Regular Articles

The divergence Borel–Cantelli Lemma revisited

Victor Beresnevich *, Sanju Velani

Department of Mathematics, University of York, Heslington, York, YO10 5DD, United Kingdom

A R T I C L E   I N F O

A B S T R A C T

Let $(\Omega, \mathcal{A}, \mu)$ be a probability space. The classical Borel–Cantelli Lemma states that for any sequence of μ-measurable sets E_i $(i = 1, 2, 3, \ldots)$, if the sum of their measures converges then the corresponding lim sup set E_∞ is of measure zero. In general the converse statement is false. However, it is well known that the divergence counterpart is true under various additional 'independence' hypotheses. In this paper we revisit these hypotheses and establish both sufficient and necessary conditions for E_∞ to have either positive or full measure.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The Borel–Cantelli Lemma is a result in probability theory with wide reaching applications to various areas of mathematics. To some extent, this note is motivated by its deep applications to number theory, in particular to metric number theory – see for example [5,6,9,28,29,46] and references within. Loosely speaking, metric number theory is concerned with the arithmetic properties of almost all numbers and many key results in the theory are underpinned by variants of the divergence part of the Borel–Cantelli Lemma (see Lemma DBC below). The divergence part is also known as the second Borel–Cantelli Lemma and it naturally shows up (in some form) in the proof of the notorious Duffin-Schaeffer Conjecture [24] recently given by Koukoulopoulos & Maynard [35] and its higher dimensional generalisation proved two decades earlier by Pollington & Vaughan [41]. Indeed, the divergence Borel–Cantelli Lemma is very much at the heart of numerous other recent advances on topical problems in metric number theory, such as those in the theory of multiplicative and inhomogeneous Diophantine approximation and Diophantine approximation on manifolds and more generally on fractals, see for example [4,8,18–21,33,43,44,49]. In a nutshell, our goal it is to revisit the Borel–Cantelli Lemma and to establish both sufficient and necessary conditions that guarantee either positive or full measure.

---

* Corresponding author.
  E-mail addresses: victor.beresnevich@york.ac.uk (V. Beresnevich), sanju.velani@york.ac.uk (S. Velani).

1.1. Background and motivation

To set the scene, let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let E_i $(i \in \mathbb{N})$ be a family of measurable subsets (events) of Ω. Also, let

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

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

Determining the measure of E_∞ turns out to be one of the fundamental problems considered within the framework of classical probability theory – see for example [14, Chp. 1 §4] and [42, Chp. 47] for general background and further details. With this in mind, the following *convergence Borel–Cantelli Lemma* provides a beautiful and truly simple criterion for zero measure.

Lemma CBC (Convergence Borel–Cantelli). *Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $\{E_i\}_{i \in \mathbb{N}}$ be a sequence of subsets (events) in $\mathcal{A}$. Suppose that $\sum_{i=1}^\infty \mu(E_i) < \infty$. Then,*

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

This powerful lemma, which is also known as the *first Borel–Cantelli Lemma*, has applications in numerous disciplines. In particular, within the context of number theory it is very much at the heart of Borel's proof that almost all numbers are normal [16].

In view of Lemma CBC, it is natural to ask whether or not there is a sufficient condition that enables us to deduce that the measure of E_∞ is positive or possibly even full; that is to say that

$$\mu(E_\infty) = \mu(\Omega) = 1 \,.$$

The divergence of the measure sum $\sum_{i=1}^\infty \mu(E_i)$ is clearly necessary but certainly not enough as the following simple example demonstrates.

Example. For $i \in \mathbb{N}$, let $E_i = (0, \frac{1}{i}) \subset \Omega := [0, 1]$ and μ be one-dimensional Lebesgue measure restricted to $[0, 1]$. Then

$$\sum_{i=1}^\infty \mu(E_i) = \sum_{i=1}^\infty i^{-1} = \infty$$

but

$$E_\infty = \bigcap_{t=1}^\infty \bigcup_{i=t}^\infty E_i = \bigcap_{t=1}^\infty (0, \tfrac{1}{t}) = \varnothing \qquad \text{and so} \qquad \mu(E_\infty) = 0 \,.$$

The problem in the above example is that the building blocks E_i of the lim sup set under consideration 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_i are in 'some sense' independent; that is, we need to control overlaps! Indeed, Borel & Cantelli showed that *mutual independence* in the classical probabilistic sense, which means that for every $n \in \mathbb{N}$

$$\mu\left(\bigcap_{t=1}^n E_{i_t}\right) = \prod_{t=1}^n \mu(E_{i_t}) \quad \text{for any indices } i_1 < \ldots < i_n \,, \tag{1}$$

implies that $\mu(E_\infty) = 1$. This full measure statement, often referred to as the *second Borel–Cantelli Lemma*, led to a flurry of activity with the aim of relaxing the mutual independence condition. Notable progress in this quest included replacing mutual independence by *pairwise independence* – this corresponds to (1) being fulfilled with $n = 2$ rather than every $n \in \mathbb{N}$. In turn, pairwise independence was replaced by the upper bound condition

$$\mu(E_s \cap E_t) \leq \mu(E_s)\mu(E_t) \qquad s \neq t \tag{2}$$

on the overlaps. Undoubtedly, verifying (2) is significantly easier than (1). However, in many applications, we rarely have (2) let alone mutual independence as in the original statement of the second Borel–Cantelli Lemma. What is much more useful is the following variant which these days is often referred to as the *divergence Borel–Cantelli Lemma.*

Lemma DBC (Divergence Borel–Cantelli). *Let $(\Omega, \mathcal{A}, \mu)$ be a probability space and let $\{E_i\}_{i \in \mathbb{N}}$ be a sequence of subsets (events) in $\mathcal{A}$. Suppose that $\sum_{i=1}^{\infty} \mu(E_i) = \infty$ and that there exists a constant $C > 0$ such that*

$$\sum_{s,t=1}^{Q} \mu(E_s \cap E_t) \leq C \left(\sum_{s=1}^{Q} \mu(E_s) \right)^2 \tag{3}$$

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

$$\mu(E_\infty) \geq C^{-1}\,.$$

In particular, if $C = 1$ then

$$\mu(E_\infty) = 1\,.$$

We refer the reader to [28,29,42,46] for the proof of the lemma which is essentially a consequence of the Cauchy-Schwarz inequality. As pointed out by Harman [29], the basic idea goes back to the works of Payley & Zygmund [38,39] from the nineteen thirties.

Remark 1. To the best of our knowledge, the in particular part of Lemma DBC first explicitly appears in the work of Erdös & Reyni [25, Lemma C] from the late fifties. Lamperti [36] in the early sixties established the weaker form of above lemma in which (3) is replaced by

$$\mu(E_s \cap E_t) \leq C\mu(E_s)\mu(E_t) \qquad s \neq t\,. \tag{4}$$

Clearly, this pairwise condition implies the average condition (3). Independently and around the same time, Kochen & Stone [34] established Lemma DBC as stated – see also [45, Lemma 2, p. 165]. It is worth mentioning, that Chung & Erdös [22] in the early fifties explored the implications of imposing condition (4) on the overlaps. Within the specific number theoretic setting, Duffin & Schaeffer [24] had carried out such an investigation in 1941 and it was a key ingredient in their proof of what today is refereed to as the Duffin-Schaeffer Theorem [24, Theorem I]. This theorem is a special case of the notorious Duffin-Schaeffer Conjecture mentioned right at the start of this paper.

Remark 2. Condition (3) is often refereed to as *quasi-independence on average* and together with the divergence of the measure sum guarantees that the associated lim sup set E_∞ is of positive measure. It does not in general guarantee full measure. However, this is not an issue if we already know by some other means

(such as Kolmogorov's theorem [14, Theorems 4.5 & 22.3] or ergodicity [14, §24]) that the lim sup set E_∞ satisfies a *zero-one law*; namely that

$$\mu(E_\infty) = 0 \quad \text{or} \quad 1.$$

Within the context of metric number theory, the existence of such a law for the lim sup set of well approximable real numbers is due to Cassels [17] and Gallagher [27] and it plays a key role in the recent proof of the Duffin-Schaeffer Conjecture [35]. For further details and higher dimensional generalisation of their zero-one laws see [7,10] and references within. Alternatively, without the presence of a general zero-one law, if we are willing to impose a little more structure on the probability space, we can guarantee full measure if the measure sum diverges locally and quasi-independence on average holds locally in the presence of an appropriate topological structure on Ω. In short, by locally we mean that the conditions under consideration hold for $E_i \cap A$ where A is an arbitrary open set with positive measure. For the precise statement see Lemma LBC below.

