Schmidt's theorem, Hausdorff measures and Slicing

Victor Beresnevich[*] Sanju Velani[†]
York York

July 18, 2005

For Bridget Bennett on her fortieth birthday

Abstract

A Hausdorff measure version of W.M. Schmidt's inhomogeneous, linear forms theorem in metric number theory is established. The key ingredient is a 'slicing' technique motivated by a standard result in geometric measure theory. In short, 'slicing' together with the Mass Transference Principle [3] allows us to transfer Lebesgue measure theoretic statements for lim sup sets associated with linear forms to Hausdorff measure theoretic statements. This extends the approach developed in [3] for simultaneous approximation. Furthermore, we establish a new Mass Transference Principle which incorporates both forms of approximation. As an application we obtain a complete metric theory for a 'fully' non-linear Diophantine problem within the linear forms setup.

2000 *Mathematics Subject Classification*: Primary 11J83, 28A78; Secondary 11J13, 11K60

Keywords and phrases: Inhomogeneous Diophantine approximation, linear forms, Hausdorff measure and dimension.

1 Introduction

Fix a vector $\mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{R}^m$ and a non–negative, real valued function $\Psi : \mathbb{Z}^n \to \mathbb{R}^+ := \{x \geq 0 : x \in \mathbb{R}\}$ such that

$$\Psi(\mathbf{a}) \to 0 \qquad \text{as} \qquad |\mathbf{a}| := \max(|a_1|, \ldots, |a_n|) \to \infty\ .$$

Let $W^{\mathbf{b}}_{n,m}(\Psi)$ be the set of $X = (\mathbf{x}_1, \ldots, \mathbf{x}_m) \in \mathbb{I}^{n\times m} := [0,1)^{n\times m}$ where $\mathbf{x}_j \in \mathbb{R}^n$ for $1 \leq j \leq m$, such that the system of inequalities

$$\|\mathbf{a}\cdot\mathbf{x}_j - b_j\| \ < \ \Psi(\mathbf{a}) \qquad\qquad (1 \leq j \leq m) \tag{1}$$

[*]Research supported by EPSRC Grant R90727/01
[†]Royal Society University Research Fellow

is satisfied for infinitely many $\mathbf{a} \in \mathbb{Z}^n$. Here and throughout $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$ is the standard inner product of two vectors $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ and $\|x\|$ is the distance from $x \in \mathbb{R}$ to the nearest integer. The following is a geometric interpretation of the set $W^{\mathbf{b}}_{n,m}(\Psi)$ and brings to the forefront its lim sup nature. For vectors $\mathbf{a} \in \mathbb{Z}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{p} \in \mathbb{Z}^m$, consider the $(n-1)m$–dimensional plane $R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}}$ given by

$$R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}} := \{X \in \mathbb{R}^{n\times m} : \ \mathbf{a} \cdot \mathbf{x}_j - b_j = p_j \ \ (1 \leq j \leq m)\} \ . \tag{2}$$

Thus, $R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}}$ is the product of the n–dimensional hyperplanes given by

$$R^{b_j}_{\mathbf{a},p_j} := \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{a} \cdot \mathbf{x} - b_j = p_j\} \qquad 1 \leq j \leq m \ .$$

For $\delta \geq 0$, let $\Delta(R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}}, \delta)$ denote the δ–neighborhood of $R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}}$; i.e. the product of the δ-neighborhoods (with respect to the Euclidean norm) of the hyperplanes $R^{b_j}_{\mathbf{a},p_j}$. Note that when $n = 1$, the δ–neighborhood $\Delta(R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}}, \delta)$ is simply a ball of radius δ in the supremum norm. It is easily verified, that

$$X \in W^{\mathbf{b}}_{n,m}(\Psi) \qquad \text{if and only if} \qquad X \in \Delta\Big(R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}}, \tfrac{\Psi(\mathbf{a})}{\sqrt{\mathbf{a}.\mathbf{a}}}\Big) \cap \mathbb{I}^{n\times m}$$

for infinitely many vectors $\mathbf{a} \in \mathbb{Z}^n$ and $\mathbf{p} \in \mathbb{Z}^m$.

1.1 The Lebesgue measure theory: Schmidt's Theorem

The following key result provides a beautiful and simple criteria for the 'size' of the set $W^{\mathbf{b}}_{n,m}(\Psi)$ expressed in terms of $n \times m$–dimensional Lebesgue measure $|\ \ |_{n\times m}$. The theorem is due to W.M. Schmidt and shows that $|W^{\mathbf{b}}_{n,m}(\Psi)|_{n\times m}$ satisfies an elegant 'zero–one' law.

Schmidt's Theorem [12] *Let Ψ be as above and $n + m > 2$. Then*

$$|W^{\mathbf{b}}_{n,m}(\Psi)|_{n\times m} = \begin{cases} 0 & \text{if} \quad \displaystyle\sum_{\mathbf{a}\in\mathbb{Z}^n} \Psi(\mathbf{a})^m < \infty \\ 1 & \text{if} \quad \displaystyle\sum_{\mathbf{a}\in\mathbb{Z}^n} \Psi(\mathbf{a})^m = \infty \end{cases} .$$

In fact, Schmidt considers a more general setup for which he obtains a quantitative result. The case that $m + n = 2$, corresponding to $n = m = 1$, is naturally excluded since the statement is known to be false – the Duffin–Schaeffer conjecture [3, 13] provides the appropriate 'expected' statement.

In order to appreciate the true significance of Schmidt's Theorem it is well worth mentioning a few special cases which in their own right represent landmarks within the classical theory of metric Diophantine approximation.

Khintchine's Theorem [10] : $n = 1$ and $\mathbf{b} = 0$ with Ψ monotonic.
(simultaneous, homogeneous approximation)

Groshev's Theorem [6] : $n > 1$ and $\mathbf{b} = 0$ with Ψ monotonic[1]. (linear forms, homogeneous approximation)

Gallagher's Theorem [5] : $n = 1$ and $m \geq 2$. (simultaneous, inhomogeneous approximation)

Note that under the condition that Ψ is monotonic, the results of Khintchine and Groshev already give the homogeneous version ($\mathbf{b} = 0$) of Schmidt's Theorem without the condition that $m + n > 2$ – Khintchine's Theorem covers the case $n = m = 1$. Generalizing the Khintchine–Groshev statement to the inhomogeneous case and entirely removing the monotonicity condition when $m + n > 2$ is by no means a trivial feat. As with Schmidt, Gallagher considers a more general setup for which he obtains a quantitative result.

1.2 The general metric theory

Let f be a dimension function and let $\mathcal{H}^f$ denote the Hausdorff f–measure – see §2.1. In short, our aim is to provide a complete metric theory for the set $W^{\mathbf{b}}_{n,m}(\Psi)$. The following result achieves this goal in that it provides a simple criteria for the 'size' of the set $W^{\mathbf{b}}_{n,m}(\Psi)$ expressed in terms of the general measure $\mathcal{H}^f$.

Theorem 1 *Let Ψ be as above and $n + m > 2$. Let f be a dimension function such that $r^{-nm} f(r)$ is monotonic. Furthermore, assume that $g : r \to r^{-(n-1)m} f(r)$ is a dimension function. Then*

$$\mathcal{H}^f(W^{\mathbf{b}}_{n,m}(\Psi)) = \begin{cases} 0 & \text{if} \quad \displaystyle\sum_{\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} g\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times |\mathbf{a}|^m < \infty \\ \mathcal{H}^f(\mathbb{I}^{n \times m}) & \text{if} \quad \displaystyle\sum_{\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} g\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times |\mathbf{a}|^m = \infty \end{cases} .$$

Notice that in the case $\mathcal{H}^f$ is $n \times m$–dimensional Lebesgue measure $| \ |_{n \times m}$, the theorem reduces to Schmidt's Theorem. As with the Lebesgue theory, the convergence part of the above theorem is relatively straightforward if not trivial – see §3.1. The main substance is the divergent part. For this, our particular strategy is straightforward enough. We establish the following:

Theorem 2 *Schmidt's Theorem (divergent part)* $\Longrightarrow$ *Theorem 1 (divergent part)* .

In [3], this strategy has recently been successfully implemented to establish the simultaneous version of Theorem 2; namely

[1]When $n > 1$, the function Ψ is in general multi-variable. To say that it is monotonic simply means that $\Psi(\mathbf{a}) := \psi(|\mathbf{a}|)$ for some monotonic ψ.

