with the subsequent discovery of the Mass Transference Principle discussed in §3.4.

The applications of ubiquity are widespread, as demonstrated in [14, §12]. We now consider a more recent application of ubiquity to the 'fibers' strengthening of Khintchine's simultaneous theorem described in §4.5.

5.3.2 Proof of Theorem 4.9: Part A

Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $\alpha \in \mathrm{I}$ such that it has Diophantine exponent $\tau(\alpha) < 2$. In view of Remark 4.17 in §4.5, establishing Theorem 4.9 is equivalent to showing that

$$m(\Pi(\psi,\alpha)) = 1 \quad \text{if} \quad \sum_{q=1}^{\infty} \psi^2(q) = \infty$$

where

$$\Pi(\psi,\alpha) := \{\beta \in \mathrm{I} : \|q\beta\| \leq \psi(q) \text{ for infinitely many } q \in \mathcal{A}_\alpha(\psi)\}\,.$$

Recall,

$$\mathcal{A}_\alpha(\psi) := \{q \in \mathbb{N} : \|q\alpha\| \leq \psi(q)\}\,.$$

Remark 5.5. Without loss of generality, we can assume that

$$q^{-\frac{1}{2}}(\log q)^{-1} \ \leq\ \psi(q)\ \leq\ q^{-\frac{1}{2}} \quad \forall\, q \in \mathbb{N}\,. \tag{5.8}$$

Exercise: Verify that this is indeed the case. For the right-hand side of (5.8), consider the auxiliary function

$$\tilde{\psi} : q \to \tilde{\psi} := \min\{q^{-\frac{1}{2}}, \psi(q)\}$$

and show that $\sum_{q=1}^{\infty} \tilde{\psi}^2(q) = \infty$. For the left-hand side of (5.8), consider the auxiliary function

$$\tilde{\psi} : q \to \tilde{\psi}(q) := \max\{\hat{\psi}(q) := q^{-\frac{1}{2}}(\log q)^{-1}, \psi(q)\}$$

and show that $m(\Pi(\hat{\psi},\alpha)) = 0$ by making use of the counting estimate (4.26) and the convergence Borel-Cantelli Lemma.

We now show that the set $\Pi(\psi,\alpha)$ can be expressed in the form of $\Lambda(\Psi)$. With this in mind, let

- $\Omega := [0,1]$ and $d(x,y) := |x-y|$,
- μ be Lebesgue measure restricted to I and $\delta := 1$,
- $J := \{(p,q) \in \mathbb{Z} \times \mathcal{A}_\alpha(\psi) : p/q \in \mathrm{I}\}$ and $\alpha := (p,q) \in J$,
- $\mathcal{R} := (p/q)_{(p,q)\in J}$ and $\beta_{(p,q)} := q$.

Thus, the resonant points R_α are simply rational points p/q in the unit interval I with denominators q restricted to the set $\mathcal{A}_\alpha(\psi)$. It is readily verified that the measure condition (5.1) and the finiteness condition (5.2) are satisfied and moreover that for any decreasing function $\psi : \mathbb{N} \to \mathbb{R}^+$,

$$\Lambda(\Psi) = \Pi(\psi, \alpha) \quad \text{with} \quad \Psi(q) := \psi(q)/q\,.$$

Note that since ψ is decreasing, the function Ψ is k-regular. Now, in view of Remark 5.5, the conditions of Proposition 4.1 are satisfied and we conclude that for k large enough, the pair $(\mathcal{R}, \beta)$ is a global m-ubiquitous system relative to (ρ, k) where

$$\rho : r \mapsto \frac{k}{r^2\psi(r)}\,.$$

Now, since ψ is monotonic

$$\sum_{t=1}^{\infty} \frac{\Psi(k^t)}{\rho(k^t)} = \sum_{t=1}^{\infty} k^{t-1}\psi^2(k^t) = \infty \quad \Longleftrightarrow \quad \sum_{q=1}^{\infty} \psi^2(q) = \infty$$

and Theorem 5.2 implies that

$$\mu\Big(\Pi(\psi, \alpha)\Big) > 0\,.$$

Now observe that $\Pi(\psi, \alpha)$ is simply the set $W(\bar{\psi})$ of $\bar{\psi}$–well approximable numbers with $\bar{\psi}(q) := \psi(q)$ if $q \in \mathcal{A}_\alpha(\psi)$ and zero otherwise. Thus, Cassels' zero-full law [38] implies the desired statement; namely that

$$\mu\Big(\Pi(\psi, \alpha)\Big) = 1\,.$$

6 Diophantine approximation on manifolds

Diophantine approximation on manifolds (as coined by Bernik & Dodson in their Cambridge Tract [31]) or *Diophantine approximation of dependent quantities* (as coined by Sprindžuk in his monograph [90]) refers to the study of Diophantine properties of points in $\mathbb{R}^n$ whose coordinates are confined by functional relations or equivalently are restricted to a sub-manifold $\mathcal{M}$ of $\mathbb{R}^n$. Thus, in the case of simultaneous Diophantine approximation one studies sets such as

$$\mathcal{M} \cap W(n, \psi)\,.$$

To some extent we have already touched upon the theory of Diophantine approximation on manifolds when we considered Gallagher multiplicative theorem on fibers in §4.4.1 and

Khintchine simultaneous theorem on fibers in §4.5. In these sections the points of interest are confined to an affine co-ordinate subspace of $\mathbb{R}^n$; namely the manifold

$$\mathrm{L}_{\boldsymbol{\alpha}} := \{\boldsymbol{\alpha}\} \times \mathbb{R}^d, \text{ where } 1 \leq d \leq n-1 \text{ and } \boldsymbol{\alpha} \in \mathrm{I}^{n-d}.$$

In general, a manifold $\mathcal{M}$ can locally be given by a system of equations, for instance, the unit sphere in $\mathbb{R}^3$ is given by the equation

$$x^2 + y^2 + z^2 = 1;$$

or it can be immersed into $\mathbb{R}^n$ by a map $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^n$ (the actual domain of $\mathbf{f}$ can be smaller than $\mathbb{R}^d$), for example, the Veronese curve is given by the map

$$x \mapsto (x, x^2, \ldots, x^n).$$

Such a map $\mathbf{f}$ is often referred to as a parameterisation and *without loss of generality we will assume that the domain of* $\mathbf{f}$ *is* I^d *and that the manifold* $\mathcal{M} \subseteq \mathrm{I}^n$. Locally, a manifold given by a system of equations can be parameterised by some map $\mathbf{f}$ and, conversely, if a manifold is immersed by a map $\mathbf{f}$, it can be written using a system of $n-d$ equations, where d is the dimension of the manifold.

Exercise: Parameterise the upper hemisphere $x^2 + y^2 + z^2 = 1$, $z > 0$, and also write the Veronese curve (see above) by a system of equations.

In these notes we will mainly concentrate on the simultaneous (rather than dual) theory of Diophantine approximation on manifolds. In particular, we consider the following two natural problems.

Problem 1. To develop a Lebesgue theory for $\mathcal{M} \cap W(n, \psi)$.

Problem 2. To develop a Hausdorff theory for $\mathcal{M} \cap W(n, \psi)$.

In short, the aim is to establish analogues of the two fundamental theorems of Khintchine and Jarník, and thereby provide a complete measure theoretic description of the sets $\mathcal{M} \cap W(n, \psi)$. The fact that the points $\mathbf{x} \in \mathbb{R}^n$ of interest are of dependent variables, which reflects the fact that $\mathbf{x} \in \mathcal{M}$, introduces major difficulties in attempting to describe the measure theoretic structure of $\mathcal{M} \cap W(n, \psi)$. This is true even in the specific case that $\mathcal{M}$ is a planar curve. More to the point, even for seemingly simple curves such as the unit circle or the parabola the above problems are fraught with difficulties. In these notes we will concentrate mainly on describing the Lebesgue theory.

Unless stated otherwise, the approximating function $\psi : \mathbb{N} \to \mathbb{R}^+$ *throughout this section is assumed to be monotonic.*

6.1 The Lebesgue theory for manifolds

The goal is to obtain a Khintchine type theorem that describes the Lebesgue measure of the set $\mathcal{M} \cap W(n, \psi)$ of simultaneously ψ–approximable points lying on $\mathcal{M}$. First of all notice that if the dimension d of the manifold $\mathcal{M}$ is strictly less than n then $m_n(\mathcal{M} \cap W(n, \psi)) = 0$ irrespective of the approximating function ψ. Thus, in attempting to develop a Lebesgue theory for $\mathcal{M} \cap W(n, \psi)$ it is natural to use the induced d-dimensional Lebesgue measure on $\mathcal{M}$. Alternatively, if $\mathcal{M}$ is immersed by a map $\mathbf{f} : \mathrm{I}^d \to \mathbb{R}^n$ we use the d-dimensional Lebesgue measure m_d on the set of parameters of $\mathbf{f}$; namely I^d. In either case, the measure under consideration will be denoted by $| \ . \ |_{\mathcal{M}}$.

Remark 6.1. Notice that for $\tau \leq 1/n$, we have that $|\mathcal{M} \cap W(n, \tau)|_{\mathcal{M}} = |\mathcal{M}|_{\mathcal{M}} :=$ FULL as it should be since, by Dirichlet's theorem, we have that $W(n, \tau) = \mathrm{I}^n$.

The two-dimension fiber problem considered in §4.5, in which the manifold $\mathcal{M}$ is a vertical line L_α, shows that it is not possible to obtain a Khintchine type theorem (both the convergence and divergence aspects) for all manifolds. Indeed, the convergent statement fails for vertical lines. Thus, in a quest for developing a general Khintchine type theory for manifolds (cf. Problem 1 above) , it is natural to avoid lines and more generally hyperplanes. In short, we insist that the manifold under consideration is "sufficiently" curved.

6.1.1 Non-degenerate manifolds

In order to make any reasonable progress with Problems 1 & 2 above, we assume that the manifolds $\mathcal{M}$ under consideration are **non-degenerate** [67]. Essentially, these are smooth sub-manifolds of $\mathbb{R}^n$ which are sufficiently curved so as to deviate from any hyperplane. Formally, a manifold $\mathcal{M}$ of dimension d embedded in $\mathbb{R}^n$ is said to be non-degenerate if it arises from a non–degenerate map $\mathbf{f} : U \to \mathbb{R}^n$ where U is an open subset of $\mathbb{R}^d$ and $\mathcal{M} := \mathbf{f}(U)$. The map $\mathbf{f} : U \to \mathbb{R}^n : \mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$ is said to be *non–degenerate at* $\mathbf{x} \in U$ if there exists some $l \in \mathbb{N}$ such that $\mathbf{f}$ is l times continuously differentiable on some sufficiently small ball centred at $\mathbf{x}$ and the partial derivatives of $\mathbf{f}$ at $\mathbf{x}$ of orders up to l span $\mathbb{R}^n$. The map $\mathbf{f}$ is *non–degenerate* if it is non–degenerate at almost every (in terms of d–dimensional Lebesgue measure) point in U; in turn the manifold $\mathcal{M} = \mathbf{f}(U)$ is also said to be non–degenerate. Any real, connected analytic manifold not contained in any hyperplane of $\mathbb{R}^n$ is non–degenerate. Indeed, if $\mathcal{M}$ is immersed by an analytic map $\mathbf{f} = (f_1, \ldots, f_n) : U \to \mathbb{R}^n$ defined on a ball $U \subset \mathbb{R}^d$, then $\mathcal{M}$ is non-degenerate if and only if the functions $1, f_1, \ldots, f_n$ are linearly independent over $\mathbb{R}$.

Without loss of generality, we will assume that U is I^d and that the manifold $\mathcal{M} \subseteq \mathrm{I}^n$

Note that in the case the manifold $\mathcal{M}$ is a planar curve $\mathcal{C}$, a point on $\mathcal{C}$ is non-degenerate if the curvature at that point is non-zero. Thus, $\mathcal{C}$ is a non-degenerate planar curve if the set of points on $\mathcal{C}$ at which the curvature vanishes is a set of one–dimensional Lebesgue measure zero. Moreover, it is not difficult to show that the set of points on a planar curve at which the curvature vanishes but the curve is non-degenerate is at most countable. In view of this, the curvature completely describes the non-degeneracy of planar curves. Clearly, a straight line is degenerate everywhere.

The claim is that the notion of non-degeneracy is the right description for a manifold $\mathcal{M}$ to be "sufficiently" curved in order to develop a general Khintchine type theory (both convergent and divergent cases) for $\mathcal{M} \cap W(n, \psi)$. With this in mind, the key then lies in understanding the distribution of rational points "close" to such manifolds.

6.1.2 Rational points near manifolds: the heuristics

Given a point $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and a set $A \subseteq \mathbb{R}^n$, let

$$\operatorname{dist}(\mathbf{x}, A) := \inf\{d(\mathbf{x}, \mathbf{a}) \,:\, \mathbf{a} \in A\}$$

where as usual $d(\mathbf{x}, \mathbf{a}) := \max\limits_{1\le i\le n} |x_i - a_i|$. Now let $\mathbf{x} \in \mathcal{M} \cap W(n, \psi)$. Then by definition there exist infinitely many $q \in \mathbb{N}$ and $\mathbf{p} \in \mathbb{Z}^n$ such that

$$\operatorname{dist}\Big(\mathcal{M}, \frac{\mathbf{p}}{q}\Big) \le d\Big(\mathbf{x}, \frac{\mathbf{p}}{q}\Big) < \frac{\psi(q)}{q}\,.$$

This means that the rational points

$$\frac{\mathbf{p}}{q} := \Big(\frac{p_1}{q}, \ldots, \frac{p_n}{q}\Big)$$

of interest must lie within the $\frac{\psi(q)}{q}$–neighbourhood of $\mathcal{M}$. In particular, assuming that ψ is decreasing, we have that the points $\mathbf{p}/q$ of interest with $k^{t-1} < q \le k^t$ are contained in the $\frac{\psi(k^{t-1})}{k^{t-1}}$–neighbourhood of $\mathcal{M}$. Let us denote this neighbourhood by $\Delta_k^+(t, \psi)$ and by $N_k^+(t, \psi)$ the set of rational points with $k^{t-1} < q \le k^t$ contained in $\Delta_k^+(t, \psi)$. In other words,

$$N_k^+(t, \psi) := \Big\{\mathbf{p}/q \in \mathrm{I}^n : k^{t-1} < q \le k^t \ \text{ and } \ \operatorname{dist}\big(\mathcal{M}, \mathbf{p}/q\big) \le \tfrac{\psi(k^{t-1})}{k^{t-1}}\Big\}\,. \tag{6.1}$$

Recall, that $\mathcal{M} \subseteq \mathrm{I}^n$. Hence, regarding the n-dimensional volume of the neighbourhood $\Delta_k^+(t, \psi)$, it follows that

$$m_n\Big(\Delta_k^+(t, \psi)\Big) \ \asymp\ \left(\frac{\psi(k^{t-1})}{k^{t-1}}\right)^{n-d}.$$

Now let $Q_k(t)$ denote the set of rational points with $k^{t-1} < q \le k^t$ lying in the unit cube I^n. Then,

$$\#Q_k(t) \ \asymp\ (k^t)^{n+1}$$

and if we assume that the points in $Q_k(t)$ are "fairly" distributed within I^n, we would expect that

the number of these points that fall into $\Delta_k^+(t, \psi)$
is proportional to the measure of $\Delta_k^+(t, \psi)$.