In short, the purpose of the present paper is to determine whether or not Lemma DBC is best possible. In other words, is it the case that the *pairwise quasi-independence on average* condition (3) cannot be replaced by a weaker condition? Recall, that in view of Lemma CBC, the divergence sum condition within Lemma DBC is not negotiable – it has to be present. We show that within a reasonably general framework, given any lim sup set $E_\infty := \limsup_{i \to \infty} E_i$ with $\mu(E_\infty) > 0$ the sets E_i can be appropriately manipulated or rather "trimmed" in such a manner that the resulting subsets E_i^* are quasi-independent on average and the sum of the measures $\mu(E_i^*)$ diverges. Thus, up to "trimming" the divergence Borel–Cantelli Lemma is best possible. Moreover, we show that quasi-independence on average for the trimmed sets is in fact not only equivalent to full measure but to three other useful properties which are of independent interest especially within the context of applications. We conclude the paper with a couple of examples that demonstrate the versatility and power of our results.

1.2. Statement of results

Throughout, $(\Omega, \mathcal{A}, \mu, d)$ will be a metric measure space equipped with a Borel probability measure μ. In what follows, $\text{supp}\,\mu$ will denote the support of the measure μ and given $x \in \Omega$ and $r > 0$, $B = B(x, r)$ will denote the ball centred at x of radius r. Also, given a real number $a > 0$, we denote by aB the ball B scaled by a factor a; i.e. $aB := B(x, ar)$. Most of the time we will assume that μ is doubling. Recall, that μ is said to be *doubling* if there are constants $\lambda \geq 1$ and $r_0 > 0$ such that for any $x \in \text{supp}\,\mu$ and $0 < r < r_0$

$$\mu(B(x, 2r)) \leq \lambda\,\mu(B(x, r)). \tag{5}$$

The doubling condition allows us to blow up a given ball by a constant factor without drastically affecting its measure. The metric measure space $(\Omega, \mathcal{A}, \mu, d)$ is also said to be *doubling* if μ is doubling [30]. Note that the doubling property is imposed only on the measure of balls centred in $\text{supp}\,\mu$. However, in many instances the doubling property can be effectively used on balls that are not necessarily centred in $\text{supp}\,\mu$ provided that they contain 'enough' of the support. In particular, when working with a given sequence of balls $\{B_i\}$ in Ω we will often impose the following weaker version of doubling:

$$\exists\, a, b > 1 \quad \text{such that} \quad \mu(aB_i) \leq b\mu(B_i) \qquad \text{for all } i \text{ sufficiently large.} \tag{6}$$

Condition (6) is not particularly restrictive and ensures that whenever $\mu(B_i) > 0$ the support of μ within B_i is not concentrated too close to the boundary of B_i. Indeed, if for some $\varepsilon > 0$ the ball $(1-\varepsilon)B_i$ contains

points in supp μ, then the inequality in (6) holds with any $a > 1$ and $b = \lambda^k$ where $k := \left\lceil \log_2 \frac{a+1-\varepsilon}{\varepsilon} \right\rceil$. Note that if the centre of B_i is in supp μ then the inequality in (6) trivially holds with $a = 2$ and $b = \lambda$.

Restricting our attention to lim sup sets of balls, we have the following 'if and only if' statement for full measure.

Theorem 1. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a doubling Borel probability measure μ. Let $\{B_i\}_{i \in \mathbb{N}}$ be a sequence of balls in Ω with $r(B_i) \to 0$ as $i \to \infty$ and such that (6) holds. Let $E_\infty := \limsup_{i \to \infty} B_i$. Then*

$$\mu(E_\infty) = 1$$

if and only if there exists a constant $C > 0$ such that for any ball B centred in supp μ there is a sub-sequence $\{L_{i,B}\}_{i \in \mathbb{N}}$ of $\{B_i\}_{i \in \mathbb{N}}$ of balls contained in B (i.e. $L_{i,B} \subset B$ for all i), such that

$$\sum_{i=1}^{\infty} \mu(L_{i,B}) = \infty \tag{7}$$

and for infinitely many $Q \in \mathbb{N}$

$$\sum_{s,t=1}^{Q} \mu\big(L_{s,B} \cap L_{t,B}\big) \leq \frac{C}{\mu(B)} \left(\sum_{s=1}^{Q} \mu(L_{s,B}) \right)^2 . \tag{8}$$

It is important to note that the constant $C > 0$ appearing in (8) is independent of the arbitrary ball B. The following is a strengthening of Theorem 1 to lim sup sets of open sets. As we shall see, the proof will follow the same line of argument as that of Theorem 1.

Theorem 2. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a doubling Borel probability measure μ. Let $\{E_i\}_{i \in \mathbb{N}}$ be a sequence of open subsets (events) in Ω and let $E_\infty := \limsup_{i \to \infty} E_i$. Then*

$$\mu(E_\infty) = 1$$

if and only if there exists a constant $C > 0$ such that for any ball B centred in supp μ there is a sequence $\{L_{i,B}\}_{i \in \mathbb{N}}$ of finite unions of disjoint balls centred in supp μ with $L_{i,B} \subset E_i \cap B$ satisfying (7) and (8) for infinitely many $Q \in \mathbb{N}$.

The upshot is that for a lim sup set E_∞ to have full measure we must be able to locally "trim" the associated sets E_i so that the resulting trimmed subsets are quasi-independent on average and the sum of their measures diverges.

It turns outs that the sufficiency part of the Theorem 2 can be made more general. In particular, the doubling condition can be altogether dropped. The following is a local variant of the (standard) divergence Borel–Cantelli Lemma which allows us to deduce full measure rather than just positive measure.

Lemma LBC (Local Borel–Cantelli). *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a Borel probability measure μ and let $\{E_i\}_{i \in \mathbb{N}}$ be a sequence of Borel subsets of Ω. Suppose there exists an increasing function $f : (0, +\infty) \to (0, +\infty)$ with $f(x) \to 0$ as $x \to 0$ such that for any open set A with $\mu(A) > 0$ there is a sequence $\{L_{i,A}\}_{i \in \mathbb{N}}$ of measurable subsets of A such that*

$$\sum_{i=1}^{\infty} \mu(L_{i,A}) = \infty \tag{9}$$

and for infinitely many $Q \in \mathbb{N}$

$$\sum_{s,t=1}^{Q} \mu\big(L_{s,A} \cap L_{t,A}\big) \leq \frac{1}{f(\mu(A))} \left(\sum_{s=1}^{Q} \mu(L_{s,A})\right)^2 . \qquad (10)$$

Then

$$\mu(E_\infty) = 1 .$$

Moreover, if in addition μ is doubling and $f(x) = cx$ for some constant $0 < c \leq 1$, it suffices to take A in the above to be an arbitrary ball of sufficiently small radius centred in $\operatorname{supp}\mu$.

Remark 3. In the case $f(x) = cx$ for some constant $0 < c \leq 1$ and $A = B$ is a ball, condition (10) becomes the same as (8) with $C = c^{-1}$. Given a measurable set A with $\mu(A) > 0$, let μ_A denote the conditional probability measure given by

$$\mu_A(E) := \frac{\mu(E \cap A)}{\mu(A)} \qquad \text{for } E \in \mathcal{A} .$$

In other words, μ_A is the re-normalised μ-measure restricted to A. Then it is easily seen that on replacing μ by μ_A, the divergence condition (9) and the overlap condition (10) with $f(x) = cx$ coincide with those of Lemma DBC. For obvious reasons, the independence condition (10) with $f(x) = cx$ is often refereed to as *local quasi-independence on average.*

Theorem 1 will follow from a more general statement that provides three more necessary and sufficient conditions for full measure. To be more precise, Theorem 1 is the equivalence between (A) and (E) within the following statement with $C = \kappa^{-2}$.

Throughout, we use the standard notation $\overset{\circ}{\bigcup}$ to denote that the union of sets under consideration is disjoint.

Proposition 1. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a doubling Borel probability measure μ. Let $\{B_i\}_{i \in \mathbb{N}}$ be a sequence of balls in Ω with $r(B_i) \to 0$ as $i \to \infty$ and such that (6) holds. Let $E_\infty := \limsup_{i \to \infty} B_i$. Then, the following statements are equivalent:*

(A) $\mu(E_\infty) = 1$.
(B) *For any ball B in Ω, we have that*

$$\mu(E_\infty \cap B) = \mu(B) . \qquad (11)$$

(C) *For any ball B in Ω centred in $\operatorname{supp}\mu$ and any $G \in \mathbb{N}$, there is a finite sub-collection $\mathcal{K}_{G,B} \subset \{B_i : i \geq G\}$ of disjoint balls contained in B such that*

$$\mu\Big(\overset{\circ}{\underset{L \in \mathcal{K}_{G,B}}{\bigcup}} L \Big) \geq \kappa \, \mu(B) \quad \text{where} \quad \kappa := \frac{1}{2} \frac{1}{\lambda^{k+1} b} \qquad (12)$$

where λ is as in (5), $k := \max\{1, \lceil \log_2 \frac{6}{a-1} \rceil\}$ and a, b are as in (6).
(D) *For any ball B in Ω centred in $\operatorname{supp}\mu$ and any $G \in \mathbb{N}$, there is a subset $E_{G,B} \subseteq B$ consisting of a finite union of disjoint balls from $\{B_i : i \geq G\}$, such that*

$$\sum_{i=1}^{\infty} \mu(E_{G_i,B}) = \infty$$

for any subsequence $(G_i)_{i\in\mathbb{N}}$ of natural numbers, and, with κ is as in (12), for any pair of natural numbers G and G'

$$\mu\big(E_{G,B} \cap E_{G',B}\big) \leq \frac{1}{\mu(B)\,\kappa^2}\,\mu(E_{G,B})\,\mu(E_{G',B}). \tag{13}$$

(E) *For any ball B in Ω centred in $\operatorname{supp}\mu$ there is a sub-sequence $\{L_{i,B}\}_{i\in\mathbb{N}}$ of $\{B_i\}_{i\in\mathbb{N}}$ of balls contained in B such that*

$$\sum_{i=1}^{\infty} \mu(L_{i,B}) = \infty$$

and, with κ is as in (12), for infinitely many $Q \in \mathbb{N}$

$$\sum_{s,t=1}^{Q} \mu\big(L_{s,B} \cap L_{t,B}\big) \leq \frac{1}{\mu(B)\,\kappa^2} \left(\sum_{s=1}^{Q} \mu(L_{s,B})\right)^2. \tag{14}$$

Remark 4. It will become apparent in the proof that we can take the subset $E_{G,B}$ in (D) to be the union of balls in the sub-collection $\mathcal{K}_{G,B}$ associated with (C).