Theorem BV *Gallagher's Theorem (divergent part)* $\Longrightarrow$ *Theorem 1 (divergent part) with $n=1$ and $m \geq 2$.*

Recall, that Schmidt's Theorem reduces to Gallagher's Theorem in the case of simultaneous approximation ($n = 1$). To be absolutely precise, in [3, §6.2] we only consider the homogeneous case of Theorem BV. However, given the method of proof adopted in [3] no extra obstacles appear in establishing the inhomogeneous version above. Indeed, the proof is essentially a simple application of the Mass Transference Principle (see §2.2) and for this it is irrelevant whether we start with a homogeneous or inhomogeneous divergent statement of Gallagher's Theorem. The only relevant aspect is that when $n = 1$, the set $W^{\mathbf{b}}_{n,m}(\Psi)$ is a limsup set naturally defined in terms of a sequence of balls – the Mass Transference Principle then does the rest! This is no longer the case when $n > 1$ and so Theorem 2 is not simply a consequence of the approach developed in [3]; namely the Mass Transference Principle.

The key aspect of this paper is the introduction of a 'slicing' technique to the theory of metric Diophantine approximation; in particular to the linear forms aspect of the theory. The technique is motivated by a relatively standard result in geometric measure theory – see §2.3. The upshot is that 'slicing' together with the Mass Transference Principle yields Theorem 2 – the 'hard' part of Theorem 1.

Remark. In all previous contributions towards the general metric theory, such as the pioneering work of Jarník [8] the function Ψ is assumed to be monotonic. For further details and references the reader is refereed to [1, Sections 1.1 & 12.1].

Before moving on, it is useful to say a little concerning the condition imposed on g in Theorem 1; namely that since g is assumed to be a dimension function we have that $g(r) \to 0$ as $r \to 0$. For the sake of clarity and ease of discussion, put $f : r \to r^s$ ($s > 0$) in Theorem 1. Then, Theorem 1 reduces to the following s–dimensional Hausdorff measure statement which in its own right is of significant importance since it characterizes the Hausdorff dimension of the set $W^{\mathbf{b}}_{n,m}(\Psi)$ as the exponent of convergence of a certain 's–volume' sum.

Corollary 1 *Let Ψ be as above and $n + m > 2$. Let $\delta > 0$ and $s := (n-1)m + \delta$. Then*

$$\mathcal{H}^s(W^{\mathbf{b}}_{n,m}(\Psi)) = \begin{cases} 0 & \text{if} \quad \displaystyle\sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \Psi(\mathbf{a})^\delta \ |\mathbf{a}|^{m-\delta} \ < \ \infty \\ \mathcal{H}^s(\mathbb{I}^{n\times m}) & \text{if} \quad \displaystyle\sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \Psi(\mathbf{a})^\delta \ |\mathbf{a}|^{m-\delta} \ = \ \infty \end{cases} .$$

It follows from the definition of Hausdorff dimension (see §2.1) that if for some $\delta > 0$ the sum in the corollary diverges, then

$$\dim W^{\mathbf{b}}_{n,m}(\Psi) \ = \ \inf\left\{ s : \textstyle\sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} \Psi(\mathbf{a})^{s-(n-1)m} \ |\mathbf{a}|^{nm-s} \ < \ \infty \right\} .$$

We suspect that the condition on s imposed in Corollary 1, namely that s is strictly greater than $(n-1)m$, cannot be relaxed. Briefly, if $\delta = 0$ so that $s = (n-1)m$, the sum in Corollary 1 diverges irrespective of Ψ. Now the 'approximating' hyperplanes as defined by (2) are themselves of dimension s and indeed of positive $\mathcal{H}^s$ measure. Thus, it is highly likely that for rapidly decreasing functions Ψ the $\mathcal{H}^s$–measure theoretic structure of $W^{\mathbf{b}}_{n,m}(\Psi)$ is purely dependent on the arithmetic properties of the approximating hyperplanes. In view of this, for appropriate Ψ and $\mathbf{b} \in \mathbb{R}^m$ one might expect that $\mathcal{H}^s(W^{\mathbf{b}}_{n,m}(\Psi))$ is finite and possibly even zero rather than $\mathcal{H}^s(\mathbb{I}^{n\times m})$ which is infinite.

2 Preliminaries

2.1 Hausdorff measures

In this section we give a brief account of Hausdorff measures. For further details see [11]. A *dimension function* $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous, non-decreasing function such that $f(r) \to 0$ as $r \to 0$. The Hausdorff f–measure with respect to the dimension function f will be denoted throughout by $\mathcal{H}^f$ and is defined as follows. Suppose F is a subset of $\mathbb{R}^k$. For $\rho > 0$, a countable collection $\{B_i\}$ of balls in $\mathbb{R}^k$ with radius $r(B_i) \leq \rho$ for each i such that $F \subset \bigcup_i B_i$ is called a *ρ-cover for F*. For a dimension function 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. The *Hausdorff f–measure* $\mathcal{H}^f(F)$ of F with respect to the dimension function f is defined by

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

A simple consequence of the definition of $\mathcal{H}^f$ is the following useful fact.

Lemma 1 *If f and g are two dimension functions such that the ratio $f(r)/g(r) \to 0$ as $r \to 0$, then $\mathcal{H}^f(F) = 0$ whenever $\mathcal{H}^g(F) < \infty$.*

In the case that $f(r) = r^s$ $(s > 0)$, the measure $\mathcal{H}^f$ is the usual s–dimensional Hausdorff measure $\mathcal{H}^s$ and the Hausdorff dimension $\dim F$ of a set F is defined by

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

In particular when s is an integer, $\mathcal{H}^s$ is comparable[2] to the s–dimensional Lebesgue measure. Actually, $\mathcal{H}^s$ is a constant multiple of the s–dimensional Lebesgue measure.

[2]The symbols $\ll$ and $\gg$ will be used to indicate an inequality with an unspecified positive constant. If $a \ll b$ and $a \gg b$ we write $a \asymp b$, and say that the quantities a and b are comparable.

2.2 The Mass Transference Principle

Given a dimension function f and a ball $B = B(x, r)$ in $\mathbb{R}^m$, we define another ball

$$B^f := B(x, f(r)^{1/m}) \ . \tag{3}$$

When $f(x) = x^s$ for some $s > 0$ we also adopt the notation B^s, *i.e.* $B^s := B^{(x \mapsto x^s)}$. It is readily verified that

$$B^m = B. \tag{4}$$

Given a sequence of balls B_i, $i = 1, 2, 3, \ldots$, as usual its limsup set is

$$\limsup_{i \to \infty} B_i := \bigcap_{j=1}^{\infty} \bigcup_{i \geqslant j} B_i \ .$$

By definition, $\limsup_{i \to \infty} B_i$ is precisely the set of points in $\mathbb{R}^m$ which lie in infinitely many balls B_i.

The following Mass Transference Principle allows us to transfer Lebesgue measure theoretic statements for lim sup subsets of $\mathbb{R}^m$ to Hausdorff measure theoretic statements.

Lemma 2 (Mass Transference Principle) *Let $\{B_i\}_{i \in \mathbb{N}}$ be a sequence of balls in $\mathbb{R}^m$ with $r(B_i) \to 0$ as $i \to \infty$. Let f be a dimension function such that $r^{-m} f(r)$ is monotonic and let Ω be a ball in $\mathbb{R}^m$. Suppose for any ball B in Ω*

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

Then, for any ball B in Ω

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

With $\Omega = \mathbb{R}^m$, the lemma is precisely Theorem 2 in [3]. It is easily seen that this implies the above modified statement which is better suited for the particular applications we have in mind.

2.3 The 'slicing' lemma

The 'slicing' lemma below is the crucial new ingredient and together with the 'slicing' technique (§3.2) makes it possible to reduce Theorem 2 to an m-dimensional problem by slicing the original set $W^{\mathbf{b}}_{n,m}(\Psi)$ into a family of subsets lying on parallel m-dimensional planes. The 'slicing' technique is motivated by the 'slicing' lemma.

In the following V will be a linear subspace of $\mathbb{R}^k$ and $V^{\perp}$ will denote the linear subspace of $\mathbb{R}^k$ orthogonal to V. Also, $V + a := \{v + a : v \in V\}$ for $a \in V^{\perp}$.