In other words and more formally, under the above distribution assumption, we would expect that

$$\#\{Q_k(t) \cap \Delta_k^+(t, \psi)\} \ \asymp\ \#Q_k(t) \times m_n\Big(\Delta_k^+(t, \psi)\Big) \tag{6.2}$$

and since the left-hand side is $\#N_k^+(t, \psi)$, we would be able to conclude that

$$\#N_k^+(t, \psi) \ \asymp\ (k^t)^{n+1}\left(\frac{\psi(k^{t-1})}{k^{t-1}}\right)^{n-d} \ \asymp\ (k^{t-1})^{d+1}\psi(k^{t-1})^{n-d}\,. \tag{6.3}$$

For the moment, let us assume that (6.2) and hence (6.3) are fact. Now

$$\begin{aligned} \mathcal{M} \cap W(n,\psi) &= \bigcap_{m=1}^{\infty} \bigcup_{t=m}^{\infty} \bigcup_{k^{t-1}<q\le k^t} \bigcup_{\mathbf{p}\in\mathbb{Z}^n:\mathbf{p}/q\in\mathrm{I}^n} B\Big(\tfrac{\mathbf{p}}{q}, \tfrac{\psi(q)}{q}\Big) \cap \mathcal{M} \\ &\subset \bigcap_{m=1}^{\infty} \bigcup_{t=m}^{\infty} A_k^+(t,\psi,\mathcal{M}) \end{aligned}$$

where

$$A_k^+(t,\psi,\mathcal{M}) := \bigcup_{k^{t-1}<q\le k^t} \bigcup_{\mathbf{p}\in\mathbb{Z}^n:\mathbf{p}/q\in\mathrm{I}^n} B\Big(\tfrac{\mathbf{p}}{q}, \tfrac{\psi(k^{t-1})}{k^{t-1}}\Big) \cap \mathcal{M}\,.$$

It is easily verified that

$$\begin{aligned} |A_k^+(t,\psi,\mathcal{M})|_{\mathcal{M}} &\le \sum_{k^{t-1}<q\le k^t} \sum_{\mathbf{p}\in\mathbb{Z}^n:\mathbf{p}/q\in\mathrm{I}^n} \underbrace{\Big|B\Big(\tfrac{\mathbf{p}}{q}, \tfrac{\psi(k^{t-1})}{k^{t-1}}\Big) \cap \mathcal{M}\Big|_{\mathcal{M}}}_{\ll(\psi(k^{t-1})/k^{t-1})^d} \\ &\ll \#N_k^+(t,\psi)\ \ (\psi(k^{t-1})/k^{t-1})^d \\ &\overset{(6.3)}{\asymp} (k^{t-1})^{d+1}\psi(k^{t-1})^{n-d}(\psi(k^{t-1})/k^{t-1})^d \\ &\asymp k^{t-1}\psi(k^{t-1})^n\,. \end{aligned}$$

Hence

$$\sum_{t=1}^{\infty} |A_k^+(t,\psi,\mathcal{M})|_{\mathcal{M}} \ll \sum_{t=1}^{\infty} k^t\psi(k^t)^n \asymp \sum_{q=1}^{\infty} \psi(q)^n\,. \tag{6.4}$$

All the steps in the above argument apart from (6.2) and hence (6.3), can be turned into a rigorous proof. Indeed, the estimate (6.3) is not always true.

Exercise. Consider the circle $\mathcal{C}_{\sqrt{3}}$ in $\mathbb{R}^2$ given by the equation $x^2+y^2=3$. Prove that $\mathcal{C}$ does not contain any rational points. Next let $\psi(q)=q^{-1-\varepsilon}$ for some $\varepsilon>0$. Prove that

$$\mathcal{C}_{\sqrt{3}} \cap W(2,\psi) = \varnothing\,.$$

The upshot is that even for non-degenerate manifolds, we cannot expect the heuristic estimate (6.3) to hold for any decreasing ψ – some restriction on the rate at which ψ decreases to zero is required. On the other hand, affine subspaces of $\mathbb{R}^n$ may contain too many rational points, for instance, if $\mathcal{M}$ is a linear subspace of $\mathbb{R}^n$ with a basis of rational vectors. Of course, such manifolds are not non-degenerate.

However, *whenever the upper bound associated with the heuristic estimate* (6.3) *is true*, inequality (6.4) together with the convergence Borel-Cantelli Lemma implies that

$$|\mathcal{M}\cap W(n,\psi)|_{\mathcal{M}} = 0 \quad \text{if} \quad \sum_{q=1}^{\infty} \psi(q)^n < \infty\,.$$

This statement represents the convergent case of the 'dream' theorem for manifolds – see §6.1.3 immediately below. Note that the associated sum $\sum \psi(q)^n$ coincides with the sum appearing in Theorem 4.5 (Khintchine in $\mathbb{R}^n$) but the associated measure $| \, . \, |_{\mathcal{M}}$ is d-dimensional Lebesgue measure (induced on $\mathcal{M}$) rather than n-dimensional Lebesgue measure.

6.1.3 The Dream Theorem and its current status

The Dream Theorem. Let $\mathcal{M}$ be a non-degenerate sub-manifold of $\mathbb{R}^n$. Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function. Then

$$|\mathcal{M} \cap W(n,\psi)|_{\mathcal{M}} = \begin{cases} 0 & \text{if} \quad \sum_{q=1}^{\infty} \ \psi(q)^n < \infty \, , \\ \\ 1 & \text{if} \quad \sum_{q=1}^{\infty} \ \psi(q)^n = \infty \, . \end{cases} \tag{6.5}$$

We emphasize that the Dream Theorem is a desired statement rather than an established fact.

As we have already demonstrated, the convergence case of the Dream Theorem would follow on establishing the upper bound estimate

$$\#N_k^+(t,\psi) \ \ll \ (k^{t-1})^{d+1}\psi(k^{t-1})^{n-d} \tag{6.6}$$

for non-degenerate manifolds. Recall that the rational points of interest are given by the set

$$N_k(t,\psi) := \Big\{ \mathbf{p}/q \in \mathrm{I}^n : k^{t-1} < q \le k^t \ \text{and} \ \operatorname{dist}\big(\mathcal{M}, \mathbf{p}/q\big) \le \tfrac{\psi(q)}{q} \Big\} \, ,$$

and that $\#N_k^+(t,\psi)$ is an upper bound for $\#N_k(t,\psi)$. Obviously, a lower bound for $\#N_k(t,\psi)$ is given by $\#N_k^-(t,\psi)$ where

$$N_k^-(t,\psi) := \Big\{ \mathbf{p}/q \in \mathrm{I}^n : k^{t-1} < q \le k^t \ \text{and} \ \operatorname{dist}\big(\mathcal{M}, \mathbf{p}/q\big) \le \tfrac{\psi(k^t)}{k^t} \Big\} \, ,$$

and if ψ is k-regular (see (5.5)) then $N_k^+(t,\psi) \asymp N_k^-(t,\psi)$. In particular, whenever we are able to establish the heuristic estimate (6.3) or equivalently the upper bound estimate (6.6) together with the lower bound estimate

$$\#N_k^-(t,\psi) \ \gg \ (k^{t-1})^{d+1}\psi(k^{t-1})^{n-d} \, , \tag{6.7}$$

we would have that

$$\#N_k(t,\psi) \ \asymp \ (k^{t-1})^{d+1}\psi(k^{t-1})^{n-d} \, . \tag{6.8}$$

It is worth stressing that the lower bound estimate (6.7) is by itself not enough to prove the divergence case of the Dream Theorem. Loosely speaking, we also need to know that rational points associated with the set $N_k^-(t,\psi)$ are "ubiquitous" within the $\frac{\psi(k^t)}{k^t}$–neighbourhood of $\mathcal{M}$. Indeed, when establishing the divergence case of Khintchine's Theorem (Theorem 2.3), we trivially have the right count of k^{2t} for the number of rational points $p/q \in \mathrm{I}$ with $k^{t-1} < q \le k^t$. The crux is to establish the associated distribution type result given by

Theorem 1.3. This in turn implies that the rational points under consideration give rise to a ubiquitous system – see §5.3.1.

We now turn our attention to reality and describe various 'general' contributions towards the Dream Theorem.

- *Extremal manifolds.* A sub-manifold $\mathcal{M}$ of $\mathbb{R}^n$ is called *extremal* if

$$\left|\mathcal{M}\cap W(n,\tfrac{1+\varepsilon}{n})\right|_{\mathcal{M}} = 0 \qquad \forall\, \varepsilon > 0\,.$$

  Note that $\mathcal{M}\cap W(n,\frac{1}{n}) = \mathcal{M}$ – see Remark 6.1. In their pioneering work [67] published in 1998, Kleinbock & Margulis proved that any non-degenerate sub-manifold $\mathcal{M}$ of $\mathbb{R}^n$ is extremal. It is easy to see that this implies the convergence case of the Dream Theorem for functions of the shape

$$\psi_\varepsilon(q) := q^{-\frac{1+\varepsilon}{n}}\,.$$

  Indeed,

$$\textstyle\sum_{q=1}^{\infty}\psi_\varepsilon(q)^n = \sum_{q=1}^{\infty} q^{-(1+\varepsilon)} < \infty$$

  and so whenever the convergent case of (6.5) is fulfilled, the corresponding manifold is extremal.

- *Planar curves.* The Dream Theorem is true when $n = 2$; that is, when $\mathcal{M}$ is a non-degenerate planar curve. The convergence case of (6.5) for planar curves was established in [91] and subsequently strengthened in [30]. The divergence case of (6.5) for planar curves was established in [15].

- *Beyond planar curves.* The divergence case of the Dream Theorem is true for analytic non-degenerate sub-manifolds of $\mathbb{R}^n$ [11]. Recently, the divergence case of (6.5) has been shown to be true for non-degenerate curves and manifolds that can be 'fibred' into such curves [20]. The latter includes C^∞ non-degenerate sub-manifolds of $\mathbb{R}^n$ which are not necessarily analytic. The convergence case of the Dream Theorem is true for a large subclass of 2-non-degenerate sub-manifolds of $\mathbb{R}^n$ with dimension d strictly greater than $(n+1)/2$ [19]. Earlier, manifolds satisfying a geometric (curvature) condition were shown to satisfy the convergence case of the Dream Theorem [47].

The upshot of the above is that the Dream Theorem is in essence fact for a fairly generic class of non-degenerate sub-manifolds $\mathcal{M}$ of $\mathbb{R}^n$ apart from the case of convergence when $n \geq 3$ and $d \leq (n+1)/2$.

Remark 6.2. The theory of Diophantine approximation stems from Mahler's problem (1932) regarding the *extremality* of the Veronese curve $\mathcal{V} := \{(x, x^2, \ldots, x^n) : x \in \mathbb{R}^n\}$. Following a substantial number of partial results (initially for $n = 2$, then $n = 3$ and some for higher n), a complete solution to the problem was given by Sprindžuk in 1965. For a historical account of the manifold theory we refer the reader to the monographs [31, 90] and the introduction given in the paper [15].

Remark 6.3. Note that in view of the Khintchine's Transference Principle, we could have easily defined extremality via the dual form of Diophantine approximation (see Remark 4.18); namely, $\mathcal{M}$ is extremal if

$$|\mathcal{M} \cap W^*(n, n+\varepsilon)|_{\mathcal{M}} = 0 \qquad \forall\, \varepsilon > 0\,.$$

The point is that both definitions are equivalent. This is not the case in the inhomogeneous setup considered in §6.3.1.

Remark 6.4. It is worth mentioning that in [67], Kleinbock & Margulis established a stronger (multiplicative) form of extremality (see §6.4.1 below) that settled the Baker-Sprindžuk Conjecture from the eighties. Not only did their work solve a long-standing fundamental problem, but it also developed new techniques utilising the link between Diophantine approximation and homogeneous dynamics. Without doubt the work of Kleinbock & Margulis has been the catalyst for the subsequent contributions towards the Dream Theorem described above.

6.2 The Hausdorff theory for manifolds

The goal is to obtain a Jarník type theorem that describes the Hausdorff measure $\mathcal{H}^s$ of the set $\mathcal{M} \cap W(n,\psi)$ of simultaneously ψ–approximable points lying on $\mathcal{M}$. In other words, we wish to obtain a Hausdorff measure version of the Dream Theorem. In view of this, by default, we consider approximating functions ψ which decrease sufficiently rapidly so that the d-dimensional Lebesgue measure of $\mathcal{M} \cap W(n,\psi)$ is zero. Now, as the example in §6.1.2 demonstrates, in order to obtain a coherent Hausdorff measure theory we must impose some restriction on the rate at which ψ decreases. Indeed, with reference to that example, the point is that $\mathcal{H}^s(\mathcal{C}_{\sqrt{3}} \cap W(2, 1+\varepsilon)) = 0$ irrespective of $\varepsilon > 0$ and the measure $\mathcal{H}^s$. On the other hand, for the unit circle $\mathcal{C}_1$ in $\mathbb{R}^2$ given by the equation $x^2 + y^2 = 1$, it can be shown [14, Theorem 19] that for any $\varepsilon > 0$

$$\mathcal{H}^s(\mathcal{C}_1 \cap W(2, 1+\varepsilon)) = \infty \quad \text{with} \quad s = \tfrac{1}{2+\varepsilon}\,.$$

Nevertheless, it is believed that if the rate of decrease of ψ is 'close' to the approximating function $q^{-1/n}$ associated with Dirichlet's Theorem, then the behaviour of $\mathcal{H}^s(\mathcal{M} \cap W(n,\psi))$ can be captured by a single, general criterion. In the following statement, the condition on ψ is captured in terms of the deviation of $\mathcal{H}^s$ from d-dimensional Lebesgue measure.

The Hausdorff Dream Theorem. Let $\mathcal{M}$ be a non-degenerate sub-manifold of $\mathbb{R}^n$, $d := \dim \mathcal{M}$ and $m := \operatorname{codim} \mathcal{M}$. Thus, $d+m = n$. Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function. Then, for any $s \in \left(\frac{m}{m+1}d, d\right)$

$$\mathcal{H}^s(\mathcal{M} \cap W(n,\psi)) = \begin{cases} 0 & \text{if} \quad \sum\limits_{q=1}^{\infty} \psi^{s+m}(q) q^{-s+d} < \infty\,, \\ \infty & \text{if} \quad \sum\limits_{q=1}^{\infty} \psi^{s+m}(q) q^{-s+d} = \infty. \end{cases} \tag{6.9}$$

We emphasize that the above is a desired statement rather than an established fact.

We now turn our attention to reality and describe various 'general' contributions towards the Hausdorff Dream Theorem.

- *Planar curves.* As with the Dream Theorem, the convergence case of (6.9) for planar curves ($n = 2, d = m = 1$) was established in [91] and subsequently strengthened in [30]. The divergence case of (6.9) for planar curves was established in [15].
- *Beyond planar curves.* The divergence case of the Hausdorff Dream Theorem is true for analytic non-degenerate sub-manifolds of $\mathbb{R}^n$ [11]. The convergence case is rather fragmented. To the best of our knowledge, the partial results obtained in [19, Corollaries 3 & 5] for 2-non-degenerate sub-manifolds of $\mathbb{R}^n$ with dimension d strictly greater than $(n+1)/2$, represent the first significant coherent contribution towards the convergence case.

Exercise. Prove the convergent case of (6.9) assuming the heuristic estimate (6.3) for the number of rational points near $\mathcal{M}$ – see §6.1.2.

Remark 6.5. Regarding the divergence case of (6.9), it is tempting to claim that it follows from the divergence case of the (Lebesgue) Dream Theorem via the Mass Transference Principle introduced in §3.4. After all, this is true when $\mathcal{M} = \mathrm{I}^n$; namely that Khintchine's Theorem implies Jarník's Theorem as demonstrated in §3.4.1. However, this is far from the truth within the context of manifolds. The reason for this is simple. With respect to the setup of the Mass Transference Principle, the set Ω that supports the $\mathcal{H}^\delta$-measure (with $\delta = \dim \mathcal{M}$) is the manifold $\mathcal{M}$ itself and is embedded in $\mathbb{R}^n$. The set $\mathcal{M} \cap W(n,\psi) \subset \Omega$ of interest can be naturally expressed as the intersection with $\mathcal{M}$ of the lim sup set arising from balls $B(\frac{\mathbf{p}}{q}, \frac{\psi(q)}{q})$ centred at rational points $\mathbf{p}/q \in \mathbb{R}^n$. However, the centre of these balls do not necessarily lie in the support of the measure $\Omega = \mathcal{M}$ and this is where the problem lies. A prerequisite for the framework of the Mass Transference Principle is that $\{B_i\}_{i\in\mathbb{N}}$ is a sequence of balls in Ω.