We now turn our attention to an 'if and only if' statement for positive measure for $\limsup$ sets of balls.

Theorem 3. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a doubling Borel probability measure μ. Let $\{B_i\}_{i\in\mathbb{N}}$ be a sequence of balls in Ω such that (6) holds. Let $E_\infty := \limsup_{i\to\infty} B_i$. Then*

$$\mu(E_\infty) > 0$$

if and only if there exists a sub-sequence $\{L_i\}_{i\in\mathbb{N}}$ of $\{B_i\}_{i\in\mathbb{N}}$ and a constant $C > 0$ such that

$$\sum_{i=1}^{\infty} \mu(L_i) = \infty \tag{15}$$

and for infinitely many $Q \in \mathbb{N}$

$$\sum_{s,t=1}^{Q} \mu\big(L_s \cap L_t\big) \leq C \left(\sum_{s=1}^{Q} \mu(L_s)\right)^2. \tag{16}$$

The following is the 'positive measure' analogue of Proposition 1 and it clearly implies Theorem 3.

Proposition 2. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a doubling Borel probability measure μ. Let $\{B_i\}_{i\in\mathbb{N}}$ be a sequence of balls in Ω such that (6) holds. Let $E_\infty := \limsup_{i\to\infty} B_i$. Then, the following statements are equivalent:*

(A) $\mu(E_\infty) > 0$.

(B) *For any $G \in \mathbb{N}$, there is a finite sub-collection $\mathcal{K}_G \subset \{B_i : i \geq G\}$ of disjoint balls such that*

$$\mu\left(\bigcup_{L \in \mathcal{K}_G}^{\circ} L\right) \geq \kappa \quad where \quad \kappa := \frac{1}{2\lambda^{k+1}b}\, \mu(E_\infty), \tag{17}$$

where λ is as in (5), *$k := \max\{1, \left\lceil \log_2 \frac{6}{a-1} \right\rceil\}$ and a, b are as in* (6).

(C) *For any $G \in \mathbb{N}$, there is a subset E_G of Ω, which is a finite union of disjoint balls from $\{B_i : i \geq G\}$, such that*

$$\sum_{i=1}^{\infty} \mu(E_{G_i}) = \infty$$

for any subsequence $(G_i)_{i \in \mathbb{N}}$ of natural numbers, and, with κ is as in (17), *for any pair of natural numbers G and G'*

$$\mu(E_G \cap E_{G'}) \leq \kappa^{-2}\, \mu(E_G)\, \mu(E_{G'}). \tag{18}$$

(D) *There is a sub-sequence $\{L_i\}_{i \in \mathbb{N}}$ of $\{B_i\}_{i \in \mathbb{N}}$ such that*

$$\sum_{i=1}^{\infty} \mu(L_i) = \infty$$

and, with κ is as in (17), *for infinitely many $Q \in \mathbb{N}$*

$$\sum_{s,t=1}^{Q} \mu(L_s \cap L_t) \leq \kappa^{-2} \left(\sum_{s=1}^{Q} \mu(L_s)\right)^2. \tag{19}$$

2. Proof of results

2.1. Preliminaries

We will make multiple use of the following basic covering lemma, see for example [30,37].

Lemma 1 *(The 5r covering lemma). Every family $\mathcal{F}$ of balls of uniformly bounded diameter in a metric space (Ω, d) contains a disjoint subfamily $\mathcal{G}$ such that*

$$\bigcup_{B \in \mathcal{F}} B \subset \bigcup_{B \in \mathcal{G}} 5B.$$

The following measure theoretic result, which is an extension of Proposition 1 in [6, §8], provides a mechanism for establishing full measure statements.

Lemma 2. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a Borel doubling probability measure μ. Let E be a Borel subset of Ω. Assume that there are constants $r_0, c > 0$ such that for any ball B centred in $\mathrm{supp}\,\mu$ with $r(B) < r_0$, we have that*

$$\mu(E \cap B) \geq c\, \mu(B). \tag{20}$$

Then, $\mu(E) = 1$.

The lemma is a standard corollary of the Lebesgue density theorem or more generally the Lebesgue differentiation theorem for doubling metric measure spaces (see for example [30, Theorem 1.8]. A slightly weaker version of this lemma can also be found in [6]. In short, the version of this lemma established as Proposition 1 in [6, §8] requires that (20) holds for arbitrary balls centred in Ω rather than just in $\operatorname{supp}\mu$ and the proof uses covering arguments rather than the Lebesgue density theorem.

Remark 5. Note that the doubling assumption in Lemma 2 can be weakened by requiring instead that $(\Omega, \mathcal{A}, \mu, d)$ is a Vitali space as defined in [30, p. 6]. Furthermore it is also possible to remove the doubling assumption altogether from Lemma 2 at the price of requiring a lower bound on $\mu(E \cap B)$ for an arbitrary open set B as opposed to an arbitrary ball of sufficiently small radius. We will state this version formally as it will be required in the proof of Lemma LBC.

Lemma 3 *(Lemma 6 in [6]). Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a Borel probability measure μ. Let E be a Borel subset of Ω and $f : (0, +\infty) \to (0, +\infty)$ be an increasing function such that $f(x) \to 0$ as $x \to 0$. Assume that*

$$\mu(E \cap A) \geq f(\mu(A))$$

for any open subset $A \subset \Omega$ with $\mu(A) > 0$. Then, $\mu(E) = 1$.

The following "obvious" but useful statement relates the standard doubling property (5) for balls centred in $\operatorname{supp}\mu$ with the weaker property corresponding to (6) in which the centre can be anywhere.

Lemma 4. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a metric measure space equipped with a Borel doubling probability measure μ. Let $a, b > 0$ be constants and let B be a ball in Ω such that $\mu(aB) \leq b\mu(B)$ and $B \cap \operatorname{supp}\mu \neq \varnothing$. Then for any $s \geq a$ we have that*

$$\mu(sB) \leq \lambda^k b\mu(B),$$

where $k \in \mathbb{N}$ satisfies $2^k \geq (1 + s)/(a - 1)$ and λ is as in (5).

Proof. Let $x \in B \cap \operatorname{supp}\mu$ and B' be the ball centred at x of radius $(a - 1)r(B)$. Then clearly $B' \subset aB$. Since B' is centred in the support of μ, the doubling property (5) is applicable to it and we have that for every $k \in \mathbb{N}$

$$\mu(2^k B') \leq \lambda^k \mu(B').$$

Since $B' \subset aB$ and $\mu(aB) \leq b\mu(B)$, we therefore have that

$$\mu(2^k B') \leq \lambda^k \mu(aB) \leq \lambda^k b\mu(B). \tag{21}$$

It remains to observe that $sB \subset 2^k B'$ provided that $(a - 1)2^k \geq 1 + s$ and so (21) implies the required inequality.　□

2.2. Proof of Lemma LBC

Let A be any open subset of Ω and $\{L_{i,A}\}_{i \in \mathbb{N}}$ be the sequence of sets as in Lemma LBC. In particular, by definition, $L_{i,A} \subset E_i \cap A$ for every $i \in \mathbb{N}$ and therefore

$$\limsup_{i \to \infty} L_{i,A} \subseteq A \cap \limsup_{i \to \infty} E_i.$$

On applying Lemma DBC (the standard divergent Borel–Cantelli Lemma) it follows that

$$\mu\Big(A \cap \limsup_{i\to\infty} E_i\Big) \geq \mu\Big(\limsup_{i\to\infty} L_{i,A}\Big) \geq f(\mu(A)).$$

Also recall that E_i is a Borel set for every $i \in \mathbb{N}$ and therefore $E_\infty := \limsup_{i\to\infty} E_i$ is a Borel subset of Ω. Then, applying Lemma 3 with $E = E_\infty$ implies that $\mu(E_\infty) = 1$ as desired.