Lemma 3 *Let $l, k \in \mathbb{N}$ such that $l \leq k$ and f and $g : r \mapsto r^{-l} f(r)$ be dimension functions. Let $A \subset \mathbb{R}^k$ be a Borel set with $\mathcal{H}^f(A) < \infty$. Then for any $(k-l)$-dimensional linear subspace V of $\mathbb{R}^k$,*

$$\mathcal{H}^g(A \cap (V + a)) < \infty \quad \textit{for } \mathcal{H}^l\textit{-almost all } a \in V^\perp.$$

When $f : r \to r^s$, the lemma constitutes the first part of Theorem 10.10 in [11]. The proof given there can be easily modified to yield the more general statement above. Nevertheless, given the importance of the lemma and for the sake of completeness we have included the proof of Lemma 3 as an appendix.

Trivially, Lemma 3 implies the following:

Lemma 4 (Slicing lemma) *Let $l, k \in \mathbb{N}$ such that $l \leq k$ and f and $g : r \mapsto r^{-l} f(r)$ be dimension functions. Let $A \subset \mathbb{R}^k$ be a Borel set and V be an $(k-l)$-dimensional linear subspace of $\mathbb{R}^k$. If for a subset S of $V^\perp$ of positive $\mathcal{H}^l$-measure*

$$\mathcal{H}^g(A \cap (V + b)) = \infty \qquad \textit{for all} \quad b \in S\,,$$

then $\mathcal{H}^f(A) = \infty$.

2.4 Additional assumption in Schmidt's theorem (divergent part)

Let $n \geq 2$ as otherwise the substance of this section becomes trivial. For each $i \in \{1, \ldots, n\}$ define the subset $\mathcal{Z}_i$ of $\mathbb{Z}^n \setminus \{\mathbf{0}\}$ to consist of vectors $\mathbf{a} \in \mathbb{Z}^n$ such that $|\mathbf{a}| = |a_i|$. Assume that

$$\sum_{\mathbf{a} \in \mathbb{Z}^n} \Psi(\mathbf{a})^m = \infty.$$

Now

$$\infty = \sum_{\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} \Psi(\mathbf{a})^m \leqslant \sum_{i=1}^{n} \sum_{\mathbf{a} \in \mathcal{Z}_i} \Psi(\mathbf{a})^m.$$

Therefore there is an index $i \in \{1, \ldots, n\}$ such that

$$\sum_{\mathbf{a} \in \mathcal{Z}_i} \Psi(\mathbf{a})^m = \infty. \tag{5}$$

Define

$$\Psi_i(\mathbf{a}) = \begin{cases} \Psi(\mathbf{a}), & \text{if } \mathbf{a} \in \mathcal{Z}_i, \\ 0, & \text{if } \mathbf{a} \notin \mathcal{Z}_i. \end{cases}$$

Hence, $|W^{\mathbf{b}}_{n,m}(\Psi_i)|_{n \times m} = 1$ by Schmidt's Theorem. Trivially, this implies that for almost all $X \in \mathbb{I}^{n \times m}$

$$\max_{1 \leqslant j \leqslant m} \|\mathbf{a} \cdot \mathbf{x}_j - b_j\| \ < \ \Psi(\mathbf{a}) \tag{6}$$

for infinitely many $\mathbf{a} \in \mathcal{Z}_i$. There is no loss of generality in assuming that (5) is satisfied with $i = 1$ as otherwise we can apply a permutation of variables (columns in X) under which Schmidt's theorem is clearly invariant. Thus, when considering the divergent part of Schmidt's theorem we can assume that

$$\Psi(\mathbf{a}) = 0 \quad \forall \ \mathbf{a} \in \mathbb{Z}^n \quad \text{with} \quad |\mathbf{a}| \neq |a_1| \ .$$

3 Proof of Theorem 1

3.1 The case of convergence

We are given that

$$\sum_{\mathbf{a} \in \mathbb{Z}^n \smallsetminus \{\mathbf{0}\}} g\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times |\mathbf{a}|^m < \infty \ .$$

The convergent part of Theorem 1 follows on using standard covering arguments. For each $N \in \mathbb{N}$, it is easily verified that

$$W_{n,m}^{\mathbf{b}}(\Psi) \subset \bigcup_{\substack{\mathbf{a} \in \mathbb{Z}^n : |\mathbf{a}| \geq N \\ \Psi(\mathbf{a}) > 0}} \ \bigcup_{\mathbf{p} \in \mathbb{Z}^m} \Delta\left(R_{\mathbf{a},\mathbf{p}}^{\mathbf{b}}, \tfrac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \cap \mathbb{I}^{n \times m} \ .$$

Note that there is no loss of generality in assuming that $\Psi(\mathbf{a}) > 0$ in the above union, for otherwise (1) has no solutions X and the integer vector $\mathbf{a}$ makes no contribution to $W_{n,m}^{\mathbf{b}}(\Psi)$.

Next notice that for any fixed $\mathbf{a} \in \mathbb{Z}^n \smallsetminus \{\mathbf{0}\}$ with $\Psi(\mathbf{a}) > 0$ and $\mathbf{p} \in \mathbb{Z}^m$, it is possible to cover

$$\Delta\left(R_{\mathbf{a},\mathbf{p}}^{\mathbf{b}}, \tfrac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \cap \mathbb{I}^{n \times m}$$

by a collection $\mathcal{C}_{\mathbf{a},\mathbf{p}}^{\mathbf{b}}$ of balls of common radius $\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}$ such that

$$\# \mathcal{C}_{\mathbf{a},\mathbf{p}}^{\mathbf{b}} \ll \left(\tfrac{|\mathbf{a}|}{\Psi(\mathbf{a})}\right)^{(n-1)m} \ .$$

Also, for a fixed $\mathbf{a} \in \mathbb{Z}^n \smallsetminus \{\mathbf{0}\}$

$$\# \left\{ \mathbf{p} \in \mathbb{Z}^m : \Delta\left(R_{\mathbf{a},\mathbf{p}}^{\mathbf{b}}, \tfrac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \cap \mathbb{I}^{n \times m} \neq \emptyset \right\} \ll |\mathbf{a}|^m \ .$$

Finally, note that since $\Psi(\mathbf{a}) \to 0$ as $|\mathbf{a}| \to \infty$, we have that for all N sufficiently large

$$\frac{\Psi(\mathbf{a})}{|\mathbf{a}|} \leq \frac{1}{N} \ .$$

It now follows from the definition of $\mathcal{H}^f$ that for N sufficiently large,

$$\begin{aligned} \mathcal{H}^f_{\rho := \frac{1}{N}}\left(W_{n,m}^{\mathbf{b}}(\Psi)\right) &\ll \sum_{\mathbf{a} \in \mathbb{Z}^n : |\mathbf{a}| \geq N} f\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times \left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right)^{-(n-1)m} \times |\mathbf{a}|^m \\ &:= \sum_{\mathbf{a} \in \mathbb{Z}^n : |\mathbf{a}| \geq N} g\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times |\mathbf{a}|^m \quad \to \quad 0 \quad \text{as} \quad N \to \infty \ . \end{aligned}$$

Thus, $\mathcal{H}^f(W^{\mathbf{b}}_{n,m}(\Psi)) = 0$ as required. $\square$

3.2 The case of divergence (Theorem 2): the 'slicing' technique

Throughout we assume that $n \geq 2$. Theorem BV covers the $n = 1$ case. We are given that

$$\sum_{\mathbf{a}\in\mathbb{Z}^n\setminus\{\mathbf{0}\}} g\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times |\mathbf{a}|^m \ = \ \infty \ . \tag{7}$$

We start by considering the case that $r^{-mn}f(r) \to L$ as $r \to 0$ and L is finite. If $L = 0$, then Lemma 1 implies that $\mathcal{H}^f(\mathbb{I}^{n\times m}) = 0$ and since $W^{\mathbf{b}}_{n,m}(\Psi) \subset \mathbb{I}^{n\times m}$ the result follows. If $L \neq 0$, then $\mathcal{H}^f$ is comparable to $\mathcal{H}^{mn}$ (in fact, $\mathcal{H}^f = L\,\mathcal{H}^{mn}$). In turn, $\mathcal{H}^{mn}$ is comparable to $n \times m$–dimensional Lebesgue measure and so the required statement follows on showing that $|W^{\mathbf{b}}_{n,m}(\Psi)|_{n\times m} = |\mathbb{I}^{n\times m}|_{n\times m}$. Well, this is a simple consequence of Schmidt's theorem since the sum appearing in (7) is comparable to $\sum_{\mathbf{a}\in\mathbb{Z}^n} \Psi(\mathbf{a})^m$.