6.3 Inhomogeneous Diophantine approximation

When considering the well approximable sets $W(n,\psi)$ or indeed the badly approximable sets $\mathbf{Bad}(i_1,\dots,i_n)$, we are in essence investigating the behaviour of the fractional part of $q\mathbf{x}$ about the origin as q runs through $\mathbb{N}$. Clearly, we could consider the setup in which we investigate the behaviour of the orbit of $\{q\mathbf{x}\}$ about some other point. With this in mind, given $\psi : \mathbb{N} \to \mathbb{R}^+$ and a fixed point $\boldsymbol{\gamma} = (\gamma_1, \dots \gamma_n) \in \mathbb{R}^n$, let

$$W_{\boldsymbol{\gamma}}(n,\psi) := \{\mathbf{x} \in \mathrm{I}^n : \|q\mathbf{x} - \boldsymbol{\gamma}\| < \psi(q) \text{ for infinitely many } q \in \mathbb{N}\}$$

denote the *inhomogeneous* set of *simultaneously ψ-well approximable* points $\mathbf{x} \in \mathrm{I}^n$. Thus, a point $\mathbf{x} \in W_{\boldsymbol{\gamma}}(n,\psi)$ if there exist infinitely many 'shifted' rational points

$$\Big(\frac{p_1 - \gamma_1}{q}, \dots, \frac{p_n - \gamma_n}{q}\Big)$$

with $q > 0$, such that the inequalities

$$|x_i - (p_i - \gamma_i)/q| \ < \ \psi(q)/q$$

are simultaneously satisfied for $1 \le i \le n$. The following is the natural generalisation of the simultaneous Khintchine-Jarník theorem to the inhomogeneous setup. For further details, see [13, 14] and references within.

Theorem 6.1 (Inhomogeneous Khintchine-Jarník)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function, $\boldsymbol{\gamma} \in \mathbb{R}^n$ and $s \in (0, n]$. Then*

$$\mathcal{H}^s(W_{\boldsymbol{\gamma}}(n,\psi)) = \begin{cases} 0 & \textit{if} \quad \displaystyle\sum_{r=1}^{\infty} r^{n-s}\psi(r)^s < \infty\,, \\[4ex] \mathcal{H}^s(\mathrm{I}^n) & \textit{if} \quad \displaystyle\sum_{r=1}^{\infty} r^{n-s}\psi(r)^s = \infty \ . \end{cases}$$

Remark 6.6. For the sake of completeness we state the inhomogeneous analogue of Hurwitz's Theorem due to Khintchine [62, §10.10]: *for any irrational $x \in \mathbb{R}$, $\gamma \in \mathbb{R}$ and $\varepsilon > 0$, there exist infinitely many integers $q > 0$ such that*

$$q\,\|qx - \gamma\| \le (1+\varepsilon)/\sqrt{5}\,.$$

Note that presence of the ε term means that the inhomogeneous statement is not quite as sharp as the homogeneous one (i.e. when $\gamma = 0$). Also, for obvious reasons, in the inhomogeneous situation it is necessary to exclude the case that x is rational.

We now swiftly move on to the inhomogeneous theory for manifolds. In short, the heuristics of §6.1.2, adapted to the inhomogeneous setup, gives evidence towards the following natural generalisation of the Dream Theorem.

The Inhomogeneous Dream Theorem. Let $\mathcal{M}$ be a non-degenerate sub-manifold of $\mathbb{R}^n$. Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $\boldsymbol{\gamma} \in \mathbb{R}^n$. Then

$$|\mathcal{M} \cap W_{\boldsymbol{\gamma}}(n,\psi)|_{\mathcal{M}} = \begin{cases} 0 & \text{if} \quad \sum_{q=1}^{\infty} \ \psi(q)^n < \infty\ , \\[3ex] 1 & \text{if} \quad \sum_{q=1}^{\infty} \ \psi(q)^n = \infty\ . \end{cases}$$

Regarding what is known, the current state of knowledge is absolutely in line with the homogeneous situation. The inhomogeneous analogue of the extremality result of Kleinbock & Margulis [67] is established in [24, 26]. We will return to this in §6.3.1 below. For planar curves, the Inhomogeneous Dream Theorem is established in [18]. Beyond planar curves, the results in [19, 20] are obtained within the inhomogeneous framework. So in summary, the Inhomogeneous Dream Theorem is in essence fact for non-degenerate sub-manifolds $\mathcal{M}$ of $\mathbb{R}^n$ apart from the case of convergence when $n \ge 3$ and $d \le (n+1)/2$.

6.3.1 Inhomogeneous extremality and a transference principle

First we need to decide on what precisely we mean by inhomogeneous extremality. With this in mind, a manifold $\mathcal{M}$ is said to be *simultaneously inhomogeneously extremal* (SIE for short) if for every $\boldsymbol{\gamma} \in \mathbb{R}^n$,

$$\left|\mathcal{M} \cap W_{\boldsymbol{\gamma}}(n, \tfrac{1+\varepsilon}{n})\right|_{\mathcal{M}} = 0 \qquad \forall\, \varepsilon > 0\,. \tag{6.10}$$

On the other hand, a manifold $\mathcal{M}$ is said to be *dually inhomogeneously extremal* (DIE for short) if for every $\gamma \in \mathbb{R}$,

$$\left|\mathcal{M} \cap W^*_{\gamma}(n, n+\varepsilon)\right|_{\mathcal{M}} = 0 \qquad \forall\, \varepsilon > 0\,.$$

Here, given $\tau > 0$ and a fixed point $\gamma \in \mathbb{R}$, $W^*_\gamma(n,\tau)$ is the *inhomogeneous* set of *dually τ-well approximable* points consisting of points $\mathbf{x} \in \mathrm{I}^n$ for which the inequality

$$\|\mathbf{q}\cdot\mathbf{x} - \gamma\| < |\mathbf{q}|^{-\tau}$$

holds for infinitely many $\mathbf{q} \in \mathbb{Z}^n$. Moreover, a manifold $\mathcal{M}$ is simply said to be *inhomogeneously extremal* if it is both SIE and DIE.

As mentioned in Remark 6.3, in the homogeneous case ($\boldsymbol{\gamma}$=0) the simultaneous and dual forms of extremality are equivalent. Recall that this is a simply consequence of Khintchine's Transference Principle (Theorem 4.13). However, in the inhomogeneous case, there is no classical transference principle that allows us to deduce SIE from DIE and vice versa. The upshot is that the two forms of inhomogeneous extremality have to be treated separately. It turns out that establishing the dual form of inhomogeneous extremality is technically far more complicated than establishing the simultaneous form [26]. The framework developed in [24] naturally incorporates both forms of inhomogeneous extremality and indeed other stronger (multiplicative) notions associated with the inhomogeneous analogue of the Baker-Sprindžuk Conjecture.

Conjecture. *Let $\mathcal{M}$ be a non-degenerate sub-manifold of $\mathbb{R}^n$. Then $\mathcal{M}$ is inhomogeneously extremal.*

The proof given in [24] of this inhomogeneous conjecture relies very much on the fact that we know that the homogeneous statement is true. In particular, the general inhomogeneous transference principle of [24, §5] enables us to establish the following transference for non-degenerate manifolds:

$$\mathcal{M} \text{ is extremal} \iff \mathcal{M} \text{ is inhomogeneously extremal.} \tag{6.11}$$

Clearly, this enables us to conclude that:

$$\mathcal{M} \text{ is SIE} \iff \mathcal{M} \text{ is DIE.}$$

In other words, a transference principle between the two forms of inhomogeneous extremality does exist at least for the class of non-degenerate manifolds.

Trivially, inhomogeneous extremality implies (homogeneous) extremality. Thus, the main substance of (6.11) is the reverse implication. This rather surprising fact relies on the fact that the inhomogeneous lim sup sets $\mathcal{M} \cap W_{\boldsymbol{\gamma}}(n, \frac{1+\varepsilon}{n})$ and the induced measure $| \, . \, |_{\mathcal{M}}$ on non-degenerate manifolds satisfy the *intersection property* and the *contracting property* described in [24, §5]. These properties are at the heart of the Inhomogeneous Transference Principle [24, Theorem 5] that enables us to transfer zero measure statements for homogeneous lim sup sets to inhomogeneous lim sup sets. The general setup, although quite natural, is rather involved and will not be reproduced in these notes. Instead, we refer the reader to the papers [24, 26]. We advise the reader to first look at [26] in which the easier statement

$$\mathcal{M} \text{ is extremal} \quad \Longrightarrow \quad \mathcal{M} \text{ is SIE} \tag{6.12}$$

is established. This has the great advantage of bringing to the forefront the main ideas of [24] while omitting the abstract and technical notions that come with describing the inhomogeneous transference principle in all its glory. In order to illustrate the basic line of thinking involved in establishing (6.12) and indeed (6.11) we shall prove the following statement concerning extremality on $\mathrm{I} = [0, 1]$:

$$m(W(1+\varepsilon)) = 0 \quad \Longrightarrow \quad m(W_\gamma(1+\varepsilon)) = 0 \;\; \forall \, \varepsilon > 0. \tag{6.13}$$

Of course it is easy to show that the inhomogeneous set $W_\gamma(1+\varepsilon)$ is of zero Lebesgue measure m by using the convergence Borel-Cantelli Lemma. However, the point here is to develop an argument that exploits the fact that we know the homogeneous set $W_0(1+\varepsilon) := W(1+\varepsilon)$ is of zero Lebesgue measure.

To prove (6.13), we make use of the fact that $W_\gamma(1+\varepsilon)$ is a lim sup set given by

$$W_\gamma(1+\varepsilon) \;=\; \bigcap_{s=1}^{\infty} \bigcup_{q=s}^{\infty} \bigcup_{p \in \mathbb{Z}} B^\gamma_{p,q}(\varepsilon) \cap \mathrm{I}\,, \tag{6.14}$$

where, given $q \in \mathbb{N}$, $p \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ and $\varepsilon > 0$

$$B^\gamma_{p,q}(\varepsilon) := \{\, y \in \mathbb{R} : |qy + p + \gamma| < |q|^{-1-\varepsilon} \,\}\,.$$

As usual, if $B = B(x, r)$ denotes the ball (interval) centred at x and of radius $r > 0$, then it is easily seen that

$$B^\gamma_{p,q}(\varepsilon) = B\Big(\tfrac{p+\gamma}{q}, |q|^{-2-\varepsilon}\Big)\,.$$

Now we consider 'blown up' balls $B^\gamma_{p,q}(\varepsilon/2)$ and observe that Lebesgue measure m satisfies the following contracting property: for any choice $q \in \mathbb{N}$, $p \in \mathbb{Z}$, $\gamma \in \mathbb{R}$ and $\varepsilon > 0$ we have that

$$m\Big(B^\gamma_{p,q}(\varepsilon)\Big) \;=\; \frac{2}{q^{2+\varepsilon}} \;=\; q^{-\frac{\varepsilon}{2}} \frac{2}{q^{2+(\varepsilon/2)}} \;=\; q^{-\frac{\varepsilon}{2}} \; m\Big(B^\gamma_{p,q}(\varepsilon/2)\Big)\,. \tag{6.15}$$

Next we separate the balls $B^\gamma_{p,q}(\varepsilon)$ into classes of disjoint and non-disjoint balls. Fix $q \in \mathbb{N}$ and $p \in \mathbb{Z}$. Clearly, there exists a unique integer $t = t(q)$ such that $2^t \le q < 2^{t+1}$. The ball $B^\gamma_{p,q}(\varepsilon)$ is said to be *disjoint* if for every $q' \in \mathbb{N}$ with $2^t \le q' < 2^{t+1}$ and every $p' \in \mathbb{Z}$

$$B^\gamma_{p,q}(\varepsilon/2) \cap B^\gamma_{p',q'}(\varepsilon/2) \cap \mathrm{I} = \varnothing\,.$$

Otherwise, the ball $B^{\gamma}_{p,q}(\varepsilon/2)$ is said to be *non-disjoint*. This notion of disjoint and non-disjoint balls enables us to decompose the $W_\gamma(1+\varepsilon)$ into the two limsup subsets:

$$D^{\gamma}(\varepsilon) \;:=\; \bigcap_{s=0}^{\infty}\bigcup_{t=s}^{\infty}\bigcup_{2^t\le|q|<2^{t+1}}\bigcup_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is disjoint}}} B^{\gamma}_{p,q}(\varepsilon)\cap \mathrm{I}\,,$$

and

$$N^{\gamma}(\varepsilon) \;:=\; \bigcap_{s=0}^{\infty}\bigcup_{t=s}^{\infty}\bigcup_{2^t\le|q|<2^{t+1}}\bigcup_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is non-disjoint}}} B^{\gamma}_{p,q}(\varepsilon)\cap \mathrm{I}\,.$$

Formally,

$$W_\gamma(1+\varepsilon) \;=\; \bigcap_{s=1}^{\infty}\bigcup_{q=s}^{\infty}\bigcup_{p\in\mathbb{Z}} B^{\gamma}_{p,q}(\varepsilon)\cap \mathrm{I} \;=\; D^{\gamma}(\varepsilon)\;\cup\;N^{\gamma}(\varepsilon)\;.$$

We now show that $m(D^{\gamma}(\varepsilon)) = 0 = m(N^{\gamma}(\varepsilon))$. This would clearly imply (6.13). Naturally, we deal with the disjoint and non-disjoint sets separately.