If μ is a doubling measure and $f(x) = cx$ for some constant $0 < c \leq 1$, then the 'moreover' part of Lemma LBC follows on applying Lemma 2 instead of Lemma 3.

2.3. Proof of Proposition 1

• Step 1: (A) $\Longrightarrow$ (B). This is obvious since μ is a probability measure.

• Step 2: (B) $\Longrightarrow$ (C). Let B be any ball in Ω centred in $\operatorname{supp}\mu$. In particular, we have that $\mu(B) > 0$. Let $\mathcal{F} := \{B_i \ : \ B_i \cap \frac{1}{2}B \cap \operatorname{supp}\mu \neq \emptyset, \ i \geq G\}$. Since, $r(B_i) \to 0$ as $i \to \infty$, we can ensure that every ball in $\mathcal{F}$ is contained in B by increasing G if necessary. In view of the $5r$ covering lemma (Lemma 1), there exists a disjoint sub-family $\mathcal{G}$ such that

$$\bigcup_{B_i\in\mathcal{F}} B_i \subset \bigcup_{B_i\in\mathcal{G}} 5B_i.$$

It follows that

$$\tfrac{1}{2}B \cap \limsup_{i\to\infty} B_i \cap \operatorname{supp}\mu \subset \bigcup_{B_i\in\mathcal{F}} B_i \cap \operatorname{supp}\mu \subset \bigcup_{B_i\in\mathcal{G}} 5B_i \cap \operatorname{supp}\mu.$$

Hence,

$$\mu\left(\bigcup_{B_i\in\mathcal{G}} 5B_i\right) \geq \mu\big(\tfrac{1}{2}B \cap \limsup_{i\to\infty} B_i\big) \overset{(11)}{=} \mu(\tfrac{1}{2}B) \overset{(5)}{\geq} \frac{1}{\lambda}\,\mu(B).$$

However, since $\mathcal{G}$ is a disjoint collection of balls, which have non-empty intersection with $\operatorname{supp}\mu$, we have that

$$\mu\left(\bigcup_{B_i\in\mathcal{G}} 5B_i\right) \leq \sum_{B_i\in\mathcal{G}} \mu\left(5B_i\right) \overset{(6)\ \&\ \text{Lemma 4}}{\leq} \lambda^k b \sum_{B_i\in\mathcal{G}} \mu\left(B_i\right) = \lambda^k b\, \mu\left(\overset{\circ}{\bigcup_{B_i\in\mathcal{G}}} B_i\right),$$

where $k := \max\{1, \big\lceil \log_2 \frac{6}{a-1} \big\rceil\}$. Thus,

$$\frac{1}{\lambda^{k+1}b}\,\mu(B) \leq \mu\left(\overset{\circ}{\bigcup_{B_i\in\mathcal{G}}} B_i\right) = \sum_{B_i\in\mathcal{G}} \mu\left(B_i\right) \leq \mu(B). \tag{22}$$

If $\mathcal{G}$ is infinite, the sum in (22) is convergent and therefore there exists some $j_0 > G$ for which

$$\sum_{B_i\in\mathcal{G}\,:\,i\geq j_0} \mu\left(B_i\right) = \mu\left(\overset{\circ}{\bigcup_{B_i\in\mathcal{G}\,:\,i\geq j_0}} B_i\right) < \frac{1}{2}\,\frac{1}{\lambda^{k+1}b}\,\mu(B). \tag{23}$$

Obviously, this is also true if $\mathcal{G}$ is finite. Now let $\mathcal{K}_{G,B} := \{B_i : B_i \in \mathcal{G}, i < j_0\}$. Clearly, this is a finite sub-collection of $\{B_i : i \geq G\}$. Moreover, in view of (22) and (23) the collection $\mathcal{K}_{G,B}$ satisfies the desired properties.

• Step 3: (C) $\implies$ (D). For any ball B centred in supp μ and any $G \in \mathbb{N}$, let $\mathcal{K}_{G,B}$ be the finite sub-collection of disjoint balls associated with (C) and define

$$E_{G,B} := \bigcup_{L \in \mathcal{K}_{G,B}}^{\circ} L \subseteq B. \tag{24}$$

It follows from (12) that

$$\mu(E_{G,B}) \geq \kappa \, \mu(B),$$

which in turn implies that $\sum_{i=1}^{\infty} \mu(E_{G_i,B}) \geq \sum_{i=1}^{\infty} \kappa \, \mu(B) = \infty$ for any subsequence $(G_i)_{i \in \mathbb{N}}$ of natural numbers, and

$$\mu\big(E_{G,B} \cap E_{G',B}\big) \leq \mu(B) \leq \frac{1}{\mu(B) \, \kappa^2} \, \mu(E_{G,B}) \, \mu(E_{G',B})$$

for any pair of natural numbers G and G'. Thus, the sets $E_{G,B}$ satisfy the desired properties.

• Step 4: (D) $\implies$ (E). Let B be any ball centred in supp μ and for any $G \in \mathbb{N}$ let $E_{G,B} \subset B$ be as in (D), and let $\mathcal{K}_{G,B}$ be a finite collection of disjoint balls from $\{B_i : i \geq G\}$ that constitute $E_{G,B}$, that is (24) holds. Observe that for any pair of natural numbers G and G'

$$
\begin{aligned}
\mu\big(E_{G,B} \cap E_{G',B}\big) &= \mu\left(\Big(\bigcup_{L \in \mathcal{K}_{G,B}}^{\circ} L \Big) \cap \Big(\bigcup_{L' \in \mathcal{K}_{G',B}}^{\circ} L' \Big) \right) \\
&= \sum_{L \in \mathcal{K}_{G,B}} \sum_{L' \in \mathcal{K}_{G',B}} \mu\big(L \cap L' \big) \\
&\overset{(13)}{\leq} \frac{1}{\mu(B) \, \kappa^2} \, \mu(E_{G,B}) \, \mu(E_{G',B}) \\
&= \frac{1}{\mu(B) \, \kappa^2} \sum_{L \in \mathcal{K}_{G,B}} \mu(L) \sum_{L' \in \mathcal{K}_{G',B}} \mu(L'). \tag{25}
\end{aligned}
$$

Let $G_1 = 1$ and fix the collection $\mathcal{K}_{G_1,B}$. Define $G_2 = t+1$ where t is the largest index such that $B_t \in \mathcal{K}_{G_1,B}$. Since $\mathcal{K}_{G_1,B}$ is finite this is clearly possible. With G_2 defined, we can fix the collection $\mathcal{K}_{G_2,B}$ and proceed by induction as follows. Suppose the integers $G_1, \ldots, G_n$ and the corresponding collections $\mathcal{K}_{G_1,B}, \ldots, \mathcal{K}_{G_n,B}$ have been determined. Define $G_{n+1} = t+1$ where t is the largest index such that $B_t \in \mathcal{K}_{G_n,B}$. With G_{n+1} defined, we can fix the collection $\mathcal{K}_{G_{n+1},B}$ and we are done. Now, let $\{L_{s,B}\}_{s \in \mathbb{N}}$ be the sequence of balls contained in B obtained by placing the balls from $\{\mathcal{K}_{G_i,B} : i \in \mathbb{N}\}$ in the same order as in $\{B_i\}_{i \in \mathbb{N}}$. In view of the choice of the integers G_i, the sequence $\{L_{s,B}\}_{s \in \mathbb{N}}$ is a well defined sub-sequence of $\{B_i\}_{i \in \mathbb{N}}$. For $M \in \mathbb{N}$, let

$$Q_M := \sum_{i=1}^{M} \#\mathcal{K}_{G_i,B}.$$

It then follows that for any $M \geq 2$

$$\sum_{s,t=1}^{Q_M} \mu\big(L_{s,B} \cap L_{t,B}\big) = \sum_{i,j=1}^{M} \mu\big(E_{G_i,B} \cap E_{G_j,B}\big)$$

$$= \sum_{i,j=1}^{M} \sum_{L \in \mathcal{K}_{G_i,B}} \sum_{L' \in \mathcal{K}_{G_j,B}} \mu\big(L \cap L'\big)$$

$$\overset{(25)}{\leq} \frac{1}{\mu(B)\,\kappa^2} \sum_{i,j=1}^{M} \sum_{L \in \mathcal{K}_{G_i,B}} \mu(L) \sum_{L' \in \mathcal{K}_{G_j,B}} \mu(L')$$

$$= \frac{1}{\mu(B)\,\kappa^2} \Big(\sum_{i=1}^{M} \sum_{L \in \mathcal{K}_{G_i,B}} \mu(L) \Big)^2$$

$$= \frac{1}{\mu(B)\,\kappa^2} \Big(\sum_{s=1}^{Q_M} \mu(L_{s,B}) \Big)^2 .$$

This together with the fact that

$$\sum_{s=1}^{\infty} \mu(L_{s,B}) = \sum_{i=1}^{\infty} \mu(E_{G_i,B}) = \infty \,,$$

shows that the sequence $\{L_{s,B}\}_{s \in \mathbb{N}}$ satisfies the desired properties.