In view of the above discussion, we can assume without loss of generality that

$$r^{-mn}f(r) \ \to \ \infty \qquad \text{as} \qquad r \to 0 \ . \tag{8}$$

Indeed, it is this situation that constitutes the main substance of Theorems 1 and 2. Trivially, (8) together with Lemma 1 implies that

$$\mathcal{H}^f(\mathbb{I}^{n\times m}) \ = \ \infty \ .$$

Similarly, since (8) is equivalent to the statement that $r^{-m}g(r) \to \infty$ as $r \to 0$, we have that $\mathcal{H}^g(B) = \infty$ for any m–dimensional ball B.

To proceed, we set

$$\widetilde{\Psi}(\mathbf{a})^m \ := \ g\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times |\mathbf{a}|^m \ .$$

In view of (7), it follows that

$$\sum_{\mathbf{a}\in\mathbb{Z}^n} \widetilde{\Psi}(\mathbf{a})^m = \infty \ ,$$

and Schmidt's Theorem (divergent part) implies that

$$|W^{\mathbf{b}}_{n,m}(\widetilde{\Psi})|_{n\times m} = 1 \ .$$

The goal is to show that this implies that

$$\mathcal{H}^f(W^{\mathbf{b}}_{n,m}(\Psi)) \ = \ \infty \ ;$$

i.e. to establish Theorem 2 under the condition imposed by (8). Recall, that the conclusion of Theorem 2 is precisely the divergent part of Theorem 1.

In view of the discussion in §2.4, we can assume without loss of generality that

$$\widetilde{\Psi}(\mathbf{a}) \ = \ 0 \qquad \forall \ \ \mathbf{a} \in \mathbb{Z}^n \qquad \text{with} \qquad |\mathbf{a}| \neq |a_1| \ . \tag{9}$$

Now, let

$$V = \{(\mathbf{x}_1, \dots, \mathbf{x}_m) \ : \ x_{j,i} = 0 \ \ \forall\, j = 1, \dots, m\,; i = 2, \dots, n, \ \},$$

where $\mathbf{x}_j = (x_{j,1}, \dots, x_{j,n})$. Thus, V is an m–dimensional subspace of $\mathbb{R}^{n\times m}$. By Fubini's theorem, there is a subset $S \subset \mathbb{I}^{m(n-1)} \subset V^{\perp}$ with $|S|_{m(n-1)} = 1$ such that for every $X_0 \in S$ the set $W^{\mathbf{b}}_{n,m}(\widetilde{\Psi})$ has full m-dimensional Lebesgue measure in $(V + X_0) \cap \mathbb{I}^{n\times m}$; i.e.

$$|(V + X_0) \cap W^{\mathbf{b}}_{n,m}(\widetilde{\Psi})|_m \ = \ 1 \ . \tag{10}$$

For every $(\mathbf{p}, \mathbf{a}) = (p_1, \dots p_m; a_1, \dots, a_n) \in \mathbb{Z}^m \times \mathbb{Z}^n$ define the set $\sigma(\mathbf{p}, \mathbf{a})$ to consist of $X \in \mathbb{I}^{n\times m}$ such that

$$\max_{1\leqslant j\leqslant m} |\mathbf{a}\cdot \mathbf{x}_j - b_j + p_j| < \widetilde{\Psi}(\mathbf{a}) \ .$$

In view of condition (9) imposed on $\widetilde{\Psi}$, the set $\sigma(\mathbf{p}, \mathbf{a})$ is empty whenever $|\mathbf{a}| \neq |a_1|$. We therefore assume that $|\mathbf{a}| = |a_1|$ throughout the rest of the proof. Then

$$\sigma(\mathbf{p}, \mathbf{a}) \cap (V + X_0) \tag{11}$$

is the product of m intervals of length $2\widetilde{\Psi}(\mathbf{a})/|a_1|$. Indeed, for all $j = 1, \dots, m$ and $i = 2, \dots, n$ we have that $x_{j,i}$ are fixed and defined by X_0 for all points in this set. That is to say that the only coordinates that may vary are $x_{j,1}$. Therefore, the set (11) is defined by the system

$$\max_{1\leqslant j\leqslant m} |a_1 x_{j,1} + (a_2 x_{j,2} + \dots + a_n x_{j,n} + p_j - b_j)| < \widetilde{\Psi}(\mathbf{a})$$

or equivalently

$$\max_{1\leqslant j\leqslant m} \left| x_{j,1} - \frac{b_j - (a_2 x_{j,2} + \dots + a_n x_{j,n} + p_j)}{a_1} \right| < \frac{\widetilde{\Psi}(\mathbf{a})}{|a_1|} \ . \tag{12}$$

The pathological situation of $a_1 = 0$ is excluded by the conditions $|\mathbf{a}| = |a_1|$ and $\mathbf{a} \neq \mathbf{0}$. On identifying $\mathbb{I}^{n\times m}$ with the $n \times m$-dimensional torus it is easily seen that every inequality in (12) defines an interval of length $2\widetilde{\Psi}(\mathbf{a})/|a_1|$. Thus (11) defines a ball of radius $\widetilde{\Psi}(\mathbf{a})/|\mathbf{a}|$. Such balls form a sequence $(A_i)_{i\in\mathbb{N}}$. Therefore

$$\limsup_{i\to\infty} A_i = (V + X_0) \cap W^{\mathbf{b}}_{n,m}(\widetilde{\Psi}) \ .$$

Hence, in view of (10) we have that $|\limsup_{i\to\infty} A_i|_m = 1$. This implies that for any ball $B \subset (V + X_0) \cap \mathbb{I}^{n\times m}$

$$\mathcal{H}^m(\limsup_{i\to\infty} A_i \cap B) = \mathcal{H}^m(B)\,.$$

For each ball A_i define the ball B_i with the same centre and radius $\Psi(\mathbf{a})/|\mathbf{a}|$. Then, by definition $B_i^g = A_i$ – see (3). It follows that

$$\limsup_{i\to\infty} B_i \ \subset \ (V + X_0) \cap W^{\mathbf{b}}_{n,m}(\Psi) \ . \tag{13}$$

Also, $r^{-m}g(r) = r^{-mn}f(r)$ is monotonic by assumption. Thus, on applying the Mass Transference Principle with $\Omega = (V + X_0) \cap \mathbb{I}^{n\times m}$, we obtain that for any ball B in Ω

$$\mathcal{H}^g(\limsup_{i\to\infty} B_i^m \cap B) = \mathcal{H}^g(B) = \infty \ .$$

Recall, that $\mathcal{H}^g(B) = \infty$ is a consequence of (8) and Lemma 1. Hence, in view of (13) and the fact that $B_i^m := B_i$ we have that for every $X_0 \in S$

$$\mathcal{H}^g((V + X_0) \cap W^{\mathbf{b}}_{n,m}(\Psi)) = \infty \ .$$

Recall that $\mathcal{H}^{m(n-1)}(S) > 0$ and so by the Slicing lemma,

$$\mathcal{H}^f(W^{\mathbf{b}}_{n,m}(\Psi)) = \infty \ .$$

This completes the proof of Theorem 2 and therefore the divergent part of Theorem 1.

$\square$

4 A Mass Transference Principle for linear forms

The Mass Transference Principle deals with lim sup sets which are defined as a sequence of balls. However, we have seen that together with the 'slicing' technique introduced in this paper we are able to deal with lim sup sets defined as a sequence of neighborhoods of 'approximating' planes – at least within the context of Schmidt's Theorem. In short, the aim of this section is to develop a single framework which enables us to combine the Mass Transference Principle and 'slicing' into a single statement. The main result (Theorem 3 below) should be viewed as a generalization to the linear forms setup of the Mass Transference Principle developed in [3] for simultaneous approximation. As applications, we deduce Theorem 2 (which constitutes the main substance of Theorem 1) as a simple corollary and more strikingly, we obtain a complete metric theory for a 'fully' non-linear Diophantine problem – see Theorem 4 of §4.6.