The disjoint case: By the definition of disjoint balls, for every fixed t we have that

$$\begin{aligned}\sum_{2^t\le q<2^{t+1}}\sum_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is disjoint}}} m(B^{\gamma}_{p,q}(\varepsilon/2)\cap \mathrm{I}) \;&=\; m\Big(\bigcup_{2^t\le q<2^{t+1}}\bigcup_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is disjoint}}} B^{\gamma}_{p,q}(\varepsilon/2)\cap \mathrm{I}\Big)\\ &\le\; m(\mathrm{I}) \;=\; 1.\end{aligned}$$

This together with the contracting property (6.15) of the measure m, implies that

$$\begin{aligned} m\Big(\bigcup_{2^t\le q<2^{t+1}}\bigcup_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is disjoint}}} B^{\gamma}_{p,q}(\varepsilon)\cap \mathrm{I}\Big) \;&=\; \sum_{2^t\le q<2^{t+1}}\sum_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is disjoint}}} m(B^{\gamma}_{p,q}(\varepsilon)\cap \mathrm{I})\\ &\le\; \sum_{2^t\le q<2^{t+1}}\sum_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is disjoint}}} q^{-\frac{\varepsilon}{2}}\;\; m(B^{\gamma}_{p,q}(\varepsilon/2)\cap \mathrm{I})\\ &\le\; 2^{-t\frac{\varepsilon}{2}}\sum_{2^t\le q<2^{t+1}}\sum_{\substack{p\in\mathbb{Z}\\ B^{\gamma}_{p,q}(\varepsilon)\text{ is disjoint}}} m(B^{\gamma}_{p,q}(\varepsilon/2)\cap \mathrm{I})\\ &\le\; 2^{-t\frac{\varepsilon}{2}}\,.\end{aligned}$$

Since $\sum_{t=1}^{\infty} 2^{-t\frac{\varepsilon}{2}} < \infty$, the convergence Borel-Cantelli Lemma implies that

$$m(D^{\gamma}(\varepsilon)) = 0\,.$$

The non-disjoint case: Let $B^{\gamma}_{p,q}(\varepsilon)$ be a non-disjoint ball and let $t = t(q)$ be as above. Clearly

$$B^{\gamma}_{p,q}(\varepsilon) \subset B^{\gamma}_{p,q}(\varepsilon/2)\,.$$

By the definition of non-disjoint balls, there is another ball $B^{\gamma}_{p',q'}(\varepsilon/2)$ with $2^t \leq q < 2^{t+1}$ such that

$$B^{\gamma}_{p,q}(\varepsilon/2) \cap B^{\gamma}_{\mathbf{p}',q'}(\varepsilon/2) \cap \mathrm{I} \neq \varnothing\,. \tag{6.16}$$

It is easily seen that $q' \neq q$, as otherwise we would have that $B^{\gamma}_{p,q}(\varepsilon/2) \cap B^{\gamma}_{p',q}(\varepsilon/2) = \varnothing$. The point here is that rationals with the same denominator q are separated by $1/q$. Take any point y in the non-empty set appearing in (6.16). By the definition of $B^{\gamma}_{p,q}(\varepsilon/2)$ and $B^{\gamma}_{p',q'}(\varepsilon/2)$, it follows that

$$|qy + p + \gamma| \;<\; q^{-1-\frac{\varepsilon}{2}} \;\leq\; 2^{t(-1-\frac{\varepsilon}{2})}$$

and

$$|q'y + p' + \gamma| \;<\; (q')^{-1-\frac{\varepsilon}{2}} \;\leq\; 2^{t(-1-\frac{\varepsilon}{2})}\,.$$

On combining these inequalities in the obvious manner and assuming without loss of generality that $q > q'$, we deduce that

$$|\,\underbrace{(q-q')}_{q''}\,y + \underbrace{(p-p')}_{p''}\,| \;<\; 2\cdot 2^{t(-1-\frac{\varepsilon}{2})} \;<\; 2^{(t+2)(-1-\frac{\varepsilon}{3})} \tag{6.17}$$

for all t sufficiently large. Furthermore, $0 < q'' \leq 2^{t+2}$ which together with (6.17) yields that

$$|q''y + p''| \;<\; (q'')^{-1-\frac{\varepsilon}{3}}\,.$$

If the latter inequality holds for infinitely many different $q'' \in \mathbb{N}$, then $y \in W(1+\varepsilon/3)$. Otherwise, there is a fixed pair $(p'', q'') \in \mathbb{Z} \times \mathbb{N}$ such that (6.17) is satisfied for infinitely many t. Thus, we must have that $q''y + p'' = 0$ and so y is a rational point. The upshot of the non-disjoint case is that

$$N^{\gamma}(\varepsilon) \;\subset\; W(1+\varepsilon/3) \;\cup\; \mathbb{Q}\,.$$

However, we are given that the homogeneous set $W(1+\varepsilon/3)$ is of measure zero and since $\mathbb{Q}$ is countable, it follows that

$$m(N^{\gamma}(\varepsilon)) \;=\; 0\,.$$

This completes the proof of (6.13).

6.4 The inhomogeneous multiplicative theory

For completeness, we include a short section surveying recent striking developments in the theory of inhomogeneous multiplicative Diophantine approximation. Nevertheless, we start by highlighting the fact that there remain gapping holes in the theory.

Given $\psi : \mathbb{N} \to \mathbb{R}^+$ and a fixed point $\boldsymbol{\gamma} = (\gamma_1, \dots \gamma_n) \in \mathbb{R}^n$, let

$$W^\times_{\boldsymbol{\gamma}}(n,\psi) := \{\mathbf{x} \in \mathrm{I}^n \colon \|qx_1 - \gamma_1\| \, \dots \, \|qx_n - \gamma_n\| < \psi(q) \text{ for infinitely many } q \in \mathbb{N}\} \quad (6.18)$$

denote the *inhomogeneous* set of *multiplicatively ψ-well approximable* points $\mathbf{x} \in \mathrm{I}^n$. When $\boldsymbol{\gamma} = \{\mathbf{0}\}$, the corresponding set $W^\times_{\boldsymbol{\gamma}}(n,\psi)$ naturally coincides with the homogeneous set $W^\times(n,\psi)$ given by (4.10) in §4.4. It is natural to ask for an inhomogeneous generalisation of Gallagher's Theorem (§4.4, Theorem 4.6). A straightforward 'volume' argument making use of the lim sup nature of $W^\times_{\boldsymbol{\gamma}}(n,\psi)$, together with the convergence Borel-Cantelli Lemma implies the following statement.

Lemma 6.1 (Inhomogeneous Gallagher: convergence)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $\boldsymbol{\gamma} \in \mathbb{R}^n$. Then*

$$m_n(W^\times_{\boldsymbol{\gamma}}(n,\psi)) = 0 \quad if \quad \sum_{q=1}^{\infty} \psi(q) \log^{n-1} q < \infty \, .$$

The context of Remark 4.9 remains valid in the inhomogeneous setup; namely, we can remove the condition that ψ is monotonic, if we replace the above convergence sum condition by $\sum \psi(q) |\log \psi(q)|^{n-1} < \infty$.

Surprisingly, the divergence counterpart of Lemma 6.1 is not known.

Conjecture 6.1 (Inhomogeneous Gallagher: divergence)**.** *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $\boldsymbol{\gamma} \in \mathbb{R}^n$. Then*

$$m_n(W^\times_{\boldsymbol{\gamma}}(n,\psi)) = 1 \quad if \quad \sum_{q=1}^{\infty} \psi(q) \log^{n-1} q = \infty \, .$$

Restricting our attention to $n = 2$, it is shown in [17, Theorem 13] that the conjecture is true if given $\boldsymbol{\gamma} = (\gamma_1, \gamma_2) \in \mathbb{R}^2$, either $\gamma_1 = 0$ or $\gamma_2 = 0$. In other words, we are able to deal with the situation in which one of the two "approximating quantities" is inhomogeneous but not both. For further details see [17, §2.2].

We now turn our attention to the Hausdorff theory. Given that the Lebesgue theory is so incomplete, it would be reasonable to have low expectations for a coherent Hausdorff theory. However, when $n = 2$, we are bizarrely in pretty good shape. To begin with note that

$$\text{if } \ s \le 1 \ \text{ then } \ \mathcal{H}^s(W^\times_{\boldsymbol{\gamma}}(2,\psi)) = \infty \ \text{ irrespective of approximating function } \ \psi. \qquad (6.19)$$

To see this, given $\boldsymbol{\gamma} = (\gamma_1, \gamma_2) \in \mathbb{R}^2$, we observe that for any $\alpha \in W_{\gamma_1}(1,\psi)$ the whole line $x_1 = \alpha$ within the unit interval is contained in $W^\times_{\boldsymbol{\gamma}}(2,\psi)$. Hence,

$$W_{\gamma_1}(1,\psi) \times \mathrm{I} \subset W^\times_{\boldsymbol{\gamma}}(2,\psi) \, . \qquad (6.20)$$

It is easy to verify that $W_{\gamma_1}(1,\psi)$ is an infinite set for any approximating function ψ and so (6.20) implies (6.19). Thus, when considering the s-dimensional Hausdorff measure of $W^\times_{\boldsymbol{\gamma}}(2,\psi)$, there is no loss of generality in assuming that $s \in (1,2]$. The following inhomogeneous multiplicative analogue of Jarník's theorem is established in [28, Theorem 1].

Theorem 6.2. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function, $\boldsymbol{\gamma} \in \mathbb{R}^2$ and $s \in (1,2)$. Then*

$$\mathcal{H}^s\big(W_{\boldsymbol{\gamma}}^{\times}(2,\psi)\big) \ = \ \begin{cases} 0 & \text{if} \ \ \sum_{q=1}^{\infty} q^{2-s}\psi^{s-1}(q) \ < \infty \ , \\ \\ \infty & \text{if} \ \ \sum_{q=1}^{\infty} q^{2-s}\psi^{s-1}(q) \ = \infty \ . \end{cases} \tag{6.21}$$

Remark 6.7. Recall that Gallagher's multiplicative statement and its conjectured inhomogeneous generalisation (Conjecture 6.1) have the extra 'log factor' in the Lebesgue 'volume' sum compared to Khintchine's simultaneous statement (Theorem 6.1 with $s = n = 2$). A priori, it is natural to expect the log factor to appear in one form or another when determining the Hausdorff measure $\mathcal{H}^s$ of $W_{\boldsymbol{\gamma}}^{\times}(2,\psi)$ for $s \in (1,2)$. This, as we see from Theorem 6.2, is very far from the truth. The 'log factor' completely disappears. Thus, genuine 'fractal' Hausdorff measures are insensitive to the multiplicative nature of $W_{\boldsymbol{\gamma}}^{\times}(2,\psi)$.

Remark 6.8. Note that in view of the previous remark, even if we had written $\mathcal{H}^s(\mathrm{I}^2)$ instead of ∞ in the divergence case of Theorem 6.2 , it is still necessary to exclude the case $s = 2$.

For $n > 2$, the proof given in [28] of Theorem 6.2 can be adapted to show that for any $s \in (n-1, n)$

$$\mathcal{H}^s\big(W_{\boldsymbol{\gamma}}^{\times}(n,\psi)\big) \ = \ 0 \qquad \text{if} \qquad \sum_{q=1}^{\infty} q^{n-s}\psi^{s+1-n}(q)\log^{n-2} q \ < \infty \ .$$

Thus, for convergence in higher dimensions we lose a log factor from the Lebesgue volume sum appearing in Gallagher's homogeneous result and indeed Lemma 6.1. This of course is absolutely consistent with the $n = 2$ situation given by Theorem 6.2. Regarding a divergent statement, the arguments used in proving Theorem 6.2 can be adapted to show that for any $s \in (n-1, n)$

$$\mathcal{H}^s\big(W_{\boldsymbol{\gamma}}^{\times}(n,\psi)\big) \ = \ \infty \qquad \text{if} \qquad \sum_{q=1}^{\infty} q^{n-s}\psi^{s+1-n}(q) \ = \infty \ .$$

Thus, there is a discrepancy in the above 's-volume' sum conditions for convergence and divergence when $n > 2$. In view of this, it remains an interesting open problem to determine the necessary and sufficient condition for $\mathcal{H}^s\big(W_{\boldsymbol{\gamma}}^{\times}(n,\psi)\big)$ to be zero or infinite in higher dimensions.

6.4.1 The multiplicative theory for manifolds

Let $\mathcal{M}$ be a non-degenerate sub-manifolds of $\mathbb{R}^n$. In a nutshell, as in the simultaneous case, the overarching problem is to develop a Lebesgue and Hausdorff theory for $\mathcal{M} \cap W_{\boldsymbol{\gamma}}^{\times}(n,\psi)$. Given that our current knowledge for the independent theory (i.e. when $\mathcal{M} = \mathbb{R}^n$) is pretty poor, we should not expect too much in terms of the dependent (manifold) theory. We start with describing coherent aspects of the Lebesgue theory. The following is the multiplicative analogue of the statement that $\mathcal{M}$ is inhomogeneously extremal. Given $\tau > 0$ and a fixed point $\boldsymbol{\gamma} \in \mathbb{R}^n$, we write $W_{\boldsymbol{\gamma}}^{\times}(n,\tau)$ for the set $W_{\boldsymbol{\gamma}}^{\times}(n,\psi)$ with $\psi(q) = q^{-\tau}$.

Theorem 6.3. *Let $\mathcal{M}$ be a non-degenerate sub-manifold of $\mathbb{R}^n$. Then*

$$\left|\mathcal{M} \cap W_{\boldsymbol{\gamma}}^{\times}(n, 1+\varepsilon)\right|_{\mathcal{M}} = 0 \qquad \forall\, \varepsilon > 0\,.$$

In the homogeneous case, the above theorem is due to Kleinbock & Margulis [67] and implies that non-degenerate manifolds are *strongly extremal* (by definition). It is easily seen that strongly extremal implies extremal. The inhomogeneous statement is established via the general Inhomogeneous Transference Principle developed in [24].

Beyond strong extremality, we have the following convergent statement for the Lebesgue measure of $\mathcal{M} \cap W_{\boldsymbol{\gamma}}^{\times}(n, \psi)$ in the case $\mathcal{M}$ is a planar curve $\mathcal{C}$.

Theorem 6.4. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function and $\boldsymbol{\gamma} \in \mathbb{R}^n$. Let $\mathcal{C}$ be a non-degenerate planar curve. Then*

$$\left|\mathcal{C} \cap W_{\boldsymbol{\gamma}}^{\times}(2, \psi)\right|_{\mathcal{C}} = 0 \qquad \textit{if} \qquad \sum_{q=1}^{\infty} \psi(q) \log q \; < \infty\,. \tag{6.22}$$

The homogeneous case is established in [5, Theorem 1]. However, on making use of the upper bound counting estimate appearing within Theorem 2 of [18], it is easy to adapt the homogeneous proof to the inhomogeneous setup. The details are left as an *exercise*. Just as in the homogeneous theory, obtaining the counterpart divergent statement for the Lebesgue measure of $\mathcal{C} \cap W_{\boldsymbol{\gamma}}^{\times}(2, \psi)$ remains a stubborn problem. However, for genuine fractal Hausdorff measures $\mathcal{H}^s$ we have a complete convergence/divergence result [28, Theorem 2].

Theorem 6.5. *Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic function, $\boldsymbol{\gamma} \in \mathbb{R}^n$ and $s \in (0, 1)$. Let $\mathcal{C}$ be a $C^{(3)}$-planar curve with non-zero curvature everywhere apart from a set of s-dimensional Hausdorff measure zero. Then*

$$\mathcal{H}^s\big(\mathcal{C} \cap W_{\boldsymbol{\gamma}}^{\times}(2, \psi)\big) \;=\; \begin{cases} 0 & \text{if} \;\; \sum_{q=1}^{\infty} q^{1-s}\psi^s(q) \; < \infty, \\[2ex] \infty & \text{if} \;\; \sum_{q=1}^{\infty} q^{1-s}\psi^s(q) \; = \infty. \end{cases}$$

It is evident from the proof of the divergence case of the above theorem [28, §2.1.3], that imposing the condition that $\mathcal{C}$ is a $C^{(1)}$-planar curve suffices.

Beyond planar curves, the following lower bound dimension result represents the current state of knowledge.

Theorem 6.6. *Let $\mathcal{M}$ be an arbitrary Lipschitz manifold in $\mathbb{R}^n$ and $\boldsymbol{\gamma} \in \mathbb{R}^n$. Then, for any $\tau \geq 1$*

$$\dim\big(\mathcal{M} \cap W_{\boldsymbol{\gamma}}^{\times}(n, \tau)\big) \;\geq\; \dim \mathcal{M} - 1 + \frac{2}{1+\tau}\,. \tag{6.23}$$

The homogeneous case is established in [23, Theorem 5]. The homogeneous proof [23, §6.2] rapidly reduces to the inequality

$$\dim\big(\mathcal{M} \cap W_{\mathbf{0}}^{\times}(n, \tau)\big) \;\geq\; \dim \mathcal{M} - 1 + \dim W_0^{\times}(1, \tau)\,.$$

But $W_0^\times(1,\tau) := W(1,\tau)$ and the desired statement follows on applying the Jarník-Besicovitch Theorem (Theorem 3.2). Now, Theorem 6.1 implies that the inhomogeneous generalisation of the Jarník-Besicovitch Theorem is valid; namely that, for any $\gamma \in \mathbb{R}$ and $\tau \geq 1$

$$\dim W_\gamma(1,\tau) \; = \; \frac{2}{1+\tau}\,.$$

Thus, the short argument given in [23, §6.2] can be adapted in the obvious manner to establish Theorem 6.6.