• Step 5: (E) $\implies$ (A). This follows immediately on applying the "moreover" part of Lemma LBC with $c = \kappa^2$.

2.4. Proof of Proposition 2

The proof is very similar to that of Proposition 1 and so we will simply provide a sketch.

• Step 1: (A) $\implies$ (B). For $G \in \mathbb{N}$, let $\mathcal{F} := \{B_i \cap \operatorname{supp} \mu \neq \varnothing \,:\, i \geq G\}$. In view of the $5r$ covering lemma (Lemma 1), there exists a disjoint sub-family $\mathcal{G}$ of $\mathcal{F}$ such that

$$\bigcup_{B_i \in \mathcal{F}} B_i \subset \bigcup_{B_i \in \mathcal{G}} 5B_i \,.$$

It follows that

$$\mu\left(\bigcup_{B_i \in \mathcal{G}} 5B_i \right) \geq \mu\big(E_\infty\big)$$

and the same argument leading to (22) shows that

$$\mu\left(\overset{\circ}{\bigcup_{B_i \in \mathcal{G}}} B_i \right) \geq \frac{1}{\lambda^{k+1}b} \,\mu\big(E_\infty\big) \,, \tag{26}$$

where k is the same integer as in (22). Furthermore, the same argument leading to (23) shows that there exists some $j_0 > G$ for which

$$\mu\left(\overset{\circ}{\bigcup_{B_i \in \mathcal{G}\,:\,i \geq j_0}} B_i \right) < \frac{1}{2} \frac{1}{\lambda^{k+1}b} \,\mu\big(E_\infty\big) := \kappa. \tag{27}$$

Then, in view of (26) and (27) the finite sub-collection $\mathcal{K}_G := \{B_i : B_i \in \mathcal{G}, i < j_0\}$ of $\{B_i : i \geq G\}$ satisfies the desired properties.

• Step 2: (B) $\Longrightarrow$ (C). For any $G \in \mathbb{N}$, let $\mathcal{K}_G$ be the finite sub-collection of disjoint balls associated with (B) and define

$$E_G := \bigcup_{L \in \mathcal{K}_G}^{\circ} L\,. \tag{28}$$

It follows from (17) that $\mu(E_G) \geq \kappa$ which in turn implies that $\sum_{i=1}^{\infty} \mu(E_{G_i}) = \infty$ for any subsequence $(G_i)_{i \in \mathbb{N}}$ of natural numbers, and that

$$\mu\big(E_G \cap E_{G'}\big) \leq 1 \leq \kappa^{-2}\, \mu(E_G)\,\mu(E_{G'})$$

for any pair of natural numbers G and G'. Thus, the sets E_G satisfy the desired properties.

• Step 3: (C) $\Longrightarrow$ (D). For any $G \in \mathbb{N}$ let $E_G \subset \Omega$ be as in (C), and let $\mathcal{K}_G$ be a finite collection of disjoint balls from $\{B_i : i \geq G\}$ that constitute E_G, that is (28) holds. Observe that the same argument leading to (25) shows that for any pair of natural numbers G and G'

$$\mu\big(E_G \cap E_{G'}\big) = \sum_{L \in \mathcal{K}_G} \sum_{L' \in \mathcal{K}_{G'}} \mu\Big(L \cap L'\Big)$$

$$\overset{(18)}{\leq} \kappa^{-2} \sum_{L \in \mathcal{K}_G} \mu(L) \sum_{L' \in \mathcal{K}_{G'}} \mu(L')\,. \tag{29}$$

Now, let $\{L_s\}_{s \in \mathbb{N}}$ be the sub-sequence of $\{B_i\}_{i \in \mathbb{N}}$ balls corresponding to the sub-collections $\{\mathcal{K}_{G_i} : i \in \mathbb{N}\}$, where the sequence of natural numbers $G_1, G_2, \ldots$ is defined in the same way as within Step 4 of the proof of Proposition 1. For $M \in \mathbb{N}$, let $Q_M := \sum_{i=1}^{M} \#\mathcal{K}_{G_i}$. Then, the same argument used within Step 4 of the proof of Proposition 1, shows that the sequence $\{L_s\}_{s \in \mathbb{N}}$ satisfies the desired properties.

• Step 4: (D) $\Longrightarrow$ (A). By definition, $\limsup_{i \to \infty} L_i \subseteq E_\infty$. Thus, on applying Lemma DBC (the standard divergent Borel–Cantelli Lemma) it follows that

$$\mu\big(E_\infty\big) \geq \mu\Big(\limsup_{i \to \infty} L_i\Big) \geq \kappa^2 > 0\,.$$

2.5. Proof of Theorem 2

The sufficiency side of Theorem 2 is an immediate consequence of the "moreover" part of Lemma LBC. Thus we only have to prove the necessity side. This would clearly follow on mimicking the proof of Proposition 1 if we could establish the analogue of Part (C) from (11) which trivially follows from our working assumption that $\mu(E_\infty) = 1$. Thus, with this in mind, let B be any ball in Ω centred in $\mathrm{supp}\,\mu$. In particular, we have that $\mu(B) > 0$. Let $\mathcal{F} := \{B(x) \subset E_i \cap B : x \in B \cap E_i \cap \mathrm{supp}\,\mu\,,\ i \geq G\}$. In view of the $5r$ covering lemma (Lemma 1), there exists a disjoint sub-family $\mathcal{G}$ such that

$$\bigcup_{L \in \mathcal{F}} L \subset \bigcup_{L \in \mathcal{G}} 5L\,.$$

It follows that

$$B \cap \limsup_{i \to \infty} E_i \cap \operatorname{supp} \mu \subset \bigcup_{L \in \mathcal{F}} L \cap \operatorname{supp} \mu \subset \bigcup_{L \in \mathcal{G}} 5L \cap \operatorname{supp} \mu \,.$$

Hence,

$$\mu \left(\bigcup_{L \in \mathcal{G}} 5L \right) \geq \mu \big(B \cap \limsup_{i \to \infty} E_i \big) \overset{(11)}{=} \mu(B) \,.$$

However, since $\mathcal{G}$ is a disjoint collection of balls centred in $\operatorname{supp} \mu$, we have that

$$\mu \left(\bigcup_{L \in \mathcal{G}} 5L \right) \leq \sum_{L \in \mathcal{G}} \mu \left(5L \right) \overset{(5)}{\leq} \lambda^3 \sum_{L \in \mathcal{G}} \mu \left(L \right) = \lambda^3 \, \mu \left(\overset{\circ}{\bigcup_{L \in \mathcal{G}}} L \right) \,.$$

Thus,

$$\frac{1}{\lambda^3} \, \mu(B) \leq \mu \left(\overset{\circ}{\bigcup_{L \in \mathcal{G}}} L \right) = \sum_{L \in \mathcal{G}} \mu \left(L \right) \leq \mu(B) \,. \tag{30}$$

The sum in (22) is convergent and therefore there exists a finite sub-collection $\mathcal{K}_{G,B} \subset \mathcal{G}$ for which

$$\sum_{L \in \mathcal{K}_{G,B}} \mu \left(L \right) = \mu \left(\overset{\circ}{\bigcup_{L \in \mathcal{K}_{G,B}}} L \right) \geq \frac{1}{2} \frac{1}{\lambda^3} \, \mu(B) \,. \tag{31}$$

In view of (30) and (31) the collection $\mathcal{K}_{G,B}$ satisfies (12) with $\kappa = \frac{1}{2\lambda^3}$. As already mentioned above, to complete the proof we simply replicate Steps 3 & 4 in the proof of Proposition 1. In remains to note that within the sequence of balls arising at Step 4 there may (and most likely will) be finite disjoint collections of balls arising from the same set E_i. These can be grouped together in an obvious manner to form the sequence $(L_{s,B})_{s \in \mathbb{N}}$ as required in the statement of Theorem 2.

3. Examples of applications

In this section we will provide two basic examples showing the conclusions of our results in action. We wish to emphasise that the applications we discuss in this section are not new – they have been chosen to demonstrate the key principles in a relatively simple format. New interesting recent applications can be found, for example, in [23]. We start with an explicit application utilising the power of trimming within a proof of Khintchine's theorem. The proof we provide is not entirely new but, to the best of our knowledge, is simpler, due to some technical simplifications, than the existing published proofs. At the same time it leads to a slightly stronger statement than the standard one.

3.1. The power of trimming: Khintchine's theorem

Let $\psi : \mathbb{N} \to (0, +\infty)$ be a real, positive function. For $q \in \mathbb{N}$, let

$$E_q = E_q(\psi) := \{x \in [0,1] : \|qx\| < \psi(q)\},$$

where $\| \cdot \|$ denotes the distance to the nearest integer, and in turn consider the $\limsup$ set

$$W(\psi) := \limsup_{q \to \infty} E_q \,.$$

For obvious reasons, $W(\psi)$ is usually referred to as the set of ψ-*well approximable numbers.* Khintchine's fundamental theorem [31] in the theory of metric Diophantine approximation dates back to 1924 and it provides an elegant criterion for the 'size' of $W(\psi)$ expressed in terms of one-dimensional Lebesgue measure λ.