4.1 A general framework for approximating by planes

Throughout $k, m \geq 1$ and $l \geqslant 0$ are integers such that $k = m + l$. Let $\mathcal{R} = (R_\alpha)_{\alpha \in J}$ be a family of planes in $\mathbb{R}^k$ of common dimension l indexed by an infinite countable set J. For every $\alpha \in J$ and $\delta \geq 0$ define

$$\Delta(R_\alpha, \delta) := \{\mathbf{x} \in \mathbb{R}^k : \operatorname{dist}(\mathbf{x}, R_\alpha) < \delta\} \ .$$

Thus $\Delta(R_\alpha, \delta)$ is simply the δ–neighborhood of the l–dimensional plane R_α. Note that by definition, $\Delta(R_\alpha, \delta) = \emptyset$ if $\delta = 0$. Next, let

$$\Upsilon : J \to \mathbb{R}^+ : \alpha \mapsto \Upsilon(\alpha) := \Upsilon_\alpha$$

be a non-negative, real valued function on J. In order to avoid pathological situations within our framework, we assume that for every $\epsilon > 0$ the set $\{\alpha \in J : \Upsilon_\alpha > \epsilon\}$ is finite. This condition implies that $\Upsilon_\alpha \to 0$ as α runs through J. We now consider the following 'lim sup' set,

$$\Lambda(\Upsilon) = \{\mathbf{x} \in \mathbb{R}^k : \mathbf{x} \in \Delta(R_\alpha, \Upsilon_\alpha) \text{ for infinitely many } \alpha \in J\} \ .$$

Note that in view of the conditions imposed on k, l and m we have that $l < k$. Thus the dimension of the 'approximating' planes R_α is strictly less than that of the ambient space $\mathbb{R}^k$. The situation when $l = k$ is of little interest and has therefore been naturally omitted.

4.2 The main result

Theorem 3 *Let $\mathcal{R}$ and Υ as above be given. Let V be a linear subspace of $\mathbb{R}^k$ such that* $\dim V = m = \operatorname{codim} \mathcal{R}$ *and*

$$(i) \quad V \cap R_\alpha \neq \emptyset \quad \textit{for all } \alpha \in J \ ,$$

$$(ii) \quad \sup_{\alpha \in J} \operatorname{diam}(V \cap \Delta(R_\alpha, 1)) \ < \ \infty \ .$$

Let f and $g : r \to g(r) := r^{-l} f(r)$ be dimension functions such that $r^{-k} f(r)$ is monotonic and let Ω be a ball in $\mathbb{R}^k$. Suppose for any ball B in Ω

$$\mathcal{H}^k\big(B \cap \Lambda\big(g(\Upsilon)^{\frac{1}{m}}\big)\big) \ = \mathcal{H}^k(B)$$

Then

$$\mathcal{H}^f\big(B \cap \Lambda(\Upsilon)\big) \ = \mathcal{H}^f(B) \ .$$

Remark : Conditions (i) and (ii) are not particularly restrictive. When $l = 0$, so that $\mathcal{R}$ is a collection of points in $\mathbb{R}^k$, conditions (i) and (ii) are trivially satisfied and Theorem 3 simply reduces to the Mass Transference Principle of §2.2. When $l \geq 1$, so that $\mathcal{R}$ is a collection of l–dimensional planes in $\mathbb{R}^k$, condition (i) excludes planes R_α parallel to V and condition (ii) simply means that the angle at which R_α 'hits' V is bounded away from zero by a fixed constant independent of $\alpha \in J$. This in turn implies that each plane in $\mathcal{R}$ intersects V at exactly one point.

4.3 Theorem 3 $\Longrightarrow$ Theorem 2

With reference to our general framework, let $k = m \times n$. Hence, $l = m(n-1)$. Furthermore, let $J := \{(\mathbf{a}, \mathbf{p}, \mathbf{b}) \in \mathbb{Z}^n \setminus \{\mathbf{0}\} \times \mathbb{Z}^m \times \{\mathbf{b}\} : |\mathbf{a}| = |a_1|\}$ where $\mathbf{b}$ is a fixed vector in $\mathbb{R}^m$, $\alpha := (\mathbf{a}, \mathbf{p}, \mathbf{b}) \in J$, $R_\alpha := R^{\mathbf{b}}_{\mathbf{a},\mathbf{p}}$ where the latter is given by (2) and $\Upsilon_\alpha := \frac{\Psi(\mathbf{a})}{\sqrt{\mathbf{a}.\mathbf{a}}}$. Then,

$$W^{\mathbf{b}}_{n,m}(\Psi) \ \supset \ \widetilde{W}^{\mathbf{b}}_{n,m}(\Psi) \ := \ \Lambda(\Upsilon) \cap \mathbb{I}^{n \times m} \ .$$

In view of §2.4, it suffices to establish Theorem 2 for the set $\widetilde{W}^{\mathbf{b}}_{n,m}(\Psi)$. As in §3.2, let

$$V := \{(\mathbf{x}_1, \ldots, \mathbf{x}_m) \ : \ x_{j,i} = 0 \ \ \forall \ j = 1, \ldots, m \, ; i = 2, \ldots, n, \ \} \ ,$$

where $\mathbf{x}_j = (x_{j,1}, \ldots, x_{j,n})$. Thus, V is an m–dimensional subspace of $\mathbb{R}^{n \times m}$ and it is easily verified that conditions (i) and (ii) of Theorem 3 are satisfied. Theorem 2 now follows on applying Theorem 3 with $\Omega = \mathbb{I}^{n \times m}$. Note that we can actually deduce the following stronger 'local' statement. For any ball B in $\mathbb{I}^{n \times m}$,

$$\mathcal{H}^f(B \cap W^{\mathbf{b}}_{n,m}(\Psi)) \ = \ \mathcal{H}^f(B) \qquad \text{if} \qquad \sum_{\mathbf{a} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} g\left(\frac{\Psi(\mathbf{a})}{|\mathbf{a}|}\right) \times \ |\mathbf{a}|^m \ = \ \infty \ .$$

4.4 Preliminaries

Before embarking on the proof of Theorem 3, we derive some crucial facts from conditions (i) and (ii) imposed in the statement of the theorem. We also state a 'shrinking' lemma which will be required in the proof of Theorem 3.

4.4.1 Crucial consequences of conditions (i) and (ii)

Let $l \geq 1$ as otherwise the substance of this section is irrelevant. Thus, $\mathcal{R}$ is a family of 'genuine' planes and not points. In view of the remark immediately after the statement of Theorem 3, we have that for every $\alpha \in J$ there is the unique point c_α given by $V \cap R_\alpha$. Clearly, the ball $B(c_\alpha, r)$ in $\mathbb{R}^k$ is contained in the r–neighborhood of R_α; i.e. $B(c_\alpha, r) \subset \Delta(R_\alpha, r)$. Hence,

$$B(c_\alpha, r) \cap V \ \subset \ \Delta(R_\alpha, r) \cap V \ .$$

It follows that the diameter of $\Delta(R_\alpha, r) \cap V$ is at least $2r$ – the diameter of the ball $B(c_\alpha, r) \cap V$. On the other hand, condition (ii) implies that the diameter of $\Delta(R_\alpha, r) \cap V$ is bounded above by a constant $C > 0$ times r (uniformly in α). Indeed, with C equal to the supremum in the left hand side of condition (ii) we have that

$$\Delta(R_\alpha, 1) \cap V \ \subset \ B(c_\alpha, C) \cap V \ .$$

Since R_α and V are planes, the set $\Delta(R_\alpha, r) \cap V$ is simply the set $\Delta(R_\alpha, 1) \cap V$ scaled by the factor r – shrunk or expanded depending on whether r is less than or greater than one. Similarly, $B(c_\alpha, Cr) \cap V$ is $B(c_\alpha, C) \cap V$ scaled by the factor r. The upshot of this, is that

$$B(c_\alpha, r) \cap V \ \subset \ \Delta(R_\alpha, r) \cap V \ \subset \ B(c_\alpha, Cr) \cap V \qquad \text{for any} \quad r > 0 \ . \tag{14}$$

Finally, we observe that since R_α and V are planes, the inclusions given by (14) remains valid if V is replaced by any parallel hyperplane. Formally, for any $r > 0$ and any $\mathbf{x}_0 \in \mathbb{R}^k$

$$B(c_{\alpha,\mathbf{x}_0}, r) \cap (V + \mathbf{x}_0) \ \subset \ \Delta(R_\alpha, r) \cap (V + \mathbf{x}_0) \ \subset \ B(c_{\alpha,\mathbf{x}_0}, Cr) \cap (V + \mathbf{x}_0) \ , \tag{15}$$

where $c_{\alpha,\mathbf{x}_0}$ is the unique point given by $R_\alpha \cap (V + \mathbf{x}_0)$.