6.4.2 Cassels' problem

A straightforward consequence of Theorem 6.1 with $s = 2$ (inhomogeneous Khintchine), is that for any $\boldsymbol{\gamma} = (\gamma_1, \gamma_2) \in \mathbb{R}^2$, the set

$$W_{\boldsymbol{\gamma}}^\times := \{\mathbf{x} \in \mathrm{I}^2 \colon \liminf_{q\to\infty} q\,\|qx_1 - \gamma_1\|\,\|qx_2 - \gamma_2\| \;=\; 0\} \tag{6.24}$$

is of full Lebesgue measure; i.e. for any $\boldsymbol{\gamma} \in \mathbb{R}^2$, we have that

$$m_2(W_{\boldsymbol{\gamma}}^\times) \;=\; 1\,.$$

Of course, one can actually deduce the stronger 'fiber' statement that for any $x \in \mathrm{I}$ and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2) \in \mathbb{R}^2$, the set

$$\{y \in \mathrm{I}\colon \liminf_{q\to\infty} q\,\|qx - \gamma_1\|\,\|qy - \gamma_2\| \;=\; 0\}$$

is of full Lebesgue measure. In a beautiful paper [89], Shapira establishes the following statement which solves a problem of Cassels dating back to the fifties.

Theorem 6.7 (U. Shapira)**.**

$$m_2\Big(\bigcap_{\boldsymbol{\gamma}\in\mathbb{R}^2} W_{\boldsymbol{\gamma}}^\times\Big) = 1\,.$$

Thus, almost any pair of real numbers $(x_1, x_2) \in \mathbb{R}^2$ satisfies

$$\forall\ (\gamma_1, \gamma_2) \in \mathbb{R}^2 \qquad \liminf_{q\to\infty} q\,\|qx_1 - \gamma_1\|\,\|qx_2 - \gamma_2\| \;=\; 0\,. \tag{6.25}$$

In fact, Cassels asked for the existence of just one pair (x_1, x_2) satisfying (6.25). Furthermore, Shapira showed that if $1, x_1, x_2$ form a basis for a totally real cubic number field, then (x_1, x_2) satisfies (6.25). On the other hand, if $1, x_1, x_2$ are linearly dependent over $\mathbb{Q}$, then (x_1, x_2) cannot satisfy (6.25).

Most recently, Gorodnik & Vishe [55] have strengthened Shapira's result in the following manner: *almost any pair of real numbers* $(x_1, x_2) \in \mathbb{R}^2$ *satisfies*

$$\forall\ (\gamma_1, \gamma_2) \in \mathbb{R}^2 \qquad \liminf_{q\to\infty} q\ \log_5 q\|qx_1 - \gamma_1\|\,\|qx_2 - \gamma_2\| \;=\; 0\,,$$

where $\log_5$ is the fifth iterate of log. This 'rate' result makes a contribution towards the following open problem.

Conjecture 6.2. *Almost any pair of real numbers* $(x_1, x_2) \in \mathbb{R}^2$ *satisfies*

$$\forall\ (\gamma_1, \gamma_2) \in \mathbb{R}^2 \qquad \liminf_{q\to\infty} q\ \log q\,\|qx_1 - \gamma_1\|\,\|qx_2 - \gamma_2\| < \infty\,. \tag{6.26}$$

Remark 6.9. It is relatively straightforward to show (*exercise*) that for any $\tau > 2$

$$\left\{\mathbf{x} \in \mathrm{I}^2 : \forall\ (\gamma_1, \gamma_2) \in \mathbb{R}^2 \quad \liminf_{q\to\infty} q\ \log^{\tau} q\,\|qx_1 - \gamma_1\|\,\|qx_2 - \gamma_2\| = 0\right\} = \varnothing\,.$$

We end this section by mentioning Cassels' problem within the context of Diophantine approximation on manifolds. By exploiting the work of Shah [88], it is shown in [56] that for any non-degenerate planar curve $\mathcal{C}$

$$\Big|\ \mathcal{C} \cap \textstyle\bigcap_{\gamma\in\mathbb{R}^2} W_{\gamma}^{\times}\ \Big|_{\mathcal{C}} = 1\,.$$

7 The badly approximable theory

We have had various discussions regarding badly approximable points in earlier sections, in particular within §1.3 and §4.2. We mentioned that the badly approximable set **Bad** and its higher dimensional generalisation $\mathbf{Bad}(i_1, \ldots, i_n)$ are small in the sense that they are of zero Lebesgue measure but are nevertheless large in the sense that they have full Hausdorff dimension. In this section we outline the basic techniques used in establishing the dimension results. For transparency and simplicity, we shall concentrate on the one-dimensional case. We begin with the classical nearly 100 years old result due to Jarník.

7.1 Bad is of full dimension

The key purpose of this section is to introduce a basic Cantor set construction and show how it can be utilised to show that **Bad** is of maximal dimension – a result first established by Jarník in [59]. Towards the end we shall mention the additional ideas required in higher dimensions.

Theorem 7.1 (Jarník, 1928)**.** *The Hausdorff dimension of* **Bad** *is one; that is*

$$\dim \mathbf{Bad} = 1\,.$$

The proof utilises the following simple **Cantor set construction.** Let $R, M \in \mathbb{N}$ and $M \leq R - 1$. Let $E_0 = [0, 1]$. Partition the interval E_0 into R equal close subinterval and remove any M of them. This gives E_1 - the union of $(R - M)$ closed intervals $\{I_{1,j}\}_{1\leq j\leq R-M}$ of length $|I_{1,j}| = R^{-1}$. Then repeat the procedure: partition each interval $I_{1,j}$ within E_1 into R equal close subinterval and remove any M intervals of the partitioning of each $I_{1,j}$. This procedure gives rise to E_2 - the union of $(R - M)^2$ closed intervals $\{I_{2,j}\}_{1\leq j\leq (R-M)^2}$ of length $|I_{2,j}| = R^{-2}$. The process goes on recurrently/inductively as follows: for $n \geq 1$, given that E_{n-1} is constructed and represents the union of $(R - M)^{n-1}$ closed intervals $\{I_{n-1,j}\}_{1\leq j\leq (R-M)^{n-1}}$ of length $|I_{n-1,j}| = R^{-(n-1)}$, to construct E_n we

(i) partition each interval $I_{n-1,j}$ within E_{n-1} into R equal closed subintervals, and

(ii) remove any M of the R intervals of the above partitioning of each $I_{n-1,j}$.

Observe that E_n will be the union of exactly $(R-M)^n$ closed intervals $\{I_{n,j}\}_{1\leq j\leq (R-M)^n}$ of length $|I_{n,j}| = R^{-n}$. The corresponding Cantor set is defined to be

$$\mathcal{K} := \bigcap_{n=0}^{\infty} E_n \,.$$

Remark 7.1. Of course the Cantor set constructed above is not unique and depends on the specific choices of M intervals being removed in each case. Indeed, there are continuum many possibilities for the resulting set $\mathcal{K}$. For example, if $R = 3$, $M = 1$ and we always remove the middle interval of the partitioning, the set $\mathcal{K}$ is the famous middle third Cantor set as described in Example 3.1 of §3.1.

Trivially, the Cantor set $\mathcal{K}$ is non-empty since it is the intersection of a nested sequence of closed intervals within $[0, 1]$. Indeed, if $0 \leq M \leq R-2$ then we have that $\mathcal{K}$ is uncountable. The following result relates the Hausdorff dimension of $\mathcal{K}$ to the parameters R and M associated with $\mathcal{K}$.

Lemma 7.1. *Let $\mathcal{K}$ be the Cantor set constructed above. Then*

$$\dim \mathcal{K} = \frac{\log(R-M)}{\log R}. \tag{7.1}$$

Proof. Let $\{I_{n,j}\}_{1\leq j\leq (R-M)^n}$ be the collection of intervals within E_n associated with the construction of $\mathcal{K}$. Recall that this is a collection of $(R-M)^n$ closed intervals, each of length R^{-n}. Naturally, $\{I_{n,j}\}_{1\leq j\leq (R-M)^n}$ is a cover of $\mathcal{K}$. Furthermore, for every $\rho > 0$ there is a sufficiently large n such that $\{I_{n,j}\}_{1\leq j\leq (R-M)^n}$ is a ρ-cover of $\mathcal{K}$ – simply make sure that $R^{-n} < \rho$. Observe that

$$\sum_j \operatorname{diam}(I_{n,j})^s = (R-M)^n R^{-ns} = 1 \qquad \text{where} \qquad s := \frac{\log(R-M)}{\log R}.$$

Hence, by definition, $\mathcal{H}^s_\rho(\mathcal{K}) \leq 1$ for all sufficiently small $\rho > 0$. Consequently, $\mathcal{H}^s(\mathcal{K}) \leq 1$ and it follows that

$$\dim \mathcal{K} \,\leq\, s\,.$$

For the lower bound, let $0 < \rho < 1$ and $\{B_i\}$ be an arbitrary ρ-cover of $\mathcal{K}$. We show that

$$\sum_i \operatorname{diam}(B_i)^s \geq \kappa,$$

where s is as above and the constant $\kappa > 0$ is independent of the cover. Without loss of generality, we will assume that each B_i is an open interval. Since $\mathcal{K}$ is the intersection of closed subsets of $[0, 1]$, it is bounded and closed and hence compact. Therefore, $\{B_i\}$ contains a finite subcover. Thus, without loss of generality, we can assume that $\{B_i\}$ is a finite ρ-cover of $\mathcal{K}$. For each B_i, let $k \in \mathbb{Z}$ be the unique integer such that

$$R^{-(k+1)} \leq \operatorname{diam}(B_i) < R^{-k}\,.$$

Then B_i intersects at most two intervals of E_k as the intervals in E_k are R^{-k} in length. If $j \geq k$, then B_i intersects at most

$$2(R-M)^{j-k} = 2(R-M)^j R^{-sk} \leq 2(R-M)^j R^s \operatorname{diam}(B_i)^s \tag{7.2}$$

intervals within E_j. These are the intervals that are contained in the (at most) two intervals of E_k that intersect B_i. Now choose j large enough so that

$$R^{-(j+1)} \leq \operatorname{diam}(B_i) \qquad \forall\ B_i\,.$$

This is possible since the cover $\{B_i\}$ is finite. Since $\{B_i\}$ is a cover of $\mathcal{K}$, it must intersect every interval of E_j. There are $(R-M)^j$ intervals within E_j. Hence, by (7.2) it follows that

$$(R-M)^j \leq \sum_i 2(R-M)^j R^s \operatorname{diam}(B_i)^s\,.$$

The upshot of this is that for any ρ-cover $\{B_i\}$ of $\mathcal{K}$, we have that

$$\sum_i \operatorname{diam}(B_i)^s \geq \tfrac{1}{2} R^{-s} = \frac{1}{2(R-M)}\,.$$

Hence, by definition, we have that $\mathcal{H}^s_\rho(\mathcal{K}) \geq \frac{1}{2(R-M)}$ for all sufficiently small $\rho > 0$. Therefore, $\mathcal{H}^s(\mathcal{K}) \geq \frac{1}{2(R-M)} > 0$ and it follows that

$$\dim \mathcal{K} \geq s = \frac{\log(R-M)}{\log R}$$

as required. □

Armed with Lemma 7.1, it is relatively straight forward to prove Jarník's full dimension result.

Proof of Theorem 7.1. Let $R \geq 4$ be an integer. For $n \in \mathbb{Z}$, $n \geq 0$ let

$$Q_n = \{p/q \in \mathbb{Q} : R^{\frac{n-3}{2}} \leq q < R^{\frac{n-2}{2}}\} \subset \mathbb{Q}\,, \tag{7.3}$$

where p/q is a reduced fraction of integers. Observe that $Q_0 = Q_1 = Q_2 = \varnothing$, that the sets Q_n are disjoint and that

$$\mathbb{Q} = \bigcup_{n=3}^{\infty} Q_n\,. \tag{7.4}$$

Furthermore, note that

$$\left|\frac{p}{q}-\frac{p'}{q'}\right| \geq \frac{1}{q'q} > R^{-n+2} \qquad \text{for different } p/q \text{ and } p'/q' \text{ in } Q_n. \tag{7.5}$$

Fix $0<\delta\leq\frac{1}{2}$. Then for $p/q \in Q_n$, define the *dangerous interval* $\Delta(p/q)$ as follows:

$$\Delta(p/q) := \left\{x\in[0,1] : \left|x-\frac{p}{q}\right| < \delta R^{-n}\right\}. \tag{7.6}$$

The goal is to construct a Cantor set $\mathcal{K}=\bigcap_{n=0}^{\infty} E_n$ such that for every $n \in \mathbb{N}$

$$E_n \cap \Delta(p/q) = \varnothing \qquad \text{for all } \; p/q \in Q_n\,. \tag{7.7}$$

To this end, let $E_0=[0,1]$ and suppose that E_{n-1} has already been constructed. Let I be any of the intervals $I_{n-1,j}$ within E_{n-1}. Then $|I|=R^{-n+1}$. By (7.5) and (7.6), there is at most one dangerous interval $\Delta(p_I/q_I)$ with $p_I/q_I \in Q_n$ that intersects I. Partition I into R closed subintervals of length $R^{-n}=R^{-1}|I|$. Note that since $\delta\leq\frac{1}{2}$, the dangerous interval $\Delta(p_I/q_I)$, if it exists, can intersect at most 2 intervals of the partitioning of I. Hence, by removing $M=2$ intervals of the partitioning of each I within E_{n-1} we construct E_n while ensuring that (7.7) is satisfied. By Lemma 7.1, it follows that for any $R\geq 4$

$$\dim \mathcal{K} \geq \frac{\log(R-2)}{\log R}\,.$$

Now take any $x\in\mathcal{K}$ and any $p/q\in\mathbb{Q}$. Then $p/q \in Q_n$ for some $n\in\mathbb{N}$ and since $\mathcal{K}\subset E_n$ we have that $x \in E_n$. Then, by (7.7), we have that $x \notin \Delta(p/q)$, which implies that

$$\left|x-\frac{p}{q}\right| \geq \delta R^{-n} \geq \delta R^{-3} q^{-2}\,. \tag{7.8}$$

Since $p/q\in\mathbb{Q}$ is arbitrary and R and δ are fixed, we have that $x\in\mathbf{Bad}$. That is, $\mathcal{K}\subset\mathbf{Bad}$ and thus it follows that

$$\dim\mathbf{Bad} \,\geq\, \dim\mathcal{K} \,\geq\, \frac{\log(R-2)}{\log R}\,.$$

This is true for any $R\geq 4$ and so on letting $R\to\infty$, it follows that $\dim\mathbf{Bad}\geq 1$. The complementary upper bound statement $\dim\mathbf{Bad}\leq 1$ is trivial since $\mathbf{Bad}\subset\mathbb{R}$. $\square$

Remark 7.2. The crucial property underpinning the proof of Theorem 7.1 is the separation property (7.5) of rationals. Indeed, without appealing to Lemma 7.1, the above proof based on (7.5) alone shows that **Bad** is uncountable. The construction of the Cantor set $\mathcal{K}$ as well as the proof of Theorem 7.1 can be generalised to higher dimensions in order to show that

$$\dim\mathbf{Bad}(i_1,\dots,i_n) = n\,.$$

Regarding the higher dimensional generalisation of the proof of Theorem 7.1, the appropriate analogue of (7.5) is the following elegant Simplex Lemma – see for example [69, Lemma 4].

Lemma 7.2 (Simplex Lemma)**.** *Let $m \geq 1$ be an integer and $Q > 1$ be a real number. Let $E \subseteq \mathbb{R}^m$ be a convex set of m-dimensional Lebesgue measure*

$$|E| \ \leq \ (m!)^{-1} Q^{-(m+1)} \ .$$

Suppose that E contains $m+1$ rational points $(p_i^{(1)}/q_i, \ldots, p_i^{(m)}/q_i)$ with $1 \leq q_i < Q$, where $0 \leq i \leq m$. Then these rational points lie in some hyperplane of $\mathbb{R}^m$.