Theorem K. *Let* $\psi : \mathbb{N} \to (0, +\infty)$ *be such that* $\psi(q)/q$ *is monotonically decreasing. Then*

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

Remark 6. The above statement of Khintchine's Theorem is in fact a slighter stronger form of the standard modern version [6] in which the $\psi(q)$ is assumed to be monotonically decreasing.

The convergence part of Khintchine's theorem is an immediate consequence of Lemma CBC on noting that $\lambda(E_q) \leq 2\psi(q)$. It does not require the monotonicity assumption or indeed any other additional assumptions. In turn, the modern-days proofs of the divergence part of Khintchine's theorem exploit the principles set out in the main theorems of this paper. For $q \in \mathbb{N}$ and $p \in \mathbb{Z}$ with $0 \leq p \leq q$ define the balls (intervals) in $\mathbb{R}$

$$E_{q,p} = \left\{ x \in [0,1] : \left| x - \frac{p}{q} \right| < \frac{\psi(q)}{q} \right\}\,.$$

Clearly, $E_q = \bigcup_{p=0}^{q} E_{q,p}$ and so $W(\psi)$ is the limsup set of the intervals $E_{q,p}$. In view of Cassels' zero-one law [17], $\lambda(W(\psi)) = 1$ if and only if $\lambda(W(\psi)) > 0$. In turn, by Theorem 3, $\lambda(W(\psi)) > 0$ if and only if there exists a subsequence $(L_i)_i$ of $(E_{q,p})_{q \in \mathbb{N}, 0 \leq p \leq q}$ satisfying (15) and (16). The upshot of this is that establishing Khintchine's theorem boils down to finding the "trimmed" subsequence $(L_i)_{i \in \mathbb{N}}$. This can be done in several ways but probably the easiest is to impose the explicit condition that the rational fractions p/q under consideration are reduced; that is

$$(L_i)_{i \in \mathbb{N}} := (E_{q,p})_{q \in \mathbb{N},\, 1 \leq p \leq q,\, \gcd(p,q)=1}\,.$$

For completeness we present an argument showing the validity of (15) and (16) for this "trimmed" subsequence, a version of which can be found in [46, §I.3].

To verify (15), we start by observing that there are exactly $\varphi(q)$ (the Euler function) positive integers $p \leq q$ such that $\gcd(p, q) = 1$, and therefore we have that

$$\sum_{\substack{p=0 \\ \gcd(p,q)=1}}^{q} \lambda(E_{q,p}) = \frac{2\varphi(q)\psi(q)}{q}. \tag{32}$$

We shall use the following well-known partial summation formula:

$$\sum_{q=1}^{T} a_q b_q = \sum_{q=1}^{T} (a_q - a_{q+1})(b_1 + \cdots + b_q) + a_{T+1}(b_1 + \cdots + b_T)\,,$$

where $(a_q)_{q \in \mathbb{N}}$ and $(b_q)_{q \in \mathbb{N}}$ are any two sequences or real numbers, and the following well known asymptotics for the average order of the Euler's function:

$$\sum_{q=1}^{T} \varphi(q) \sim \frac{3}{\pi^2} T^2 \quad \text{as } T \to \infty\,.$$

Let $0 < C_1 < 3/\pi^2$. Then, using the fact that $\psi(q)/q$ is decreasing, (32) and the trivial estimate $1 + \cdots + q \leq q^2$, by the partial summation formula with $a_q = 2\psi(q)/q$, $b_q = \varphi(q)$, we have that for sufficiently large T

$$\sum_{q=1}^{T} \sum_{\substack{p=0 \\ \gcd(p,q)=1}}^{q} \lambda(E_{q,p}) \geq \sum_{q=1}^{T} (a_q - a_{q+1}) C_1 q^2 + a_{T+1} C_1 T^2$$

$$\geq C_1 \left(\sum_{q=1}^{T} (a_q - a_{q+1})(1 + \cdots + q) + a_{T+1}(1 + \cdots + T) \right).$$

And again by the partial summation formula, this time with $a_q = 2\psi(q)/q$ and $b_q = q$, we get that the above equals $C_1 \sum_{q=1}^{T} q a_q = 2 C_1 \sum_{q=1}^{T} \psi(q)$. Hence,

$$\sum_{q=1}^{T} \sum_{\substack{p=0 \\ \gcd(p,q)=1}}^{q} \lambda(E_{q,p}) \geq 2 C_1 \sum_{q=1}^{T} \psi(q) \tag{33}$$

for sufficiently large T. In particular, this implies (15).

To verify (16), first observe that if $q < m$, $1 \leq p \leq q$ and $1 \leq k \leq m$ then $\lambda(E_{q,p} \cap E_{m,k}) \leq \lambda(E_{m,k}) \leq 2\psi(m)/m$. Then, for fixed $q < m$ we get that

$$\sum_{\substack{p=0 \\ \gcd(p,q)=1}}^{q} \sum_{\substack{k=0 \\ \gcd(k,m)=1}}^{m} \lambda(E_{q,p} \cap E_{m,k}) \leq \frac{2\psi(m)}{m} \times \#\{(k,p) : E_{q,p} \cap E_{m,k} \neq \varnothing\}. \tag{34}$$

Further, if $x \in E_{q,p} \cap E_{m,k}$ then trivially $|qx - p| < \psi(q)$ and $|mx - k| < \psi(m)$, whence

$$|pm - qk| \leq m\psi(q) + q\psi(m) \leq 2m\psi(q).$$

Also, since the fractions p/q and k/m are reduced and different (for we assumed that $m > q$) we must have that $|pm - qk| \geq 1$. Thus the number of (p,k) in the right hand side of (34) is less than or equal to the number of integer points (p,k) satisfying

$$1 \leq p \leq q, \ 1 \leq k \leq m, \ 1 \leq |pm - qk| \leq 2m\psi(q). \tag{35}$$

If all such points (p,k) lie on a line, then from the last inequality of (35) we immediately get that their number is $\leq 4m\psi(q)$. Otherwise, assuming such points exit, the set of these points has rank 2 and, by (35), lies in the convex body given by $|p| \leq q$, $|k| \leq m$, $|pm - qk| \leq 2m\psi(q)$ which has volume $\leq 16m\psi(q)$. In this case, the number of such points is bounded by $32m\psi(q) + 2 \leq 36m\psi(q)$ as a consequence of Blichfeldt's theorem [15]. Either way, the right hand side of (34) is bounded by $72\psi(m)\psi(q)$. Clearly, the same holds when $q > m$. Therefore, in view of the divergence sum condition

$$\sum_{q=1}^{T} \sum_{m=1}^{T} \sum_{\substack{p=0 \\ \gcd(p,q)=1}}^{q} \sum_{\substack{k=0 \\ \gcd(k,m)=1}}^{m} \lambda(E_{q,p} \cap E_{m,k}) \leq 72 \left(\sum_{q=1}^{T} \psi(q) \right)^2 + 2 \sum_{q=1}^{T} \psi(q)$$

$$\leq 73 \left(\sum_{q=1}^{T} \psi(q) \right)^2$$

for sufficiently large T. Together with (33) this verifies (16) with $C = 73/(4C_1^2)$.

Remark 7. The question regarding the relevance of monotonicity in Khintchine's theorem remained a prominent open problem in probabilistic number theory for nearly 80 years. Indeed, in 1941 Duffin & Schaeffer showed that the monotonicity could not be removed (by providing a counterexample) and they formulated an alternative statement. This attracted much work (by Erdös, Vaaler, Pollington, Vaughan and Harman amongst others) and was eventually proved by Koukoulopoulos & Maynard [35]. All these works used trimming as the basis for their approaches very much in line with the outline above. Of course, the process and implementation of trimming are significantly more sophisticated.

Remark 8. The above example makes use of the power of trimming within the context of Theorem 3, a statement dealing with positive measure. In turn, the "ubiquity" technique [6] represents an example of the power of trimming within the context of Theorem 1, a statement dealing with full measure. In short, the theory of ubiquitous systems provides a general framework for deducing full measure statements for a large class of lim sup sets and in view of Theorem 1, it is not at all surprising that "trimming" plays a central role when developing the theory.

Returning to Theorem K, note that the convergence part implies that

$$\lambda\big(W(\tau)\big) = 0 \qquad \text{for any} \quad \tau > 1\,,$$

where for any $\tau > 0$ we 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*. The upshot of the above is that for any $\tau > 1$, the set of τ-well approximable numbers is of measure zero and 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 τ increases. 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.