4.4.2 The shrinking lemma

Given a ball B and a positive constant $\delta < 1$, let δB denote the ball B shrunk by the factor δ. The following result formally states that the measure of lim sup sets arising from a sequence of balls in $\mathbb{R}^k$ is not effected by shrinking the balls by a constant factor.

Lemma 5 *Let B_i be a sequence of balls in $\mathbb{R}^k$ such that $\limsup_{i\to\infty} B_i$ has full measure in an open subset U of $\mathbb{R}^k$. Let $\delta < 1$ be a positive constant. Then, the set $\limsup_{i\to\infty} \delta B_i$ has full measure in U.*

The lemma is a simple consequence of Lemma 6 in [3].

4.5 Proof of Theorem 3

Without loss of generality, we assume that $l \geq 1$. The case when $l = 0$ corresponds to the Mass Transference Principle of §2.2.

The proof of Theorem 3 follows the basic strategy as the proof of Theorem 2 (see §3.2); i.e. that of combining the Mass Transference Principle (Lemma 2) and the Slicing lemma (Lemma 4) in an appropriate manner. In view of this we shall give a sketch proof and leave the details to the reader.

As in the proof of Theorem 2, we can assume without loss of generality that

$$r^{-k} f(r) \;\to\; \infty \qquad \text{as} \qquad r \to 0 \ . \tag{16}$$

Indeed, it is this situation that constitutes the main substance of Theorem 3. Recall, that (16) together with Lemma 1 implies that $\mathcal{H}^f(B) = \infty$ for any k–dimensional ball B and that $\mathcal{H}^g(B) = \infty$ for any m–dimensional ball B. For the sake of clarity we introduce the following notation. Let V be as in the statement of the theorem. For a subset A of $\mathbb{R}^k$ and $\mathbf{x}_0$ in $V^\perp$ let

$$A'_{\mathbf{x}_0} \;:=\; A \cap (V + \mathbf{x}_0) \ .$$

By definition,

$$\Lambda'_{\mathbf{x}_0}(\Upsilon) = \{\mathbf{x} \in \mathbb{R}^k : \mathbf{x} \in \Delta'_{\mathbf{x}_0}(R_\alpha, \Upsilon_\alpha) \text{ for infinitely many } \alpha \in J\} \ .$$

Fix a ball D in Ω. The aim is to show that

$$\mathcal{H}^f(D \cap \Lambda(\Upsilon)) \;=\; \infty \ .$$

We are given that

$$\mathcal{H}^k\big(D \cap \Lambda\big(g(\Upsilon)^{\frac{1}{m}}\big)\big) \;=\; \mathcal{H}^k(D) \ . \tag{17}$$

Now let $D^* := \{\mathbf{x}_0 \in V^\perp : D'_{\mathbf{x}_0} \neq \emptyset\}$. Then, (17) together with Fubini's theorem implies the existence of a set $S \subset D^* \subset V^\perp$ with $|S|_l = |D^*|_l$ such that for every $\mathbf{x}_0 \in S$

$$\mathcal{H}^m\big(D'_{\mathbf{x}_0} \;\cap\; \Lambda'_{\mathbf{x}_0}\big(g(\Upsilon)^{\frac{1}{m}}\big)\big) \;=\; \mathcal{H}^m(D'_{\mathbf{x}_0}) \ . \tag{18}$$

In view of (15), we have that

$$\limsup_{\alpha \in J} \; B'_{\mathbf{x}_0}\big(c^*_\alpha, g(\Upsilon_\alpha)^{\frac{1}{m}}\big) \;\subset\; \Lambda'_{\mathbf{x}_0}\big(g(\Upsilon)^{\frac{1}{m}}\big) \;\subset\; \limsup_{\alpha \in J} \; B'_{\mathbf{x}_0}\big(c^*_\alpha, Cg(\Upsilon_\alpha)^{\frac{1}{m}}\big) \ . \tag{19}$$

For each $\alpha \in J$, the ball $B'_{\mathbf{x}_0}(c^*_\alpha, r)$ is by definition a subset of $V + \mathbf{x}_0$ with centre $c^*_\alpha := R_\alpha \cap (V + \mathbf{x}_0)$. It follows via (18) and (19), that

$$\mathcal{H}^m\big(\, D'_{\mathbf{x}_0} \cap \, \limsup\nolimits_{\alpha \in J} \, B'_{\mathbf{x}_0}\big(c^*_\alpha, Cg(\Upsilon_\alpha)^{\frac{1}{m}}\big) \,\big) \;=\; \mathcal{H}^m(D'_{\mathbf{x}_0}) \ . \tag{20}$$

As a consequence of the shrinking lemma (Lemma 5), we can put $C = 1$ in (20); i.e.

$$\mathcal{H}^m\big(\, D'_{\mathbf{x}_0} \cap \, \limsup\nolimits_{\alpha \in J} \, B'_{\mathbf{x}_0}\big(c^*_\alpha, g(\Upsilon_\alpha)^{\frac{1}{m}}\big) \,\big) \;=\; \mathcal{H}^m(D'_{\mathbf{x}_0}) \ . \tag{21}$$

Now for any ball B in $D'_{\mathbf{x}_0}$, (21) implies that

$$\mathcal{H}^m\big(\, B \cap \, \limsup_{\alpha\in J} \, B'_{\mathbf{x}_0}\big(c^*_\alpha, g(\Upsilon_\alpha)^{\frac{1}{m}}\big) \,\big) \;=\; \mathcal{H}^m(B) \ .$$

On applying the Mass Transference Principle with $\Omega = D'_{\mathbf{x}_0}$, we obtain that

$$\mathcal{H}^g\big(\, D'_{\mathbf{x}_0} \,\cap\, \limsup_{\alpha\in J} \, B'_{\mathbf{x}_0}(c^*_\alpha, \Upsilon_\alpha) \,\big) \;=\; \mathcal{H}^g(D'_{\mathbf{x}_0}) \;=\; \infty \ . \tag{22}$$

In view of (15), we have that

$$\limsup_{\alpha\in J} \, B'_{\mathbf{x}_0}(c^*_\alpha, \Upsilon_\alpha) \;\subset\; \Lambda'_{\mathbf{x}_0}(\Upsilon) \;\subset\; \limsup_{\alpha\in J} \, B'_{\mathbf{x}_0}(c^*_\alpha, C\Upsilon_\alpha) \ .$$

This together with (22), implies that for every $\mathbf{x}_0 \in S$

$$\mathcal{H}^g\big(\, D'_{\mathbf{x}_0} \,\cap\, \Lambda'_{\mathbf{x}_0}(\Upsilon) \,\big) \;=\; \infty \ .$$

On applying the Slicing lemma, we obtain that $\mathcal{H}^f(D \cap \Lambda(\Upsilon)) = \infty$ as desired.

$\square$

4.6 'Fully' non-linear Diophantine problems

Schmidt's theorem underpins the metric theory of non-linear Diophantine approximation – the integer points $\mathbf{a}$ associated with the definition of $W^{\mathbf{b}}_{n,m}(\Psi)$ can be restricted to lie in a subset $\mathcal{A}$ of $\mathbb{Z}^n$ which is completely free of any linear structure. Indeed, one simply sets Ψ to be zero for points $\mathbf{a}$ outside of $\mathcal{A}$ so that the points $\mathbf{a}$ that make any contribution to $W^{\mathbf{b}}_{n,m}(\Psi)$ lie only in $\mathcal{A}$. However, Schmidt's theorem is not non-linear in the full sense, since the integer variable $\mathbf{p}$, implicit in the symbol $\|\cdot\|$ is a linear term. Theorems 1 is therefore of the same nature; i.e. it provides a complete metric theory of non-linear Diophantine approximation but fails to be fully non-linear. Sprindžuk, in his 1979 monograph [13] writes: 'As of now, no metric theory of (fully) non-linear Diophantine approximation has been constructed. The working out of such a theory is a very topical problem.' Since then, substantial progress has been made within the one dimensional setting – the numerator and denominator of the rational approximates a/p are restricted to sets of number theoretic interest such as primes (see, for example [7, Chapter 6] for the Lebesgue measure theory and [1, §12.5] for the complete metric theory). There has also been some progress within the simultaneous setting [9]. To our knowledge, there has been no progress what so ever within the linear forms setting. We now demonstrate the power of Theorem 3 – it naturally allows us to consider fully non-linear problems; in particular within the linear forms setting.