7.2 Schmidt's games

In his pioneering work [85], Wolfgang M. Schmidt introduced the notion of (α, β)-games which now bear his name. These games are an extremely powerful tool for investigating badly approximable sets. The simplified account which we are about to present is sufficient to bring out the main features of the games.

Suppose that $0 < \alpha < 1$ and $0 < \beta < 1$. Consider the following game involving the two arch rivals **A**yesha and **B**hupen – often simply referred to as players **A** and **B**. First, **B** chooses a closed ball $\mathbf{B}_0 \subset \mathbb{R}^m$. Next, **A** chooses a closed ball $\mathbf{A}_0$ contained in $\mathbf{B}_0$ of diameter $\alpha\, \rho(\mathbf{B}_0)$ where $\rho(\ .\ )$ denotes the diameter of the ball under consideration. Then, **B** chooses at will a closed ball $\mathbf{B}_1$ contained in $\mathbf{A}_0$ of diameter $\beta\, \rho(\mathbf{A}_0)$. Alternating in this manner between the two players, generates a nested sequence of closed balls in $\mathbb{R}^m$:

$$\mathbf{B}_0 \supset \mathbf{A}_0 \supset \mathbf{B}_1 \supset \mathbf{A}_1 \supset \ldots \supset \mathbf{B}_n \supset \mathbf{A}_n \supset \ldots \tag{7.9}$$

with diameters

$$\rho(\mathbf{B}_n) \ = \ (\alpha\, \beta)^n\, \rho(\mathbf{B}_0) \quad \text{and} \quad \rho(\mathbf{A}_n) \ = \ \alpha\, \rho(\mathbf{B}_n)\,.$$

A subset X of $\mathbb{R}^m$ is said to be (α, β)-*winning* if **A** can play in such a way that the unique point of the intersection

$$\bigcap_{n=0}^{\infty} \mathbf{B}_n \ = \ \bigcap_{n=0}^{\infty} \mathbf{A}_n$$

lies in X, regardless of how **B** plays. The set X is called α-*winning* if it is (α, β)-winning for all $\beta \in (0, 1)$. Finally, X is simply called *winning* if it is α-winning for some α. Informally, player **B** tries to stay away from the 'target' set X whilst player **A** tries to land on X. As shown by Schmidt in [85], the following are the key consequences of winning.

- *If $X \subset \mathbb{R}^m$ is a winning set, then* $\dim X = m$.
- *The intersection of countably many α-winning sets is α-winning.*

Schmidt [85] proved the following fundamental result for the symmetric case of the higher dimensional analogue of **Bad** which, given the above properties, has implications well beyond simply full dimension.

Theorem 7.2 (Schmidt, 1966)**.** *For any $m \in \mathbb{N}$, the set* $\mathbf{Bad}(\frac{1}{m}, \ldots, \frac{1}{m})$ *is winning.*

Proof. To illustrate the main ideas involved in proving the theorem we shall restrict our attention to when $m = 1$. In this case, we are able to establish the desired winning statement by naturally modifying the proof of Theorem 7.1. Without loss of generality, we can restrict $\mathbf{Bad} := \mathbf{Bad}(1)$ to the unit interval $[0, 1]$. Let $0 < \alpha < \frac{1}{2}$ and $0 < \beta < 1$. Let $R = (\alpha\beta)^{-1}$ and define Q_n by (7.3). Again $Q_0 = Q_1 = Q_2 = \varnothing$; the sets Q_n are disjoint; (7.4) and (7.5) are both true. Furthermore, for $p/q \in Q_n$ the corresponding dangerous interval $\Delta(p/q)$ is defined by (7.6), where $0 < \delta < 1$ is to be specified below and will be dependent on α and the first move made by **B**hupen .

Our goal is to show that **A**yesha has a strategy to ensure that sequence (7.9) satisfies

$$\mathbf{A}_n \cap \Delta(p/q) = \varnothing \qquad \text{for all } \; p/q \in Q_n\,. \tag{7.10}$$

Then the single point x corresponding to the intersection over all the closed and nested intervals $\mathbf{A}_n$ would satisfy (7.8) for all $p/q \in \mathbb{Q}$ meaning that x is badly approximable. By definition, this would implying that $\mathbf{Bad}$ is α-winning as desired.

Let $\mathbf{B}_0 \subset [0, 1]$ be any closed interval. Now we set

$$\delta := \rho(\mathbf{B}_0)(\tfrac{1}{2} - \alpha).$$

Suppose that

$$\mathbf{B}_0 \supset \mathbf{A}_0 \supset \mathbf{B}_1 \supset \mathbf{A}_1 \supset \ldots \supset \mathbf{B}_{n-1} \supset \mathbf{A}_{n-1}$$

are already chosen and satisfy the required properties; namely (7.10). Suppose that $\mathbf{B}_n \subset \mathbf{A}_{n-1}$ is any closed interval of length

$$\rho(\mathbf{B}_n) = \beta\rho(\mathbf{A}_{n-1}) = (\alpha\,\beta)^n\,\rho(\mathbf{B}_0) = R^{-n}\rho(\mathbf{B}_0).$$

Next, $\mathbf{A}$ has to choose a closed interval $\mathbf{A}_n$ contained in $\mathbf{B}_n$ of diameter

$$\rho(\mathbf{A}_n) = \alpha\,\rho(\mathbf{B}_n) = \alpha R^{-n}\rho(\mathbf{B}_0)$$

and satisfying (7.10). If (7.10) is satisfied with $\mathbf{A}_n$ replaced by $\mathbf{B}_n$, then choosing $\mathbf{A}_n$ obviously represents no problem. Otherwise, using (7.5) one readily verifies that there is exactly one point $p_n/q_n \in Q_n$ such that $\Delta(p_n/q_n)$ intersects $\mathbf{B}_n$. In this case $\mathbf{B}_n \setminus \Delta(p_n/q_n)$ is either the union of two closed intervals, the larger one being of length

$$\geq \tfrac{1}{2}\Big(\rho(\mathbf{B}_n) - \rho(\Delta(p_n/q_n))\Big) = \tfrac{1}{2}R^{-n}\Big(\rho(\mathbf{B}_0) - 2\delta\Big) = \alpha R^{-n}\rho(\mathbf{B}_0) = \alpha\rho(\mathbf{B}_n)$$

or a single closed interval of even greater length. Hence, it is possible to choose a closed interval $\mathbf{A}_n \subset \mathbf{B}_n \setminus \Delta(p_n/q_n)$ of length $\rho(\mathbf{A}_n) = \alpha\rho(\mathbf{B}_n)$. By construction, (7.10) is satisfied, thus proving the existence of a winning strategy for $\mathbf{A}$. □

Remark 7.3. For various reasons, over the last decade or so there has been an explosion of interest in Schmidt's games. This has given rise to several ingenious generalisations of the original game leading to stronger notions of winning, such as modified winning, absolute winning, hyperplane winning and potential winning. For details see [51, 68] and references within.

The framework of Schmidt games and thus the notion of winning is defined in terms of balls. Thus, it is naturally applicable when considering the symmetric case ($i_1 = \ldots = i_n = 1/n$) of the badly approximable sets $\mathbf{Bad}(i_1 \ldots, i_n)$. Recall, that in the symmetric case, points in $\mathbf{Bad}(\frac{1}{n}, \ldots, \frac{1}{n})$ avoid squares (which are essentially balls) centred around rational points were as in the general case the points avoiding rectangles (far from being balls). We now turn our attention to the general case. Naturally, it would be desirable to be able to show that the general set $\mathbf{Bad}(i_1 \ldots, i_n)$ is winning.

7.3 Properties of general $\mathbf{Bad}(i_1 \ldots, i_n)$ sets beyond full dimension

Despite the fact that the sets $\mathbf{Bad}(i_1, \ldots, i_n)$ have long been know to be uncountable and indeed of full dimension, see [42, 68, 69, 77], the following conjecture of Schmidt dating back to 1982 remained unresolved until reasonably recently.

Schmidt's Conjecture .

$$\mathbf{Bad}(\tfrac{1}{3}, \tfrac{2}{3}) \cap \mathbf{Bad}(\tfrac{2}{3}, \tfrac{1}{3}) \neq \varnothing \,.$$

As is already highlighted in Remark 4.2, if false then it would imply that Littlewood's Conjecture is true.

Schmidt's Conjecture was proved in [7] by establishing the following stronger statement regarding the intersection of $\mathbf{Bad}(i, j)$ sets with vertical lines $\mathrm{L}_\alpha := \{(\alpha, y) : y \in \mathbb{R}\} \subset \mathbb{R}^2$. To some extent it represents the badly approximable analogue of the 'fiber' results that appeared in §4.5.

Theorem 7.3. *Let (i_k, j_k) be a countable sequence of non-negative reals such that $i_k + j_k = 1$ and let $i := \sup\{i_k : k \in \mathbb{N}\}$. Suppose that*

$$\liminf_{k\to\infty} \min\{i_k, j_k\} > 0 \,. \tag{7.11}$$

Then, for any $\alpha \in \mathbb{R}$ such that $\liminf\limits_{q\to\infty} q^{1/i} \|q\alpha\| > 0$, we have that

$$\dim \textstyle\bigcap_k \mathbf{Bad}(i_k, j_k) \cap \mathrm{L}_\alpha = 1. \tag{7.12}$$

Remark 7.4. The Diophantine condition imposed on α associated with the vertical line L_α is easily seen to be necessary – see [7, §1.3]. Note that the condition is automatically satisfied if $\alpha \in \mathbf{Bad}$. On the other hand, condition (7.11) is present for technical reason and can be removed – see Theorem 7.5 and discussion below. At the point, simply observe that it is automatically satisfied for any finite collection of pairs (i_k, j_k) and thus Theorem 7.3 implies Schmidt's Conjecture. Indeed, together with a standard 'slicing' result from fractal geometry one obtains the following full dimension statement – see [7, §1.2] for details.

Corollary 7.1. *Let (i_k, j_k) be a countable sequence of non-negative reals such that $i_k + j_k = 1$ and satisfying condition* (7.11)*. Then,*

$$\dim \textstyle\bigcap_k \mathbf{Bad}(i_k, j_k) = 2. \tag{7.13}$$

At the heart of establishing Theorem 7.3 is the 'raw' construction of the generalised Cantor sets framework formulated in [8]. For the purposes of these notes, we opt to follow the framework of *Cantor rich sets* introduced in [12] which is a variation of the aforementioned generalised Cantor sets.

Let $R \geq 3$ be an integer. Given a collection $\mathcal{I}$ of compact intervals in $\mathbb{R}$, let $\frac{1}{R}\mathcal{I}$ denote the collection of intervals obtained by splitting each interval in $\mathcal{I}$ into R equal closed subintervals with disjoint interiors. Given a compact interval $I_0 \subset \mathbb{R}$, the sequence $(\mathcal{I}_q)_{q\geq 0}$ such that

$$\mathcal{I}_0 = \{I_0\} \qquad \text{and} \qquad \mathcal{I}_q \subset \tfrac{1}{R}\mathcal{I}_{q-1} \quad \text{for } q \geq 1$$

is called *an R-sequence in I_0*. It defines the corresponding *generalised Cantor set*:

$$\mathcal{K}((\mathcal{I}_q)_{q\geq 0}) \ := \ \bigcap_{q\geq 0} \ \bigcup_{I_q \in \mathcal{I}_q} I_q. \tag{7.14}$$

Given $q \in \mathbb{N}$ and any interval J, let

$$\widehat{\mathcal{I}}_q \ := \ \left(\tfrac{1}{R}\mathcal{I}_{q-1}\right) \setminus \mathcal{I}_q \qquad \text{and} \qquad \widehat{\mathcal{I}}_q \sqcap J \ := \ \{I_q \in \widehat{\mathcal{I}}_q : I_q \subset J\}\,.$$

Furthermore, define

$$d_q(\mathcal{I}_q) \ := \ \min_{\{\widehat{\mathcal{I}}_{q,p}\}} \ \sum_{p=0}^{q-1} \left(\frac{4}{R}\right)^{q-p} \max_{I_p \in \mathcal{I}_p} \#\big(\widehat{\mathcal{I}}_{q,p} \sqcap I_p\big)\ , \tag{7.15}$$

where the minimum is taken over all partitions $\{\widehat{\mathcal{I}}_{q,p}\}_{p=0}^{q-1}$ of $\widehat{\mathcal{I}}_q$; that is $\widehat{\mathcal{I}}_q = \bigcup_{p=0}^{q-1} \widehat{\mathcal{I}}_{q,p}$.

The following dimension statement was established in [8, Theorem 4], see also [12, Theorem 5].

Lemma 7.3. *Let $\mathcal{K}((\mathcal{I}_q)_{q\geq 0})$ be the Cantor set given by* (7.14)*. Suppose that*

$$d_q(\mathcal{I}_q) \leq 1 \tag{7.16}$$

for all $q \in \mathbb{N}$. Then

$$\dim \mathcal{K}((\mathcal{I}_q)_{q\geq 0}) \geq 1 - \frac{\log 2}{\log R}\,.$$

Although the lemma can be viewed as a generalisation of Lemma 7.1, we stress that its proof is substantially more involved and requires new ideas. At the heart of the proof is the 'extraction' of a 'local' Cantor type subset $\mathcal{K}$ of $\mathcal{K}((\mathcal{I}_q)_{q\geq 0})$. By a local Cantor set we mean a set arising from a construction as described in §7.1. The parameter M associated with the extracted local Cantor set $\mathcal{K}$ is essentially $\frac{1}{2}R$.

It is self evident from Lemma 7.3, that if a given set $X \subset \mathbb{R}$ contains a generalised Cantor set given by (7.14) with arbitrarily large R, then $\dim X = 1$. The following definition of Cantor rich [12], imposes a stricter requirement than (7.16) in order to ensure that the countable intersection of generalised Cantor sets is of full dimension. To some extent, building upon the raw construction of [7, §7.1], the full dimension aspect for countable intersections had previously been investigated in [8, §7].

Definition 7.1. Let $M > 1$, $X \subset \mathbb{R}$ and I_0 be a compact interval. The set X is said to be *M-Cantor rich in I_0* if for any $\varepsilon > 0$ and any integer $R \geq M$ there exists an R-sequence $(\mathcal{I}_q)_{q\geq 0}$ in I_0 such that $\mathcal{K}((\mathcal{I}_q)_{q\geq 0}) \subset X$ and

$$\sup_{q\in\mathbb{N}} d_q(\mathcal{I}_q) \leq \varepsilon\,.$$

The set X is said to be *Cantor rich in I_0* if it is *M-Cantor rich in I_0* for some M, and it is said to be *Cantor rich* if it is *Cantor rich in I_0* for some compact interval I_0.

The following summarises the key properties of Cantor rich sets.

(i) Any Cantor rich set X in $\mathbb{R}$ satisfies $\dim X = 1$.

(ii) For any given compact interval I_0 and any given fixed $M \in \mathbb{N}$, any countable intersection of M-Cantor rich sets in I_0 is M-Cantor rich in I_0.

The framework of Cantor-rich sets was utilised in the same paper [12] to establish the following result concerning badly approximable points on manifolds.