Let (Ω, d) be a metric space and let X be a subset of Ω. For $\rho > 0$, a countable collection $\{B_i\}$ of balls in Ω of radius $r_i \le \rho$ for each i such that $X \subset \bigcup_i B_i$ is called a *ρ-cover* for X. Let s be a non-negative number and define

$$\mathcal{H}_\rho^s(X) \; := \; \inf\Big\{ \sum_i r_i^s : \{B_i\} \text{ is a } \rho\text{-cover of } X \Big\}\,,$$

where the infimum is taken over all possible ρ-covers of X. The *s-dimensional Hausdorff measure $\mathcal{H}^s(X)$* of X is defined by

$$\mathcal{H}^s(X) := \lim_{\rho \to 0} \mathcal{H}_\rho^s(X) = \sup_{\rho > 0} \mathcal{H}_\rho^s(X)$$

and the *Hausdorff dimension* $\dim X$ of X by

$$\dim_{\mathrm{H}} X := \inf \{s \ge 0 : \mathcal{H}^s(X) = 0\}\,.$$

It is worth emphasising that when s is a positive integer, then $\mathcal{H}^s$ is a constant multiple of Lebesgue measure in $\mathbb{R}^s$. Indeed, when $s = 1$ $\mathcal{H}^s$ is $\frac{1}{2}\lambda$. In particular, $\mathcal{H}^1([0,1]) = \frac{1}{2}\lambda([0,1])$ and it follows from the definition of Hausdorff dimension that

$$\mathcal{H}^s\big([0,1]\big) = \begin{cases} 0 & \text{if} \quad s > 1\,, \\ \infty & \text{if} \quad s < 1\,. \end{cases}$$

For further details concerning Hausdorff measure and dimension see [26,30,37].

The following statement is a Hausdorff measure analogue of Khintchine's Theorem. It provides an elegant criterion for the 'size' of the set $W(\psi)$ expressed in terms of the measure $\mathcal{H}^s$. The convergent part is an immediate consequence of the natural generalisation of Lemma CBC to Hausdorff measures (see for example [13, Lemma 3.10]). As with Khintchine's theorem, the main substance is very much the divergence part.

Theorem K-J. *Let $\psi : \mathbb{N} \to (0, +\infty)$ be such that $\psi(q)/q$ is monotonically decreasing and let $s \in (0, 1]$. Then*

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

Recall, that $\mathcal{H}^1 = \frac{1}{2}\lambda$ and so when $s = 1$ the above reduces to Theorem K. When $s < 1$, the above Hausdorff measure statement is essentially due to Jarník and dates back to 1931. Note that in this case $\mathcal{H}^s([0,1]) = \infty$ and Jarník Theorem (i.e. Theorem K-J with $s < 1$) implies that

$$\dim W(\tau) = \frac{2}{1 + \tau} \qquad (\tau \ge 1).$$

Hence the 'size' of $W(\tau)$ decreases as τ increases which is inline with our intuition. For further details and a gentle introduction to the theory of metric Diophantine approximation see [9].

The second application of our results constitutes the key element of the so-called Mass Transference Principle which enable us to deduce Theorem K-J from Theorem K. At first glance this seems rather odd since Hausdorff measures are regarded as a natural refinement of Lebesgue measure.

3.2. The power of full measure: mass transference principle

The second key example exhibits the power of full measure. To set the scene, let $(\Omega, \mathcal{A}, \mu, d)$ be a locally compact metric measure space equipped with a Borel regular probability measure μ. Without loss of generality we will assume that Ω is the support of μ. With this in mind, suppose there exist constants $\delta > 0$, $0 < a \le 1 \le b < \infty$ and $r_0 > 0$ such that

$$a \, r^\delta \le \mu(B) \le b \, r^\delta \tag{36}$$

for any ball $B = B(x, r)$ with $x \in \Omega$ and radius $r \le r_0$. Such a measure is said to be *Ahlfors δ-regular*. It is well known that if Ω supports an Ahlfors δ-regular measure μ, then $\dim_{\mathrm{H}} \Omega = \delta$ and moreover that μ is strongly equivalent to δ-dimensional Hausdorff measure $\mathcal{H}^\delta$ – see [26,30,37] for details. The latter simply means that there exists a constant $C \ge 1$ such that for every μ-measurable subset E of Ω

$$C^{-1}\mathcal{H}^\delta(E) \le \mu(E) \le C\mathcal{H}^\delta(E)$$

and so (36) is equally valid with μ replaced by $\mathcal{H}^\delta$. Also note that it is easily verified that a δ-Ahlfors regular measure is a doubling measure. Finally, throughout this section, given $s > 0$ and a ball B we define the scaled ball

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

Note, by definition $B^\delta = B$ and if $r < 1$ and $s < \delta$ then B^s is a scaled up ball.

Let $\{B_i\}_{i\in\mathbb{N}}$ be a sequence of balls in Ω with radius $r(B_i) \to 0$ as $i \to \infty$ and suppose that

$$\sum_{i=1}^{\infty} \mu(B_i) < \infty\,.$$

In view of Lemma CBC, it follows that

$$\mu\big(\limsup_{i\to\infty} B_i\big) = 0 = \mathcal{H}^\delta\big(\limsup_{i\to\infty} B_i\big)\,.$$

However, now suppose there exists some $s > 0$ such that the lim sup set associated with the scaled up balls B_i^s has full measure; that is

$$\mu\big(\limsup_{i\to\infty} B_i^s\big) = 1 = \mathcal{H}^\delta\big(\limsup_{i\to\infty} B_i^s\big)\,.$$

It turns out that knowing such a full measure statement for the "scaled up" balls enables us to deduce an analogous statement for the original balls. Indeed, the following Mass Transference Principle [11, Theorem 3] allows us to transfer $\mathcal{H}^\delta$-measure theoretic statements for lim sup subsets of Ω to general $\mathcal{H}^s$-measure theoretic statements.

Theorem MTP. *Let $(\Omega, \mathcal{A}, \mu, d)$ be a locally compact metric measure space equipped with a Borel regular δ-Ahlfors regular probability measure μ supported on Ω. Let $\{B_i\}_{i\in\mathbb{N}}$ a sequence of balls in Ω with radius $r(B_i) \to 0$ as $i \to \infty$. Let $s \geq 0$ and suppose that*

$$\mathcal{H}^\delta\big(\limsup_{i\to\infty} B_i^s\big) = \mathcal{H}^\delta(\Omega) \qquad \text{or equivalently} \qquad \mu\big(\limsup_{i\to\infty} B_i^s\big) = \mu(\Omega).$$

Then,

$$\mathcal{H}^s\big(\limsup_{i\to\infty} B_i\big) = \mathcal{H}^s(\Omega).$$

Remark 9. Note that by the definition of Hausdorff dimension, Theorem MTP implies that $\dim_{\mathrm{H}}(\limsup_{n\to\infty} B_n) \geq s$, and moreover that $\mathcal{H}^s(\limsup_{n\to\infty} B_n) = \infty$ if $s < \delta$.

With reference to Proposition 1, the key towards establishing the Mass Transference Principle is to make use of the fact that the full measure statement (A) implies the existence of the finite sub-collection $\mathcal{K}_{G,B}$ of balls satisfying (C). In [11], this implication is explicitly the subject of Section 4. In short it provides deep information regarding the local distribution of the centres of the balls under consideration. This is very much at the heart of the "optimal" Cantor construction carried out in [11, Section 5] that enables one to show that $\mathcal{H}^s(\limsup_{n\to\infty} B_n) = \infty \;(= \mathcal{H}^s(\Omega))$ if $s < \delta$. The Cantor construction itself is more technical rather than innovative – the existence of the collection $\mathcal{K}_{G,B}$ is the crux!

Remark 10. There have been a steady series of works [1,2,12,32,40,47,48,50] that extend the Mass Transference Principle in numerous directions, such as to systems of linear forms, iterated function schemes and large intersection sets. For an overview of the first ten years after Theorem MTP, we refer the reader to the review article [3]. The more recent work of Wang & Wu [47] is particularly notable in that it deals with lim sup sets defined via rectangles rather than simply balls. It is well worth stressing that all the above cited variants of Theorem MTP have at their heart a common feature. In one form or another, they all exploit the fact that any full measure statement such as (A) in Proposition 1 implies the existence of the finite sub-collection $\mathcal{K}_{G,B}$ of balls satisfying (C).

We bring this section to a close by using Theorem MTP to show that within the world of classical metric Diophantine approximation as described in §3.1, the Lebesgue theory of lim sup sets underpins the general Hausdorff theory. This is rather surprising since the latter theory is regarded to be a subtle refinement of the former.

The claim is that in view of the Mass Transference Principle we have that

$$\text{Khintchine's Theorem} \quad \implies \quad \text{Jarník's Theorem};$$

i.e., Theorem K (which is of course Theorem K-J with $s = 1$) implies Theorem K-J for all $s \in (0, 1)$. First of all let us dispose of the case that $\psi(q)/q \nrightarrow 0$ as $q \to \infty$. Then trivially, $W(\psi) = [0, 1]$ and the result is obvious. Without loss of generality, assume that $\psi(q)/q \to 0$ as $q \to \infty$. With respect to the Mass Transference Principle, let $\Omega = [0, 1]$, d be the supremum norm, $\delta = 1$ and $s \in (0, 1)$. We are given that $\psi(q)/q$ is monotonically decreasing and that $\sum q^{1-s}\psi(q)^s = \infty$. Let $\theta(q) := q^{1-s}\psi(q)^s$. Then it follows that $\theta(q)/q$ is monotonically decreasing and $\sum \theta(q) = \infty$. Thus, Khintchine's Theorem implies that $\mathcal{H}^1(W(\theta)) = \mathcal{H}^1([0, 1])$. It now follows via the Mass Transference Principle that $\mathcal{H}^s(W(\psi)) = \mathcal{H}^s([0, 1]) = \infty$ and this completes the proof of the divergence part of Jarník's Theorem – the main substance of Theorem K-J. As mentioned in §3.1, the convergence part of Theorem K-J is a straight forward consequence of Lemma CBC for Hausdorff measures [13, Lemma 3.10].