A natural source of fully non-linear Diophantine problems is the theory of partial differential equations (PDE's). The following is a concrete example of a fully non-linear problem arising in such a manner – it is related to the solubility of the two-dimensional inhomogeneous wave equation (see [2] for details). Given a vector $\mathbf{a} = (a_1, a_2) \in \mathbb{Z}^2$, let $\mathbf{a}^2 := (a_1^2, a_2^2)$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a non-negative, real valued function and consider the set

$$S_2(\psi) := \{\mathbf{x} \in \mathbb{I}^2 : |\,\mathbf{a}^2 \cdot \mathbf{x} - p^2| < \psi(|\mathbf{a}|) \text{ for infinitely many } (\mathbf{a}, p) \in \mathbb{Z}^2 \times \mathbb{Z} \ \} \ .$$

Naturally, the problem is to determine a complete metric theory for $S_2(\psi)$. Clearly, this is a fully non-linear problem since the coefficients of the 'approximating planes' are restricted to perfect squares.

In [2], the following criteria for the 'size' of the set $S_2(\psi)$ expressed in terms of 2–dimensional Lebesgue measure $|\ \ |_2$ is established.

Theorem BDKL *Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a monotonic function such that $\lim_{h\to\infty} \psi(h) = 0$. Then*

$$|S_2(\psi)|_2 = \begin{cases} 0 & \text{if} \quad \sum_{h=1}^{\infty} \psi(h) < \infty \\ 1 & \text{if} \quad \sum_{h=1}^{\infty} \psi(h) = \infty \end{cases} .$$

In view of our general framework and Theorem 3, we are able to give a complete measure theoretic description of the set $S_2(\psi)$.

Theorem 4 *Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a monotonic function such that $\lim_{h\to\infty} \psi(h) = 0$. Let f be a dimension function such that $r^{-2}f(r)$ is monotonic. Furthermore, assume that $g : r \to r^{-1}f(r)$ is a dimension function. Then*

$$\mathcal{H}^f(S_2(\psi)) = \begin{cases} 0 & \text{if} \quad \sum\limits_{h=1}^{\infty} g\left(\dfrac{\psi(h)}{h^2}\right) \times h^2 < \infty \\ \mathcal{H}^f(\mathbb{I}^2) & \text{if} \quad \sum\limits_{h=1}^{\infty} g\left(\dfrac{\psi(h)}{h^2}\right) \times h^2 = \infty \end{cases} .$$

With $f : r \to r^s$ $(s > 0)$, the theorem reduces the the following s–dimensional Hausdorff measure statement. Naturally, it coincides with Theorem BKDL when $s = 2$.

Corollary 2 *Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a monotonic function such that $\lim_{h\to\infty} \psi(h) = 0$. For $1 < s \leq 2$, we have that*

$$\mathcal{H}^s(S_2(\psi)) = \begin{cases} 0 & \text{if} \quad \sum_{h=1}^{\infty} \psi(h)^{s-1} h^{4-2s} < \infty \\ \mathcal{H}^s(\mathbb{I}^2) & \text{if} \quad \sum_{h=1}^{\infty} \psi(h)^{s-1} h^{4-2s} = \infty \end{cases} .$$

Consider the case $\psi : r \to r^{-\tau}$ $(\tau > 0)$ and write $S_2(\tau)$ for $S_2(\psi)$. For $\tau > 1$, the above corollary not only implies that

$$\dim S_2(\tau) = \tfrac{5+\tau}{2+\tau} ,$$

but that $\mathcal{H}^s(S_2(\tau))$ is infinite at the critical exponent $s = \dim S_2(\tau)$.

4.6.1 Proof of Theorem 4

We start be rewriting the set $S_2(\psi)$ in terms of 'approximating' planes. For $\mathbf{a} \in \mathbb{Z}^2$ and $p \in \mathbb{Z}$, let

$$R_{\mathbf{a},p} := \{\mathbf{x} \in \mathbb{R}^2 : \ \mathbf{a}^2 \cdot \mathbf{x} = p^2\} \ . \tag{23}$$

It is easily verified, that

$$\mathbf{x} \in S_2(\psi) \qquad \text{if and only if} \qquad \mathbf{x} \in \Delta\Big(R_{\mathbf{a},p} \ , \tfrac{\psi(|\mathbf{a}|)}{\mathbf{a}.\mathbf{a}}\Big) \cap \mathbb{I}^2$$

for infinitely many vectors $\mathbf{a} \in \mathbb{Z}^2$ and $p \in \mathbb{Z}$. The proof of Theorem 4 follows on establishing the convergent and divergent parts separately. We make use of the fact that:

$$\sum_{h=1}^{\infty} g\left(\frac{\psi(h)}{h^2}\right) \times \ h^2 \ \asymp \sum_{\mathbf{a} \in \mathbb{Z}^2 \smallsetminus \{\mathbf{0}\}} g\left(\frac{\psi(|\mathbf{a}|)}{|\mathbf{a}|^2}\right) \times \ |\mathbf{a}| \ .$$

The case of convergence. The assumption that ψ is monotonic is irrelevant to this case. The proof follows on modifying the argument of §3.1 in the obvious manner with $n = 2$, $m = 1$ and with $\Psi(\mathbf{a})/|\mathbf{a}|$ replaced by $\psi(|\mathbf{a}|)/|\mathbf{a}|^2$.

The case of divergence. With reference to our general framework §4.1, let $k = 2$ and $m = 1$. Hence, $l = 1$. Furthermore, let $J := \{(\mathbf{a}, p) \in \mathbb{Z}^2 \setminus \{\mathbf{0}\} \times \mathbb{Z} : |\mathbf{a}| = |a_1|\}$, $\alpha := (\mathbf{a}, p) \in J$, $R_\alpha := R_{\mathbf{a},p}$ where the latter is given by (23) and $\Upsilon_\alpha := \psi(|\mathbf{a}|)/\mathbf{a}.\mathbf{a}$. Then,

$$S_2(\psi) \ \supset \ \widetilde{S}_2(\psi) \ := \ \Lambda(\Upsilon) \cap \mathbb{I}^2 \ .$$

It is easily verified that $|\widetilde{S}_2(\psi)|_2 = 1$ whenever $|S_2(\psi)|_2 = 1$. Thus, it suffices to consider the set $\widetilde{S}_2(\psi)$. Let $V := \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 \ : \ x_2 = 0\}$. Trivially, conditions (i) and (ii) of Theorem 3 are satisfied and with $\Omega = \mathbb{I}^2$ the divergence case now follows.

$\square$

4.7 Generalizing Theorem 3 to fractal subsets X of $\mathbb{R}^k$

On making use of the general Mass Transference Principle established in [3, §6.1] and adapting the Slicing lemma in an appropriate manner, it is possible to generalize Theorem 3 to the following 'fractal' setup. With k, l and m as in §4.1, let K be a compact subset of $\mathbb{R}^l$. Suppose there exists a dimension function h and constants $0 < c_1 < 1 < c_2 < \infty$ and $r_0 > 0$ such that

$$c_1 \ h(r) \ \leqslant \ \mathcal{H}^h(B(x,r)) \ \leqslant \ c_2 \ h(r) \ ,$$

for any ball $B(x,r)$ with $x \in X$ and $r \leqslant r_0$. In the case $h : r \to r^\delta$ for some $\delta > 0$, the above measure condition on balls implies that $\dim K = \delta$ and moreover that $\mathcal{H}^\delta(K)$ is strictly positive and finite. The simplest example of a fractal set K satisfying these measure theoretic properties is the standard middle third Cantor set – simple take $h : r \to r^\delta$ with

$\delta := \log 2/\log 3$. More sophisticated examples include the attractor K arising from a family of contracting self similarity maps of $\mathbb{R}^l$ satisfying the open set condition [4, 11]. Now let

$$X \;:=\; K \times \mathbb{R}^m \ .$$

Thus, X is a subset of $\mathbb{R}^k$ equipped with the product measure $\mu := \mathcal{H}^h \times |\ \ |_m$. Note that if $\dim K = \delta$, then $\dim X = \delta + m$ and furthermore if $\delta < l$, then X is a set of k–dimensional Lebesgue measure zero. Finally, let B be an arbitrary ball in X and consider the set $B \cap \Lambda(\Upsilon)$. Thus, the points of interest are restricted to X since B is by definition a subset of X. In short, it is possible to establish an analogue of Theorem 3 which enables us to transfer full measure theoretic statements with respect to the measure μ on X to general Hausdorff measure theoretic statements for $B \cap \Lambda(\Upsilon)$. The details of this and its many consequences will be the subject of a forthcoming article.