Theorem 7.4. *For any non-degenerate analytic sub-manifold $\mathcal{M} \subset \mathbb{R}^n$ and any sequence $(i_{1,k}, \ldots, i_{n,k})$ of non-negative reals such that $i_{1,k} + \cdots + i_{n,k} = 1$ and*

$$\inf\{i_{j,k} > 0 : 1 \leq j \leq n,\ k \in \mathbb{N}\} > 0\,, \tag{7.17}$$

one has that

$$\dim \textstyle\bigcap_k \mathbf{Bad}(i_{1,k}, \ldots, i_{n,k}) \cap \mathcal{M} = \dim \mathcal{M}\,. \tag{7.18}$$

The condition of analyticity from Theorem 7.4 can be omitted in the case the sub-manifold $\mathcal{M} \subset \mathbb{R}^n$ is a curve. Indeed, establishing the theorem for curves is very much the crux since any manifold can be 'fibred' into an appropriate collection of curves – see [12, §2.1] for details. In the case $n = 2$, so that $\mathcal{M}$ is a non-degenerate planar curve, the theorem was previously established in [9] and provides a solution to an explicit problem of Davenport dating back to the swinging sixties concerning the existence of badly approximable pairs on the parabola. Furthermore, in [9] partial results for lines (degenerate curves) with slopes satisfying certain Diophantine constraints are also obtained. Although not optimal, they naturally extend Theorem 7.3 beyond vertical lines. As already mentioned, Theorem 7.4 as stated for general n was established in [12] and it settles the natural generalisations of Schmidt's Conjecture and Davenport's problem in arbitrary dimensions.

Remark 7.5. Building upon the one-dimensional, generalised Cantor sets framework formulated in [8], an abstract 'metric space' framework of higher dimensional generalised Cantor sets, branded as 'Cantor winning sets', has recently been introduced in [6]. Projecting this framework onto the specific one-dimensional construction of Cantor rich sets given above, the definition of Cantor-winning sets reads as follows. Let $\varepsilon_0 > 0$, $X \subset \mathbb{R}$ and I_0 be a compact interval. Then the set X is *ε_0-Cantor-winning in I_0* if for any positive $\varepsilon < \varepsilon_0$ there exists a positive integer R_ε such that for any integer $R \geq R_\varepsilon$ there exists an R-sequence $(\mathcal{I}_q)_{q\geq 0}$ in I_0 such that $\mathcal{K}((\mathcal{I}_q)_{q\geq 0}) \subset X$ and

$$\max_{I_p \in \mathcal{I}_p} \#\big(\widehat{\mathcal{I}}_{q,p} \sqcap I_p\big) \leq R^{(q-p)(1-\varepsilon)} .$$

The latter key condition implies that $d_q(\mathcal{I}_q)$ is no more than $8R^{-\varepsilon}$ provided that $8R^{-\varepsilon} < 1$. Most recently, David Simmons has shown that the notion of Cantor winning as defined in [6] is equivalent to the notion of potential winning as defined in [51].

The use of Cantor rich sets in establishing statements such as Theorems 7.3 & 7.4, comes at a cost of having to impose, seemingly for technical reasons, conditions such as (7.11) and (7.17). Although delivering some additional benefits, unfortunately the framework of Cantor winning sets described above does not seem to resolve this issue. However, if for example, we could show that $\mathbf{Bad}(i_1, \ldots, i_n)$ is (Schmidt) winning, then we would be able to intersect countably many such sets without imposing any technical conditions. When $n = 2$, this has been successfully accomplished by Jinpeng An in his elegant paper [2].

Theorem 7.5 (J. An)**.** *For any pair of non-negative reals (i, j) such that $i + j = 1$, the two-dimensional set $\mathbf{Bad}(i, j)$ is winning.*

A simple consequence of this is that we can remove condition (7.11) from the statement of Corollary 7.1. Prior to [2], it is important to note that An in [1] had shown that $\mathbf{Bad}(i, j) \cap \mathrm{L}_\alpha$ is winning, where L_α is a vertical line as in Theorem 7.3. Of course, this implies that Theorem 7.3 is true without imposing condition (7.11). On combining the ideas and techniques introduced in the papers [1, 9, 12], it is shown in [3] that $\mathbf{Bad}(i, j) \cap \mathcal{C}$ is winning, where $\mathcal{C}$ is a non-degenerate planar curve. This implies that we can remove condition (7.17) from the $n = 2$ statement of Theorem 7.4. In higher dimensions ($n > 2$), removing condition (7.17) remains very much a key open problem. The recent work of Guan and Yu [44] makes a contribution toward this problem. Building upon the work of An [2], they show that the set $\mathbf{Bad}(i_1, \ldots, i_n)$ is winning whenever $i_1 = \cdots = i_{n-1} \geq i_n$.

So far we have discussed the homogeneous theory of badly approximable sets. We now turn our attention to the inhomogeneous theory.

7.4 Inhomogeneous badly approximable points

Given $\theta \in \mathbb{R}$ the natural inhomogeneous generalisation of the one-dimensional set $\mathbf{Bad}$ is the set

$$\mathbf{Bad}(\theta) := \{x \in \mathbb{R} : \exists\ c(x) > 0 \ \text{ so that } \|qx - \theta\| \ > \ c(x)\ q^{-1} \quad \forall\ q \in \mathbb{N}\} .$$

Within these notes we shall prove the following inhomogeneous strengthening of Theorem 7.1.

Theorem 7.6. *For any $\theta \in \mathbb{R}$, we have that*

$$\dim \mathbf{Bad}(\theta) = 1 .$$

The basic philosophy behind the proof is simple and exploits the already discussed homogeneous 'intervals construction'; namely

$$\text{(homogeneous construction)} \quad + \quad (\boldsymbol{\theta} - \boldsymbol{\theta} = \mathbf{0}) \quad \Longrightarrow \quad \text{(inhomogeneous statement)}.$$

Remark 7.6. Recall that we have already made use of this type of philosophy in establishing the inhomogeneous extremality conjecture stated in §6.3.1, where the proof very much relies on the fact that we already know that any non-degenerate manifold is (homogeneously) extremal.

Proof of Theorem 7.6. Let $R \geq 4$ be an integer and $\delta = \frac{1}{2}$. For $n \in \mathbb{Z}$, $n \geq 0$, define the sets Q_n by (7.3) and additionally define the following sets of 'shifted' rational points

$$Q_n(\theta) = \{(p+\theta)/q \in \mathbb{R} : p, q \in \mathbb{Z},\ R^{\frac{n-5}{2}} \leq q < R^{\frac{n-4}{2}}\}. \tag{7.19}$$

Clearly, $Q_0(\theta) = \cdots = Q_4(\theta) = \varnothing$ and the union $Q(\theta) := \bigcup_{n=5}^{\infty} Q_n(\theta)$ contains all the possible points $(p+\theta)/q$ with $p, q \in \mathbb{Z}$, $q > 0$.

Next, for $p/q \in Q_n$ define the dangerous interval $\Delta(p/q)$ by (7.6) and additionally define the inhomogeneous family of dangerous intervals given by

$$\Delta((p+\theta)/q) := \left\{x \in [0,1] : \left|x - \frac{p+\theta}{q}\right| < \delta R^{-n}\right\}, \tag{7.20}$$

where $(p+\theta)/q \in Q(\theta)$. With reference to the Cantor construction of §7.1, our goal is to construct a Cantor set $\mathcal{K} = \bigcap_{n=0}^{\infty} E_n$ such that for every $n \in \mathbb{N}$

$$E_n \cap \Delta(p/q) = \varnothing \qquad \text{for all } \ p/q \in Q_n \tag{7.21}$$

and simultaneously

$$E_n \cap \Delta((p+\theta)/q) = \varnothing \qquad \text{for all } \ (p+\theta)/q \in Q_n(\theta). \tag{7.22}$$

To this end, let $E_0 = [0,1]$ and suppose that E_{n-1} has been constructed as required. Let I be any interval within E_{n-1}. Then $|I| = R^{-n+1}$. When constructing E_n, I is partitioned into R subintervals. We need to decide how many of these subintervals have to be removed in order to satisfy (7.21) and (7.22). As was argued in the proof of Theorem 7.1, removing 2 intervals of the partitioning of I ensures that (7.21) is satisfied. We claim that the same applies to (7.22), that is removing 2 intervals of the partitioning of I ensures (7.22). Indeed, since the length of $\Delta((p+\theta)/q)$ is no more that R^{-n}, to verify this claim it suffices to show that there is only one point $(p+\theta)/q \in Q_n(\theta)$ such that

$$\Delta((p+\theta)/q) \cap I \neq \varnothing.$$

This condition implies that

$$|qx - p - \theta| < R^{\frac{n-4}{2}}(\delta R^{-n} + R^{-n+1}) \qquad \text{for any } x \in I. \tag{7.23}$$

For a contradiction, suppose there are two distinct points $(p_1+\theta)/q_1$ and $(p_2+\theta)/q_2$ in $Q_n(\theta)$ satisfying (7.23). Then, by (7.23) and the triangle inequality, we get that

$$|(q_1 - q_2)x - (p_1 - p_2)| < 2R^{\frac{n-4}{2}}(\delta R^{-n} + R^{-n+1}) \qquad \text{for any } x \in I. \tag{7.24}$$

Clearly $q_1 \neq q_2$ as otherwise we would have that $|p_1 - p_2| < 2R^{\frac{n-4}{2}}(\delta R^{-n} + R^{-n+1}) < 1$, implying that $p_1 = p_2$ and contradicting to the fact that $(p_1+\theta)/q_1$ and $(p_2+\theta)/q_2$ are distinct. In the above we have used that $n \geq 5$. Also without loss of generality we assume that $q_1 > q_2$. Then define $d = \gcd(q_1 - q_2, p_1 - p_2)$, $q = (q_1 - q_2)/d$, $p = (p_1 - p_2)/d$ and let m be the unique integer such that

$$p/q \in Q_m \,.$$

Thus, $R^{\frac{m-3}{2}} \leq q < R^{\frac{m-2}{2}}$. Since $q < q_1 < R^{\frac{n-4}{2}}$ we have that $m \leq n-2$. Then, by (7.24),

$$\left| x - \frac{p}{q} \right| < R^{-\frac{m-3}{2}} 2R^{\frac{n-4}{2}}(\delta R^{-n} + R^{-n+1}) \leq \delta R^{-m} \tag{7.25}$$

for any $x \in I$ provided that $R \geq 36$ (recall that $\delta = \frac{1}{2}$). It means that $\Delta(p/q) \cap I \neq \varnothing$. But this is impossible since (7.21) is valid with n replaced by m and $I \subset E_{n-1} \subset E_m$. This proves our above claim. The upshot is that by removing $M = 4$ intervals of the partitioning of each I within E_{n-1} we construct E_n while ensuring that the desired conditions (7.21) and (7.22) are satisfied. The finale of the proof makes use of Lemma 7.1 and is almost identical to that of the proof of Theorem 7.1. We leave the details to the reader. □

Remark 7.7. Note that in the above proof of Theorem 7.6, we actually show that

$$\dim \mathbf{Bad} \cap \mathbf{Bad}(\theta) = 1 \,.$$

It seems that proving this stronger statement is simpler than any potential 'direct' proof of the implied fact that $\dim \mathbf{Bad}(\theta) = 1$.

Remark 7.8. In the same way that the proof of Theorem 7.1 can be modified to show that **Bad** is winning (see the proof of Theorem 7.2 for the details), the proof of Theorem 7.6 can be adapted to show that $\mathbf{Bad}(\theta)$ is winning.

In higher dimensions, the natural generalisation of the one-dimensional set $\mathbf{Bad}(\theta)$ is the set $\mathbf{Bad}(i_1, \ldots, i_n; \boldsymbol{\theta})$ defined in the following manner. For any $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$ and n-tuple of real numbers $i_1, ..., i_n \geq 0$ such that $i_1 + \cdots + i_n = 1$, we let $\mathbf{Bad}(i_1, \ldots, i_n; \boldsymbol{\theta})$ to be the set of points $(x_1, ..., x_n) \in \mathbb{R}^n$ for which there exists a positive constant $c(x_1, ..., x_n)$ such that

$$\max\{\, ||qx_1 - \theta_1||^{1/i_1} \,, ..., \, ||qx_n - \theta_n||^{1/i_n} \,\} \; > \; c(x_1, ..., x_n)\, q^{-1} \quad \forall \quad q \in \mathbb{N}.$$

The ideas used in the proof of Theorem 7.6 can be naturally generalised to show that

$$\dim \mathbf{Bad}(i_1 \ldots, i_n; \boldsymbol{\theta}) = n \,.$$

In the case $n = 2$, the details of the proof are explicitly given in [28, §3]. Indeed, as mentioned in [28, Remark 3.4], in the symmetric case $i_1 = \ldots = i_n = 1/n$, we actually have that $\mathbf{Bad}(\frac{1}{n}, \ldots, \frac{1}{n}; \boldsymbol{\theta})$ is winning; i.e. the inhomogeneous strengthening of Theorem 7.2.

Remark 7.9. The basic philosophy exploited in proving Theorem 7.6 has been successfully incorporated within the context of Schmidt games to establish the inhomogeneous generalisation of the homogeneous winning statements discussed at the end of §7.3. In particular, let $\boldsymbol{\theta} \in \mathbb{R}^2$ and (i, j) be a pair of non-negative real numbers such that $i+j = 1$. Then, it is shown in [3] that (i) the set $\mathbf{Bad}(i, j; \boldsymbol{\theta})$ is winning and (ii) for any non-degenerate planar curve $\mathcal{C}$, the set $\mathbf{Bad}(i, j; \boldsymbol{\theta}) \cap \mathcal{C}$ is winning. Also, in [3] the following almost optimal winning result for the intersection of $\mathbf{Bad}(i, j)$ sets with arbitrary lines (degenerate curves) is obtained. It substantially extends and generalises the previous 'line' result obtained in [9].

Theorem 7.7. *Let (i, j) be a pair of non-negative real numbers such that $i + j = 1$ and given $a, b \in \mathbb{R}$ with $a \neq 0$, let $\mathrm{L}_{a,b}$ denote the line defined by the equation $y = ax + b$. Suppose there exists $\epsilon > 0$ such that*

$$\liminf_{q\to\infty} q^{\frac{1}{\sigma}-\epsilon} \max\{\|qa\|, \|qb\|\} > 0 \qquad \text{where} \ \ \sigma := \min\{i, j\}\,. \tag{7.26}$$

Then, for any $\boldsymbol{\theta} \in \mathbb{R}^2$ we have that $\mathbf{Bad}_{\boldsymbol{\theta}}(i, j) \cap \mathrm{L}_{a,b}$ is winning. Moreover, if $a \in \mathbb{Q}$ the statement is true with $\epsilon = 0$ in (7.26).

The condition (7.26) is optimal up to the ϵ – see [3, Remark 4]. It is indeed, both necessary and sufficient in the case $a \in \mathbb{Q}$. Note that the argument presented in [3, Remark 4] showing the necessity of (7.26) with $\epsilon = 0$ only makes use of the assumption that $\mathbf{Bad}(i, j) \cap \mathrm{L}_{a,b} \neq \varnothing$. It is plausible to suggest that this latter assumption is a necessary and sufficient condition for the conclusion of Theorem 7.7 to hold.

Conjecture 7.1. *Let (i, j) be a pair of non-negative real numbers such that $i + j = 1$ and given $a, b \in \mathbb{R}$ with $a \neq 0$, let $\mathrm{L}_{a,b}$ denote the line defined by the equation $y = ax + b$. Then*

$$\mathbf{Bad}(i, j) \cap \mathrm{L}_{a,b} \neq \varnothing$$

if and only if

$$\forall\ \boldsymbol{\theta} \in \mathbb{R}^2 \qquad \mathbf{Bad}_{\boldsymbol{\theta}}(i, j) \cap \mathrm{L}_{a,b} \quad \textit{is winning.}$$

Observe that the conjecture is true in the case $a \in \mathbb{Q}$ and when the line $\mathrm{L}_{a,b}$ is horizontal or vertical in the homogenous case.