Acknowledgments

References

[1] D. Allen, S. Baker, A general mass transference principle, Sel. Math. New Ser. 25 (2019), Art. 39.
[2] D. Allen, V. Beresnevich, A mass transference principle for systems of linear forms and its applications, Compos. Math. 154 (2018) 1014–1047.
[3] D. Allen, S. Troscheit, The mass transference principle: ten years on, in: Horizons of Fractal Geometry and Complex Dimensions, in: Contemporary Mathematics, vol. 731, American Mathematical Society, Providence, RI, 2019, pp. 1–33.
[4] D. Allen, S. Chow, H. Yu, Dyadic approximation in the middle-third Cantor set, preprint, arXiv:2005.09300, 2020.
[5] V. Beresnevich, V. Bernik, M. Dodson, S. Velani, Classical metric Diophantine approximation revisited, in: Analytic Number Theory, Cambridge Univ. Press, Cambridge, 2009, pp. 38–61.
[6] V. Beresnevich, D. Dickinson, S. Velani, Measure theoretic laws for lim sup sets, Mem. Am. Math. Soc. 179 (846) (2006) 1–91.
[7] V. Beresnevich, A. Haynes, S. Velani, Multiplicative zero-one laws and metric number theory, Acta Arith. 160 (2) (2013) 101–114.
[8] V. Beresnevich, A. Haynes, S. Velani, Sums of reciprocals of fractional parts and multiplicative Diophantine approximation, Mem. Am. Math. Soc. 263 (1276) (2020) 1–77.
[9] V. Beresnevich, F. Ramírez, S. Velani, Metric Diophantine approximation: aspects of recent work, in: Dmitry Badziahin, Alex Gorodnik, Norbert Peyerimhoff (Eds.), Dynamics and Analytic Number Theory, in: LMS Lecture Note Series, vol. 437, Cambridge University Press, 2016, pp. 1–95.
[10] V. Beresnevich, S. Velani, A note on zero-one laws in metrical Diophantine approximation, Acta Arith. 133 (4) (2008) 363–374.
[11] V. Beresnevich, S. Velani, A mass transference principle and the Duffin-Schaeffer conjecture for Hausdorff measures, Ann. Math. (2) 164 (2) (2006) 971–992.
[12] V. Beresnevich, S. Velani, Schmidt's theorem, Hausdorff measures, and slicing, Int. Math. Res. Not. (2006) 48794, 1–24.
[13] V. Bernik, M. Dodson, Metric Diophantine Approximation on Manifolds, Cambridge Tracts in Mathematics, vol. 137, Cambridge University Press, Cambridge, 1999, xii+172 pp.
[14] P. Billingsley, Probability and Measure, third edition, Wiley Series in Probability and Mathematical Statistics, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1995.
[15] H.F. Blichfeldt, Notes on geometry of numbers, Bull. Am. Math. Soc. 27 (1921) 150–153.
[16] E. Borel, Lesprobabilités dénombrables et leurs applications arithmétiques, Rend. Circ. Mat. Palermo 27 (1909) 247–271.
[17] J.W.S. Cassels, Some metrical theorems in Diophantine approximation. I, Proc. Camb. Philos. Soc. 46 (1950) 209–218.
[18] S. Chow, Bohr sets and multiplicative Diophantine approximation, Duke Math. J. 167 (9) (2018) 1623–1642.
[19] S. Chow, N. Technau, Higher-rank Bohr sets and multiplicative Diophantine approximation, Compos. Math. 155 (11) (2019) 2214–2233.

[20] S. Chow, N. Technau, Littlewood and Duffin–Schaeffer-type problems in Diophantine approximation, Mem. Am. Math. Soc. (2020), in press, preprint, arXiv:2010.09069.
[21] S. Chow, L. Yang, An effective Ratner equidistribution theorem for multiplicative Diophantine approximation on planar lines, preprint, arXiv:1902.06081, 2019.
[22] K.L. Chung, P. Erdös, On the application of the Borel–Cantelli lemma, Trans. Am. Math. Soc. 72 (1952) 179–186.
[23] É. Daviaud, Extraction of optimal subsequences of sequence of balls, and application to optimality estimates of mass transference principles, preprint, arXiv:2204.01304, 2022.
[24] R.J. Duffin, A.C. Schaeffer, Khintchine's problem in metric Diophantine approximation, Duke Math. J. 8 (1941) 243–255.
[25] P. Erdös, A. Renyi, On Cantor's series with convergent $\sum 1/q_n$, Ann. Univ. Sci. Budapest Sect. Math. 2 (1959) 93–109.
[26] K. Falconer, The Geometry of Fractal Sets, Cambridge Tracts in Mathematics, vol. 85, Cambridge University Press, 1985.
[27] P.X. Gallagher, Approximation by reduced fractions, J. Math. Soc. Jpn. 13 (1961) 342–345.
[28] G. Harman, Metric Number Theory, LMS Monographs New Series, vol. 18, Clarendon Press, 1998.
[29] G. Harman, Variants of the second Borel–Cantelli lemma and their applications in metric number theory, in: Number Theory, in: Trends Math., Birkhäuser, Basel, 2000, pp. 121–140.
[30] J. Heinonen, Lectures on Analysis on Metric Spaces, Universitext, Springer-Verlag, New York, 2001.
[31] A.Ya. Khintchine, Einige Sätze über Kettenbruche, mit Anwendungen auf die Theorie der Diophantischen Approximationen, Math. Ann. 92 (1924) 115–125.
[32] H. Koivusalo, M. Rams, Mass transference principle: from balls to arbitrary shapes, Int. Math. Res. Not. 2021 (8) (2021) 6315–6330.
[33] O. Khalil, M. Luethi, Random walks, spectral gaps, and Khintchine's Theorem on fractals, preprint, arXiv:2101.05797, 2021.
[34] S. Kochen, C. Stone, A note on the Borel–Cantelli lemma, Ill. J. Math. 8 (1964) 248–251.
[35] D. Koukoulopoulos, J. Maynard, On the Duffin-Schaeffer conjecture, Ann. Math. (2) 192 (1) (2020) 251–307.
[36] J. Lamperti, Wiener's test and Markov chains, J. Math. Anal. Appl. 6 (1963) 58–66.
[37] P. Mattila, Geometry of Sets and Measures in Euclidean Spaces, Cambridge Studies in Advanced Mathematics, vol. 44, CUP, 1995.
[38] R.E.A.C. Paley, A. Zygmund, On some series of functions (1) and (2), Proc. Camb. Philol. Soc. 26 (1930) 337–357, 458–474.
[39] R.E.A.C. Paley, A. Zygmund, On some series of functions (3), Proc. Camb. Philol. Soc. 28 (1933) 190–205.
[40] T. Persson, A mass transference principle and sets with large intersections, Real Anal. Exch. 47 (2022) 191–206.
[41] A.D. Pollington, R.C. Vaughan, The k-dimensional Duffin and Schaeffer conjecture, Mathematika 37 (1990) 190–200.
[42] S. Port, Theoretical Probability for Applications, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1994.
[43] F.A. Ramírez, Counterexamples, covering systems, and zero-one laws for inhomogeneous approximation, Int. J. Number Theory 13 (3) (2017) 633–654.
[44] F.A. Ramírez, Khintchine's theorem with random fractions, Mathematika 66 (1) (2020) 178–199.
[45] V.G. Sprindžuk, Problema Malera v Metricheskoǐ Teorii Chisel, Izdat. "Nauka i Tehnika", Minsk, 1967 (in Russian).
[46] V.G. Sprindžuk, Metric Theory of Diophantine Approximations, Translated by R.A. Silverman, Scripta Series in Mathematics, V. H. Winston & Sons, A Halsted Press Book, John Wiley & Sons, Washington, D.C., New York-Toronto, Ont.-London, 1979.
[47] B.-W. Wang, J. Wu, Mass transference principle from rectangles to rectangles in Diophantine approximation, Math. Ann. 381 (2021) 243–317.
[48] B.-W. Wang, J. Wu, J. Wu, Mass transference principle for limsup sets generated by rectangles, Math. Proc. Camb. Philos. Soc. 158 (2015) 419–437.
[49] H. Yu, On the metric theory of multiplicative Diophantine approximation, preprint, arXiv:2010.09004, 2020.
[50] W. Zong, Mass transference principle: from balls to arbitrary shapes: measure theory, J. Math. Anal. Appl. 495 (2021) 1–23.