5 Appendix: Proof of Lemma 3

On taking $\phi : \mathbb{R}^k \to V^\perp$ to be the orthogonal projection map in the following statement, one easily deduces Lemma 3.

Lemma 3* *Let $l, k \in \mathbb{N}$ such that $l \leq k$ and f and $g : r \mapsto r^{-l} f(r)$ be dimension functions. Furthermore, let $A \subset \mathbb{R}^k$ and let $\phi : A \to \mathbb{R}^l$ be a Lipschitz map. Then*

$$\int^*_{\mathbb{R}^l} \mathcal{H}^g(A \cap \phi^{-1}\{y\})\, d\mathcal{L}^l y \;\leq\; \alpha(l)\, 2^l\, \mathrm{Lip}(\phi)^l\, \mathcal{H}^f(A) \ .$$

Remark. This is essentially Lemma 7.7 in [11]. For the sake of comparison, the notation adopted above is as far as possible the same as in [11]. Thus, $\int^*$ denotes the upper integral, $\mathcal{L}^l$ is the l–dimensional Lebesgue measure on $\mathbb{R}^l$, $\alpha(l) := \mathcal{L}^l\{x \in \mathbb{R}^l : |x| \leq 1\}$ is the volume of the l–dimensional unit ball and $\mathrm{Lip}(\phi)$ is the Lipschitz constant of ϕ. To avoid unnecessary confusion when comparing Lemma 3* with Lemma 7.7 in [11], it is worth pointing out that our statement contains an extra factor of 2^l since we have defined Hausdorff measure in terms of radii of balls rather than diameters. This extra factor has no effect in deducing Lemma 3 since all that we require is that the right hand side of the inequality appearing in Lemma 3* is finite whenever $\mathcal{H}^f(A)$ is finite.

5.1 Proof of Lemma 3*

The statement of Lemma 3* follows on making the obvious modifications to the proof of Lemma 7.7 in [11]. It follows from the definition of Hausdorff f–measure that for each $n \in \mathbb{N}$, there exists a cover of A by closed balls $B_{n,1}, B_{n,2}, \ldots$ such that $r(B_{n,i}) \leq 1/n$ and

$$\sum_i f(r(B_{n,i})) \;\leq\; \mathcal{H}^f_{1/n}(A) \,+\, 1/n \ . \tag{24}$$

Let, $F_{n,i} := \{y \in \mathbb{R}^l : B_{n,i} \ \cap \ \phi^{-1}\{y\} \neq \emptyset\}$. By definition, if $y, z \in F_{n,i}$ then there exist $u, v \in A \cap B_{n,i}$ such that $\phi(u) = y$ and $\psi(v) = z$. It follows that $|y - z| \leq \mathrm{Lip}(\phi)\, |u - v|$ and so

$$\mathcal{L}^l(F_{n,i}) \ \leq \ \alpha(l) \ (\mathrm{Lip}(\phi)\, 2\, r(B_{n,i})\,)^l \ . \tag{25}$$

For $y \in \mathbb{R}^l$, let $B^{\phi}_{n,i}(y)$ denote a ball of diameter $d(B_{n,i} \cap \phi^{-1}\{y\})$ such that $B_{n,i} \cap \phi^{-1}\{y\} \subseteq B^{\phi}_{n,i}(y)$. On applying Fatou's lemma and using the fact that g is non-decreasing, we obtain that

$$\begin{aligned}
\int^*_{\mathbb{R}^l} \mathcal{H}^g(A \cap \phi^{-1}\{y\})\, d\mathcal{L}^l y \quad &= \quad \int^*_{\mathbb{R}^l} \lim_{n\to\infty} \mathcal{H}^g_{1/n}(A \cap \phi^{-1}\{y\})\, d\mathcal{L}^l y \\
&\leq \quad \int_{\mathbb{R}^l} \liminf_{n\to\infty} \sum_i g(\, r(B^{\phi}_{n,i}(y))\,)\, d\mathcal{L}^l y \\
&\leq \quad \liminf_{n\to\infty} \ \sum_i \ \int_{F_{n,i}} g\Big(\, \tfrac{1}{2}\, d(B_{n,i} \cap \phi^{-1}\{y\})\Big)\, d\mathcal{L}^l y \\
&\leq \quad \liminf_{n\to\infty} \ \sum_i \ g(r(B_{n,i})\,) \ \ \mathcal{L}^l(F_{n,i}) \\
&\overset{(25)}{\leq} \quad \alpha(l) \ (2\, \mathrm{Lip}(\phi)\,)^l \ \liminf_{n\to\infty} \ \sum_i \ f(r(B_{n,i})\,) \\
&\overset{(24)}{\leq} \quad \alpha(l) \ (2\, \mathrm{Lip}(\phi)\,)^l \ \liminf_{n\to\infty} \ \Big(\mathcal{H}^f_{1/n}(A) + 1/n\Big) \\
&\leq \quad \alpha(l) \ (2\, \mathrm{Lip}(\phi)\,)^l \ \mathcal{H}^f(A) \ \ .
\end{aligned}$$

□

Acknowledgments: SV would like to thank the Ayesha and Iona for keeping him well focused on those important things in life – namely good times, mangoes and simplicity. Also, many thanks to Bridget for sharing nearly half of her years with me – poor thing!

References

[1] V. Beresnevich, H. Dickinson and S. Velani : *Measure Theoretic Laws for limsup sets*, Pre-print (97pp): arkiv:math.NT/0401118. To appear: Memoirs of the AMS.

[2] V. Beresnevich, M. Dodson, S. Kristensen and J. Levesley : *Approximation by perfect squares and the solvability of the wave equation*, In preparation.

[3] V. Beresnevich and S. Velani : *A Mass Transference Principle and the Duffin–Schaeffer conjecture for Hausdorff measures*, Pre-print (22pp): arkiv:math.NT/0401118. To appear: Annals of Math.

[4] K. Falconer : *Fractal Geometry : Mathematical Foundations and Applications.* John Wiley & Sons, (1990).

[5] P.X. Gallagher : *Metric simultaneous Diophantine approximation*, Jour. London Math. Soc. **37** (1962), 387–390.

[6] A.V. Groshev : *Un théorème sur les systèmes des formes linéaires*, Dokl. Akad. Nauk SSSR, **19** (1938), 151–152.

[7] G. Harman : *Metric Number Theory.* LMS Monographs 18, Clarendon Press, Oxford, (1998).

[8] I. Jarník : *Zur metrischen Theorie der diophantischen Appoximationen*, Proc. Mat. Fyz., 36 (1928), 91–106.

[9] H. Jones : *Khinchin's theorem in k–dimensions with prime numerator and denominator*, Acta Arith. **99** (2001), 205–225.

[10] A. Khintchine : *Zur metrischen Theorie der diophantischen Approximationen*, Math. Z. **24** (1926), 706–714.

[11] P. Mattila : Geometry of sets and measures in Euclidean space, *CUP, Cambridge studies in advance mathematics* **44** (1995)

[12] W.M. Schmidt : *Metrical theorems on fractional parts of sequences*, Trans. American Math. Soc. **110** (1964), 493–518.

[13] V.G. Sprindžuk : *Metric theory of Diophantine approximation*, John Wiley & Sons, New York-Toronto-London, 1979, (English transl.).

Victor V. Beresnevich: Department of Mathematics, University of York,
Heslington, York, YO10 5DD, England.

e-mail: vb8@york.ac.uk

Sanju L. Velani: Department of Mathematics, University of York,
Heslington, York, YO10 5DD, England.

e-mail: slv3@york.ac.uk