Acknowledgements. SV would like to thank the organisers of the 2014 Durham Easter School "Dynamics and Analytic Number Theory" for giving him the opportunity to give a mini-course – it was a stimulating and enjoyable experience. Subsequently, the subject matter of that mini-course formed the foundations for a MAGIC graduate lecture course on metric number theory given jointly by VB and SV at the University of York in Spring 2015. We would like to thank the participants of these courses for providing valuable feedback on both the lectures and the accompanying notes. In particular, we thank Demi Allen and Henna Koivusalo for their detailed comments (well beyond the call of duty) on earlier drafts of this end product. For certain their input has improved the clarity and the accuracy of the exposition. Of course, any remaining typos and mathematical errors are absolutely their fault!

References

[1] J. An, *Badziahin-Pollington-Velani's theorem and Schmidt's game*. Bull. London Math. Soc. **45** (2013), no. 4, 721–733.

[2] J. An, *Two dimensional badly approximable vectors and Schmidt's game*. Duke Math. J. (2015), doi:10.1215/00127094-3165862

[3] J. An, V. Beresnevich, and S. L. Velani, *Badly approximable points on planar curves and winning*. Pre-print: arXiv:1409.0064.

[4] D. Badziahin *On multiplicatively badly approximable numbers*. Mathematika **59** (2013), no. 1, 31–55.

[5] D. Badziahin, J. Levesley, *A note on simultaneous and multiplicative Diophantine approximation on planar curves*. Glasgow Mathematical Journal **49** (2007), no. 2, 367–375.

[6] D. Badziahin, S. Harrap, *Cantor-winning sets and their applications*. Preprint: arXiv:1503.04738.

[7] D. Badziahin, A. Pollington, S. L. Velani, *On a problem in simultaneous diophantine approximation: Schmidt's conjecture*. Ann. of Math. (2) **174** (2011), no. 3, 1837–1883.

[8] D. Badziahin, S. L. Velani, *Multiplicatively badly approximable numbers and generalised Cantor sets*. Adv. Math. **225** (2011), 2766–2796.

[9] D. Badziahin, S. L. Velani, *Badly approximable points on planar curves and a problem of Davenport*. Mathematische Annalen **359** (2014), no. 3, 969–1023.

[10] A. Baker, W. M. Schmidt, *Diophantine approximation and Hausdorff dimension*, Proc. Lond. Math. Soc. **21** (1970), 1–11.

[11] V. Beresnevich, *Rational points near manifolds and metric Diophantine approximation*. Ann. of Math. (2) **175** (2012), no. 1, 187–235.

[12] V. Beresnevich, *Badly approximable points on manifolds*. Inventiones Mathematicae **202** (2015), no. 3, 1199–1240.

[13] V. Beresnevich, V. Bernik, M. Dodson, S. L. Velani, *Classical metric Diophantine approximation revisited*, Roth Festschrift - essays in honour of Klaus Roth on the occasion of his 80th birthday. Editors: W. Chen, T. Gowers, H. Halberstam, W.M. Schmidt and R.C. Vaughan. Cambridge University Press. (2009), pp. 38–61.

[14] V. Beresnevich, D. Dickinson, S. L. Velani, *Measure Theoretic Laws for limsup Sets*, Mem. Amer. Math. Soc. **179** (2006), no. 846, 1–91.

[15] V. Beresnevich, D. Dickinson, S. L. Velani, *Diophantine approximation on planar curves and the distribution of rational points*, Ann. of Math. (2) **166** (2007), pp. 367–426. With an Appendix II by R. C. Vaughan.

[16] V. Beresnevich, A. Haynes, S. L. Velani, *Multiplicative zero-one laws and metric number theory*. Acta Arith. **160** (2013), no. 2, 101-114.

[17] V. Beresnevich, A. Haynes, S. L. Velani, *Sums of reciprocals of fractional parts and multiplicative Diophantine approximation* , Pre-print: arXiv:1511.06862 .

[18] V. Beresnevich, R. Vaughan, S. L. Velani, *Inhomogeneous Diophantine approximation on planar curves.* Math. Ann. **349** (2011), 929–942.

[19] V. Beresnevich, R. Vaughan, S. L. Velani, E. Zorin, *Diophantine approximation on manifolds and the distribution of rational points: contributions to the convergence theory.* Pre-print: arXiv:1506.09049

[20] V. Beresnevich, R. Vaughan, S. L. Velani, E. Zorin, *Diophantine approximation on manifolds and the distribution of rational points: contributions to the divergence theory.* In preperation.

[21] V. Beresnevich, S. L. Velani, *A Mass Transference Principle and the Duffin-Schaeffer conjecture for Hausdorff measures*, Ann. Math. **164** (2006), 971–992.

[22] V. Beresnevich, S. L. Velani, *Schmidt's theorem, Hausdorff Measures and Slicing*, IMRN 2006, Article ID 48794, 24 pages.

[23] V. Beresnevich, S. L. Velani, *A note on simultaneous Diophantine approximation on planar curves.* Math. Ann. **337** (2007), no. 4, 769–796.

[24] V. Beresnevich, S. L. Velani, *An inhomogeneous transference principle and Diophantine approximation.* Proc. Lond. Math. Soc. (3) **101** (2010), no. 3, 821–851.

[25] V. Beresnevich, S. L. Velani, *Ubiquity and a general logarithm law for geodesics*, Dynamical Systems and Diophantine Approximation. Sminaires et Congrs 20, (2009), 21-36. Editors: Yann Bugeaud, Francoise Dal'Bo and Cornelia Drutu.

[26] V. Beresnevich, S. L. Velani, *Simultaneous inhomogeneous diophantine approximation on manifolds*, Diophantine and Analytical Problems in Number Theory, Conference Proceedings in honour of A.O. Gelfond: Moscow State University, Fundam. Prikl. Mat., 16:5, 2010, 3-17.

[27] V. Beresnevich, S. L. Velani, *Classical metric Diophantine approximation revisited: the Khintchine-Groshev theorem.* International Math. Research Notes, (2010), 69–86.

[28] V. Beresnevich, S. L. Velani, *A note on three problems in metric Diophantine approximation.* Dynamical system and Ergodic Theory Conference Proceedings in honour of S.G. Dani on the occasion of his 65th birthday: University of Baroda, India 26-29 Dec 2012. Contemporary Mathematics (2015), Volume 631. Editors: Siddhartha Bhattacharya, Tarun Das, Anish Ghosh and Riddhi Shah.

[29] V. Beresnevich, S. L. Velani, *Divergence Borel-Cantelli Lemma revisited.* In preperation.

[30] V. Beresnevich, E. Zorin, *Explicit bounds for rational points near planar curves and metric Diophantine approximation.* Adv. Math. **225** (2010), no. 6, 3064–3087.

[31] V. Bernik, M. Dodson, *Metric Diophantine approximation on manifolds.* Cambridge Tracts in Mathematics, 137. Cambridge University Press, Cambridge, 1999. xii+172 pp.

[32] A. S. Besicovitch, *Sets of fractional dimensions (IV): on rational approximation to real numbers*, J. Lond. Math. Soc. **9** (1934), 126–131.

[33] E. Borel, *Sur un problème de probabilités aux fractions continues*, Math. Ann. **72** (1912), 578–584.

[34] Sky Brewer, *Continued Fractions of Algebraic Numbers.* Preprint.

[35] Y. Bugeaud, *A note on inhomogeneous Diophantine approximation*, Glasgow Math. J. **45** (2003), 105–110.

[36] Y. Bugeaud, N. Chevallier, *On simultaneous inhomogeneous Diophantine approximation*, Acta Arith. **123** (2006), 97-123.

[37] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge University Press, 1957.

[38] J. W. S. Cassels, *Some metrical theorems in Diophantine approximation. I*, Proc. Cambridge Philos. Soc., **46** (1950), 209–218.

[39] J. W. S. Cassels, H. P. F. Swinnerton-Dyer, *On the product of three homogeneous linear forms and indefinite ternary quadratic forms.* Philos. Trans. Roy. Soc. London. Ser. A, **248** (1955), 73–96.

[40] P. A. Catlin, *Two problems in metric Diophantine approximation I*, J. Number Th. **8** (1976), 282–288.

[41] H. Davenport, *A note on Diophantine approximation.* Studies in Math. Analysis and Related Topics, Stanford Univ. Press (1962) 77–82.

[42] H. Davenport, *A note on Diophantine approximation II.* Mathematika **11** (1964), 50–58

[43] H. Davenport, W. M. Schmidt, *Dirichlet's theorem on Diophantine approximation*, Inst. Alt. Mat. Symp. Math. **4** (1970), 113–132.

[44] L. Guan, J. Yu, *Badly approximable vectors in higher dimension.* Preprint: arXiv:1509.08050.

[45] H. Dickinson, S. L. Velani, *Hausdorff measure and linear forms*, J. reine angew. Math. **490** (1997), 1–36.

[46] M. Dodson, B. Rynne, J. Vickers, *Diophantine approximation and a lower bound for Hausdorff dimension.* Mathematika, **37** (1990), 59–73.

[47] M. Dodson, B. Rynne, J. Vickers, *Khintchine type theorems on manifolds.* Acta Arithmetica **57** (1991), no. 2, 115–130.

[48] R. J. Duffin, A. C. Schaeffer, *Khintchine's problem in metric Diophantine approximation*, Duke Math. J. **8** (1941), 243–255.

[49] M. Einsiedler, A. Katok, E. Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture.* Ann. of Math. (2) **164** (2006), no. 2, 513-560.

[50] K. Falconer, *The geometry of fractal sets*, Cambridge Tracts in Mathematics, No. 85, Cambridge University Press, 1985.

[51] L. Fishman, D. Simmons, M. Urbański, *Diophantine approximation and the geometry of limit sets in Gromov hyperbolic metric spaces*. Preprint: arXiv:1301.5630.

[52] P. X. Gallagher, *Approximation by reduced fractions*, J. Math. Soc. Japan **13** (1961), 342–345.

[53] P. X. Gallagher, *Metric Simultaneous Diophantine Approximation.* Jou. L.M.S., **37** (1962), 387–390.

[54] P. X. Gallagher, *Metric simultaneous Diophantine approximation II*, Mathematika **12** (1965), 123–127.

[55] A. Gorodnik, P. Vishe, *Inhomogeneous multiplicative Littlewood conjecture and* log *savings.* arXiv:1601.03525

[56] A. Gorodnik, F. Ramírez, P. Vishe, In preparation.

[57] G. H. Hardy, E. M. Wright, *An introduction to the theory of numbers*. Sixth edition, Oxford University Press, Oxford, 2008.

[58] G. Harman, *Metric number theory*, LMS Monographs New Series, vol. 18, Clarendon Press, 1998.

[59] I. Jarník, *Zur metrischen Theorie der diophantischen Appoximationen.* Prace Mat.-Fiz. **36** (1928–29), 91–106.

[60] V. Jarník, *Diophantischen Approximationen und Hausdorffsches Mass*, Mat. Sbornik **36** (1929), 371–382.

[61] V. Jarník, *Über die simultanen diophantischen Approximationen*, Math. Z. **33** (1931), 505–543.

[62] L. K. Hua, *Introduction to Number Theory*, Springer-Verlag, 1982.

[63] Jing-Jing Huang, *Rational points near planar curves and Diophantine approximation*, Advances in Mathematics **274** (2015), 490–515.

[64] A. I. Khintchine, *Einige Sätze über Kettenbruche, mit Anwendungen auf die Theorie der Diophantischen Approximationen*, Math. Ann. **92** (1924), 115–125.

[65] A. I. Khintchine, *Über die angenäherte Auflösung linearer Gleichungen in ganzen Zahlen*, Rec. math. Soc. Moscou Bd. **32** (1925), 203–218.

[66] A. Y. Khintchine, *Continued fractions*. Dover Publications, 1997.

[67] D. Kleinbock, G. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. Math. **148** (1998), 339–360.

[68] D. Kleinbock, B. Weiss, *Modified Schmidt games and Diophantine approximation with weights.* Adv. Math. **223** (2010), no.4, 1276–1298.

[69] S. Kristensen, R. Thorn, S. L. Velani, *Diophantine approximation and badly approximable sets.* Adv. Math. **203** (2006), no. 1, 132–169.

[70] J. Kurzweil, *On the metric theory of inhomogeneous Diophantine approximations*, Studia mathematica **XV** (1955), 84-112.

[71] S. Lang, H. Trotter, *Continued fractions for some algebraic numbers.* Introduction to Diophantine Approximations, page 93, 1972.

[72] P. Mattila, *Geometry of sets and measures in Euclidean space*, Cambridge University Press, 1995.

[73] L. G. Peck, *Simultaneous rational approximations to algebraic numbers.* Bull. A.M.S. **67** (1961), 197–201.

[74] O. Perron, *Über diophantische Approximationen.* Math. Ann. **83** (1921), 77–84.

[75] A. D. Pollington, R. C. Vaughan, *The k-dimensional Duffin and Schaeffer conjecture.* Mathematika **37** (1990), 190–200.

[76] A. D. Pollington, S. L. Velani, *On a problem in simultaneously Diophantine approximation: Littlewood's conjecture.* Acta Math. **66** (2000), 29–40.

[77] A. Pollington, S. L. Velani, *On simultaneously badly approximable numbers.* J. London Math. Soc. (2) **66** (2002), no.1, 29–40.

[78] F. Ramírez, *Khintchine types of translated cooordinate hyperplanes.* Acta Arith. to appear.

[79] F. Ramírez, D. Simmons, F. Süess, *Rational approximation of affine coordinate subspaces of Euclidean spaces.* Pre-print: arXiv:1510.05012.

[80] A. Rockett, P. Szüsz, *Continued Fractions.* World Scientific, Singapore, 1992.

[81] K. F. Roth, *On irregularities of distribution.* Mathematika **7** (1954), 73–79

[82] K. F. Roth, *Rational approximation to algebraic numbers.* Mathematika **2** (1955), 1–20, with a corrigendum on p. 168.

[83] J. Schmeling, S. Troubetzkoy, *Inhomogeneous Diophantine Approximation and Angular Recurrence for Polygonal Billards.* Math. Sbornik **194** (2003), 295–309.

[84] W. M. Schmidt, *Metrical theorems on fractional parts of sequences.* Trans. Amer. Math. Soc. **110** (1964), 493–518.

[85] W. M. Schmidt, *On badly approximable numbers and certain games.* Trans. Amer. Math. Soc. **123** (1966), 178–199.

[86] W. M. Schmidt, *Irregularities of Distribution. VII.* Acta. Arith. **21** (1972), 45—50.

[87] W. M. Schmidt, *Diophantine approximation.* Lecture notes in Maths. 785, Springer – Verlag, (1975).

[88] N. A. Shah, *Equidistribution of expanding translates of curves and Dirichlet's theorem on Diophantine approximation*. Invent. Math. **177** (2009), no. 3, 509–532.

[89] U. Shapira, *A solution to a problem of Cassels and Diophantine properties of cubic numbers.* Ann. of Math. (2) **173** (2011), no. 1, 543–557.

[90] V. G. Sprindžuk, *Metric theory of Diophantine approximations*. John Wiley, 1979, Translated by R. A. Silverman.

[91] R. C. Vaughan, S. L. Velani, *Diophantine approximation on planar curves: the convergence theory.* Invent. Math. **166** (2006), 103–124.

[92] M. Waldschmidt, *Open Diophantine Problems*. Moscow Mathematical Journal **4** (2004), 245–305.

Victor V. Beresnevich:
Department of Mathematics, University of York,
Heslington, York, YO10 5DD, England
E-mail: `victor.beresnevich@york.ac.uk`

Felipe A. Ramírez:
Department of Mathematics and Computer Science, Wesleyan University,
265 Church Street Middletown, CT 06459
E-mail: `framirez@wesleyan.edu`

Sanju L. Velani:
Department of Mathematics, University of York,
Heslington, York, YO10 5DD, England
E-mail: `sanju.velani@york.ac.uk`